ABOUT THE JOURNAL

The Berkeley Journal of International Law (BJIL) (ISSN 1085-5718) is edited by students at U.C. Berkeley School of Law. As one of the leading international law journals in the United States, BJIL infuses international legal scholarship and practice with new ideas to address today’s most complex legal challenges. BJIL is committed to publishing high-impact pieces from established and newer scholars likely to advance scholarly and policy debates in international and comparative law. As the center of U.C. Berkeley’s international law community, BJIL hosts professional and social events with students, academics, and practitioners on pressing international legal issues. The Journal also seeks to sustain and strengthen U.C. Berkeley’s international law program and to cultivate critical learning and legal expertise amongst its members.

Website: http://www.berkeleyjournalofinternationallaw.com/
http://scholarship.law.berkeley.edu/bjil/


Subscriptions: To receive electronic notifications of future issues, please send an email to bjil@law.berkeley.edu. To order print copies of the current issue or past issues, contact Journal Publications, The University of California at Berkeley School of Law, University of California, Berkeley, CA 94720. Telephone: (510) 643-6600, Fax: (510) 643-0974, or email JournalPublications@law.berkeley.edu.

Indexes: The Berkeley Journal of International Law is indexed in the Index to Legal Periodicals, Browne Digest for Corporate & Securities Lawyers, Current Law Index, Legal Resource Index, LegalTrac, and PAIS International in Print. Selected articles are available on LexisNexis and Westlaw.

Citation: Cite as Berkeley J. Int’l L.

Submissions: The editors of the Berkeley Journal of International Law invite the submissions of manuscripts. Manuscripts will be accepted with the understanding that their content is unpublished previously. If any part of a paper has been published previously, or is to be published elsewhere, the author must include this information at the time of the submission. Citations should conform to THE BLUEBOOK: A UNIFORM SYSTEM OF CITATION (20th ed.) except where common sense dictates otherwise. Please review our website for additional submissions guidance and send all manuscripts to bjilsubmissions@law.berkeley.edu.

Sponsorship: Individuals and organizations interested in sponsoring the Berkeley Journal of International Law should contact the Editors-in-Chief at bjil@law.berkeley.edu.

Copyright: Copyright for material published in the Journal is held by the University of California Regents except where otherwise noted.
ARTICLES

BETWEEN THE KINGDOM AND THE DESERT SUN: HUMAN RIGHTS, IMMIGRATION, AND BORDER WALLS
Moria Paz

CHINESE RECESSION AND TRANSPLANTATION OF WESTERN CONTRACT LAW
Wang Jingen and Larry A. DiMatteo

RECOVERING LOST TAX REVENUE THROUGH TAXATION OF TRANSNATIONAL HOUSEHOLDS
Ariel Stevenson

A COMITY OF ERRORS: THE RISE, FALL, AND RETURN OF INTERNATIONAL COMITY IN TRANSACTIONAL DISCOVERY
Diego Zambrano

COOPERATING ALONE: THE GLOBAL REACH OF U.S. REGULATIONS ON CONFLICT MINERALS
Remi Moncel
BOARD OF EDITORS

Editors-in-Chief
AVERY BROWN
VICTORIA HARGIS

Managing Editor
MARSHA RUCH

Senior Publishing Editors
MARI SAHAKYAN
MARISSA RHoades

Submissions Editors
RAY Duer
GUILHERME DURAES
BRIAN HALL
AARON MURPHY

Senior Articles Editors
JUAN ARAGON
LILY DOBSON
EMILY SHULMAN

Supervising Editors
SAYA WALLACE
LAUREN ASSAF

Operations Director
NEELAM MOHAMMED

Online Editors
KELSEY QUIGLEY
AARON VOIT
TANIA SWEIS

Web Editor
JANE PENNEBAKER

Book Review Editor
CHRISTOPHER CASEY

Senior Executive Editor
ALEXANDRA BRANDT

Executive Editor
DANIEL LOEVINSOHN

Symposium Editors
FRIDA ALIM
CHELSEA GEIGER
AYLIN KUZUCAN
MARISSA RHoades
ASHIK SHAH

Article Editors
NATE BERRY
RACHEL BOSLEY
TARA BRAILEY

ANDREA ORTEGA
MARI SAHAKyAN
REMI SALTER
CHRISTINA SHIN

OLIVIER SINONCELLI
GINNY SMITH
CHRIS YANDELL

Assistant Editors
EMMANUELINE BERDUGO
ELLEN CHOI
RACHEL FLEIG-GOLDSnTAIN

HEIDI LII
SARAH MIRZA

SEAN PINCKnEY
WHITNEY TOLAr
CHELSEA TnERNER

http://scholarship.law.berkeley.edu/bjil/vol34/iss1/1
Members

PETER ANDREWS
ALFREDO DIAZ
MATT GAUTHIER
NISHA GIRIDHAR
SARAH GODWIN
GRACE GOHLKE

SARAH HUNTER
HANNAH KHALIFEH
LIZ KING
JENNA KLEIN

DANIEL LUECKE
DAVID NAHIMIAS
NAM PHAN
JOSE RAMOS
IOSIF SOROKIN
AYN WOODWARD

Faculty Advisor

MARC HOFFMAN
KATE JAstrom
Between the Kingdom and the Desert Sun: 
Human Rights, Immigration, and Border 
Walls

Moria Paz*

ABSTRACT

A peculiar construction boom is in progress worldwide: border walls are being installed by wealthy countries at an unprecedented rate in order to control unwanted immigration by poor people. This Article asks why, almost a quarter of a century after the Iron Curtain came down, the walls are now going up again. It suggests a provocative answer: these separation barriers are a logical response by States to the way in which human rights law has been enforced in cases bearing on immigration. In other words, and counter-intuitively, the recent boom in border wall construction signals the success of the human rights tradition, rather than its failure to establish an alternative to territorial sovereignty.

At the same time, this Article also uses the case study of walls to make a larger point on the intractability of the human rights regime that bears on immigration. Building on a systematic analysis of jurisprudence, I argue that human rights courts and quasi-judicial bodies utilize an arbitrary category—territory—to balance the policy interests of the individual non-national and the State. The result is essentially random from the perspective of both of these stakeholders. Walls make concrete a perverse side effect of this compromise: because the regime conflates access with territory, it disproportionately rewards strong young men who already have sufficient capacity (in age, gender, or resources) to scale the barrier, even if their predicament may not actually call for protection. But it privileges them only after they have risked themselves, and if they survive that risk at all. And so, at least when it comes to immigration, the

DOI: http://dx.doi.org/10.15779/Z380P11

*  Moria Paz is a fellow at Stanford Law School. She is thankful to Gabriella Blum, William Forbath, Martti Koskenniemi, David Luban, Itamar Mann, Samuel Moyn, Jaya Ramji-Nogales, Daria Roithmayr, Paul K. Saint-Amour, Hila Shamir, Nomi Stolzenberg, Markus Wagner and Ralph Wilde for reading earlier drafts. In particular, she is grateful to Barbara Fried, Mark Kelman, Amalia Kessler, and Peter Kozodoy for conversation as she developed the Article. She also wants to thank Rich Porter and Sergio Stone from SLS library for invaluable help with the research.
human rights regime operates in effect as a natural selection mechanism. This is fundamentally unstable and unjust.

There is no larger eternity than a door marked: closed today.
Closed forever; no one’s opening it, no one’s coming.
There are no clouds in the sky. Accept the verdict; sign.
No one’s opening. Go home, dream on.

(Yehuda Amichai1)

“The distribution of membership is not pervasively subject to the constraints of justice. . . . [S]tates are simply free to take in strangers (or not) . . . . [T]he right to choose an admissions policy . . . is not merely a matter of acting in the world, exercising sovereignty, and pursuing national interests. At stake here is the shape of the community that acts in the world, exercises sovereignty, and so on. Admission and exclusion are at the core of communal independence. They suggest the deepest meaning of self-determination.”

(Michael Walzer)2

Abstract ................................ ................................ ................................ ..................1
Introduction ................................ ................................ ................................ ...........2
I. Back Door Strategies of Immigration Control ................................ ................10
II. Front Door Strategies of Immigration Control ................................ ...............21
   A. First Method—Adopt the Universalist Tradition. ...........................26
   B. Second Method—Adopt the Exclusionist Tradition. ......................29
   C. Third Method—Adopt the Compromise Approach that Human
    Rights Courts Tailored Between Universality and Exclusion
    and that is Structured Through Territory ................................ ........32
III. A Case Study: The Israel-Egypt Wall ................................ ...........................34
Conclusion ................................ ................................ ................................ ...........39

INTRODUCTION

Sometimes walls become the world all around. Take the razor wire fence that Spain built in North Africa around its enclave Melilla that borders Morocco. Spain installed the physical barrier to close itself—and Europe—off from Africa. On one day in May, around 1,000 Sub-Saharan and Syrian migrants rushed this wall, seeking to cross.3 They devoted many months to preparing for their attack on the fence, including studying the movements of the guards and accumulating specialized gear, such as hooks to attach to their wrists and screws to stick to their shoes for a better grip.4 They coordinated D-day-style mass

2.  MICHAEL WALZER, SPHERES OF JUSTICE: A DEFENSE OF PLURALISM AND EQUALITY 61–
4.  Id.
attempts on the wall, seeking to overwhelm guards so that some might make it across uncaught, or to topple a section of the fence by their sheer weight. And one attack came after another; in fact, in that same month, May 2014, there were three mass attempts on the barrier involving 1,000 to 2,000 people each.5 These numbers, though massive, are still only a small fraction of the 80,000 individuals that by the middle of 2014 were already approaching the fence.6

Spain, in turn, spends considerable resources to stop these men. It employs nearly 1,000 police and Guardia civil officers to guard these fences,7 making this border “one of the most closely guarded borders in the EU,”8 and has already announced that it is planning to deploy more.9 It equips the barrier with motion sensors, cameras, and watchtowers, and patrols by car and helicopter.10 More recently, Spain has resorted to live ammunition to deter men from scaling the fences,11 and, in cooperation with Morocco, it is now also building an extra ditch and fence, crowned with concertina wire, about 500 meters from the existing Spanish fences.12

This border war zone is far from unique. Border walls like the one in Melilla, that are (i) substantially designed to block illegal immigration, and (ii) constructed on undisputed State territory,13 are quickly multiplying around us. Along with the physical barrier in Melilla (10.5 kilometers of border), Spain has also installed another six-meter-high double fence around its second land border

5. Id.
6. Raphael Minder, Spain Struggles to Halt Migrants at Two Enclaves, N.Y. TIMES, Mar. 6, 2014 (the majority are Sub-Saharan Africans, but more recently, they were joined by Syrians fleeing their country) [hereinafter Minder, Spain Struggles]; see Europe’s Huddled Masses: Rich Countries Must Take on More of the Migration Burden, ECONOMIST, Aug. 16, 2014.
8. Id.
11. See id.
13. See Save Our Heritage Org. v. Gonzales, 533 F. Supp. 2d 58, 61 (D.D.C. 2008) (this definition includes barriers such as the U.S.-Mexico wall aimed at deterring “illegal crossings in areas of high illegal entry”); HCJ 7957/04 Mara’abe v. Prime Minister of Israel 60(2) PD 57–58 [June 21, 2005] (Isr.) (discussing the decision-making process to construct the separation barrier and the process of land seizure); HCJ 2056/04 Be’it Sourik Village Council v. The Government of Israel 58(5) PD 16–17 [2004] (Isr.) (It also excludes walls built for security purposes. Here, again the example is the Israeli Security Fence, that, according to the Israeli High Court’s definition, is built to enhance security and is “motivated by security concerns” that do not “express a political border, or any other border.”); Legal Consequences of the Construction of a Wall in the Occupied Palestinian Territory, Advisory Opinion, 2004 I.C.J. 136, ¶ 121 (July 9) (but it excludes walls that are constructed for political purposes such as the Israeli Security Fence that, according to the International Court of Justice, is built to achieve “de facto annexation”); see also Streletz v. Germany, 2001-II Eur. Ct. H.R. 409 ¶ 69 (similarly, this definition also excludes walls like the Berlin Wall that are designed to stop emigration and to “staunch the endless flow of fugitives”).
in North Morocco, Ceuta (7.8 kilometers long). These two fences are higher than the Berlin Wall and cost thirty million euros. According to Spain’s Interior Minister, the goal of both walls is to “impede anyone from climbing” and to deter the thousands of migrants who arrive at the borders of its two enclaves. In 2006, only a year after Spain reinforced its fences, the United States began constructing its own massive 1,100 kilometer, double-layer fence between El Paso and Ciudad Juarez, and between San Diego and Tijuana, at a price tag of $21 million per mile. In signing a bill into law to construct these barriers, U.S. President George W. Bush declared that there had been an increase in illegal immigration and that the fence “is an important step” among several “to secure our borders.” Israel has also installed an immigration wall. In 2010, the country began to construct a 245-mile-long, five-meter-high fence (twice the height of the Israeli Security Fence, the separation barrier built by Israel in the West Bank). This “monster of a fence” stretches almost the entire border between Israel and Egypt and cost $450 million dollars to build, making it one of the largest projects in Israel’s history.

14. For details of these two walls, see Lisa-Maria Leipersberger, The European Hard Borders, MIGRABLOG (Feb. 5, 2015), https://migrablog.wordpress.com/2015/02/05/the-european-hard-borders/.


17. Ashifa Kassam, supra note 9.


22. Gidon Ben-zvi, supra note 20. When thinking about walls and Israel, the Israeli Security Fence immediately comes to mind. The wall that I discuss here, on the Israel-Egypt border, is surprisingly under-researched.

23. Barak Ravid, Israel to Build NIS 1.5b Fence Along Egypt Border, HAARETZ, Jan. 10, 2010. In 2010, Benjamin Netanyahu, explained that he is ordering the construction of the fence in order to keep out African asylum seekers he claims are threatening the country’s Jewish character:
Greece joined. In 2011, it began building the 12.89 kilometer “Evros Wall” on the land border it shares with Turkey. The cash-strapped country completed the almost $10 million barbed wire fence in less than a year without EU support.24 The objective, said the Greek Minister of the Interior, is that “no illegal migrant will be left in the country.”25 And now others are joining. Bulgaria is erecting a thirty-three kilometer, four-meter-high wall on its rugged Turkish border to “prevent[] the illegal crossing of the border,”26 while Hungary is building a wall to secure its 177 kilometers on the border with Serbia. “This is a necessary step,” said the Hungarian government’s spokesman, adding “[w]e need to stop the flood.”27 Finally, more walls are coming: Austria recently declared that it would build a wall along its border with Slovenia to “control the refugees in an orderly way,”28 while the Slovenian Prime Minister announced that “[i]f necessary, we are ready to put up [a] fence immediately.”29

But despite the rapid increase in wall construction, and notwithstanding the mounting brutality surrounding them, the international legal community still has not decided how to treat these walls as legal objects. Surprisingly, border walls are under-researched in international legal scholarship, including in international law, human rights law, and refugee law.30

he made “a strategic decision to secure Israel’s Jewish and democratic character.” Id. Only later, in 2011, after the trampling of the Mubarak regime in Egypt, a security function was also added to the wall. See Shuki Sadeh, The Money Fence, MARKER, Nov. 12, 2011 (original source in Hebrew).


26. Stoyan Nenov, Bulgaria’s Fence to Stop Migrants on Turkey Border Nears Completion, REUTERS, July 17, 2014.

27. Patrick Kingsley, Migrants on Hungary’s Border Fence: ‘This Wall, We Will Not Accept It’, GUARDIAN, June 22, 2015. Turkey announced that it too is building a wall—this one to secure its 900 kilometer southeastern border with Syria. The wall, the Turkish Interior Ministry made clear, is being built both for “for security reasons,” and “to curb smuggling and illegal crossings.” Suriye Sınırına Seyyar Duvar [Syria Border to the Mobile Wall], RADIKAL, Apr. 27, 2014 (Turk.); see also Dasha Afanasieva, Turkey Builds Wall in Token Effort to Secure Border with Syria, REUTERS, May 5, 2014.


29. Id.

30. In human rights law and international law, by far most of the scholarly attention to walls is given to the Berlin Wall and the Israeli Security Fence. An important exception is scholarship out of the University of Texas at Austin Law School that deals mainly with the U.S.-Mexico wall, but also with other walls. See, e.g., Denise Gilman, Seeking Breaches in the Wall: An International Human Rights Law Challenge to the Texas-Mexico Border Wall, 46 TEX. INT’L L.J. 257 (2011); see also Yishai Blank, Legalizing the Barrier: The Legality and Materiality of the Israel/Palestine Separation Barrier, 46 TEX. INT’L L.J. 310–11 (2011) [hereinafter Blank, Legalizing the Barrier] (focusing mainly on the Israeli Security Fence, but also border walls in general); Marta Tavares, Fencing Out the Neighbors: Legal Implications of the U.S.-Mexico Border Security Fence, 14 HUM. RTS. BRIEF 33 (2007). Similarly, in refugee law there is very little discussion of walls as immigration exclusion modes. In fact, prominent scholars do not mention walls. See, inter alia, ROSEMARY BYRNE ET AL., NEW ASYLUM COUNTRIES? MIGRATION CONTROL AND REFUGEE
In this Article, I work through the unstable, uncertain international legal ontology of these border walls. I suggest that they reflect a disappointing story about human rights law: at least when it comes to immigration, the regime is both inherently arbitrary and fundamentally unjust. To tell this story of disillusionment, I begin with a familiar tension.

When human rights courts and quasi-judicial bodies decide cases that bear on immigration control, they can choose between two, often competing, doctrinal traditions. The first is a universal framework that views human rights as inherent in the individual, whether or not the individual complied with formal conditions for immigration. In this approach, the human rights of non-nationals may impose substantive constraints on the State’s discretion to expel them. The second is an exclusionist international legal regime that gives the State sole authority to decide who may enter its domain, under what conditions, and with what legal consequences. Here, strangers who reach a State’s shores have no claim to rights that the State does not willingly grant.31

These traditions represent two prevailing normative outlooks and descriptions of behavior that conflict with one another: if individuals have certain basic rights because they are human, then, at least under certain

31. For a detailed analysis of this tension, including detailed survey of relevant treaty law, review of the writing of legal scholars and philosophers, see Chantal Thomas, *Convergences and Divergences in International Legal Norms on Migrant Labor*, 32 COMP. LAB. L. & POL’Y J. 405 (2011), and Chantal Thomas, *What Does the Emerging International Law of Migration Mean for Sovereignty*, 14 MELB. J. INT’L L. 1, 4 (2013). For the roots of this tension, see IMMANUEL KANT, *PERPETUAL PEACE* 21 (FQ Classics 2007) (1795) and the right to temporary sojourn (“It is not the right to be a permanent visitor that one may demand. A special beneficent agreement would be needed in order to give an outsider a right to become a fellow inhabitant for a certain length of time. It is only a right of temporary sojourn, a right to associate, which all men have.”). For a typology of positions on this, see MARTTI KOSKENNIEMI, FROM APOLOGY TO UTOPIA: THE STRUCTURE OF INTERNATIONAL LEGAL ARGUMENT 184–85 (1989) (laying out four variants of the combination of the normative and the concrete in international law: the “rule-approach” emphasizing power politics; the “policy-approach” that sees all (governmental or non-governmental) global processes as part of international law; the “idealistic position”; and the “skeptical position”).
circumstances, non-nationals can enter or remain in a State outside norms of the State’s sovereign interests. If, however, the State has absolute power over who belongs in the national community, then non-nationals are not allowed to enter or remain in the State without government consent. Starting from this tension between universality and exclusion (or human rights and sovereignty), I make three claims.

First, in the past ten years, moving from case to case, human rights courts and quasi-judicial bodies have worked out a compromise between the two legal traditions that is biased in favor of human rights. In particular, they read norms more strictly and more absolutely, and develop substantive standards of protection beyond the five grounds specified under the Refugee Convention: race, religion, nationality, membership in a particular social group, or political opinion. But these enforcement bodies stop short of going all the way in the direction of universality; they do not suggest the extreme step of open borders (one’s place of birth is irrelevant in the exercise of rights). Instead, they constrain the reach of increasingly expansive human rights protections by linking jurisdiction to variants of what I term ‘physicality’: human rights courts and quasi-judicial bodies bootstrap expansive rights on either establishing territorial presence in the host State (jurisdiction grounded in territory) or coming within the effective control of the State or its agents (jurisdiction grounded in contact). To be protected, then, an individual must get close to the State or its agents.

Paradoxically, therefore, this human-rights-leaning compromise has ended up reinforcing territoriality and thus also the exclusionist (statist) tradition. Because courts and quasi-judicial bodies attach access to territorial presence, every time that they enforce human rights, despite their emphasis on universality, they re-consecrate the centrality of territory. And so, more human rights also means more exclusion.

Second, border walls are a predictable strategic response by States that seek to regain exclusion capabilities, reacting to the way in which human rights courts and quasi-judicial bodies balance the tension between universality and exclusion. In other words, the recent boom in border wall construction may

32. Convention Relating to the Status of Refugees art. 1(A)(2), Apr. 22, 1954, 189 U.N.T.S. 150 [hereinafter Refugees Convention]. I am focusing here only on legal obligations. I do not discuss emerging obligations around burden sharing such as, for example, obligations to promote economic justice either in the form of donation to United Nations Human Rights Council, contributions by rich donor States to enhance welfare in developing States, for example the Millennium Development Goals, or burden sharing with countries neighboring those in crisis (including things such as resettlement, aid in supplying sanitary, education, housing and other facilities, etc.).

33. An example of jurisdiction grounded in contact is interdiction on the high seas; for further discussion, see infra, Part III.

34. For discussion, see infra, Part II.

35. My claim here is not causal. A causal analysis requires extensive empirical data to account for the real efficacy of, first, human rights courts’ decisions on actually shaping States’ immigration policy on the ground, and, second, of border walls to control immigration in practice. But this is well
signal the success of the human rights tradition, rather than its failure to establish an alternative to territorial sovereignty. The compromise that courts created assigns human rights protection only after would-be migrants and asylum seekers enter the territory of the host State (strong territoriality) or come within its effective control (neo-territoriality). But this balance impels States to tighten their borders to prevent a territorially-based human rights regime from being triggered by border crossing: in other words, the States’ protective duties can be avoided if there is no one to protect.

Walls are not the only possible mode of deterring access left for States to utilize after international human rights courts have made it more difficult to exercise traditional exclusion authority. But they do present a unique challenge for regulation. To begin, a wall is a relatively passive interdiction method: it does not require extensive State agency after its initial construction. In addition, a wall is installed on the very border of the State and physically reinforces that boundary: it simply marks the border that was always there. Borders, in turn, are central to the operation of the larger international legal and political regime. Thus, a legal attack on such a wall also calls into question the larger international bargain. I use the treatment by the Israeli Supreme Court of the Israel/Egypt fence to demonstrate how the conflation of walls and borders made these walls legally permissible \textit{ex ante}. This case study comes from a national, not an international court, but the Court is interpreting international law and human rights law. This jurisprudence suggests the inherent challenge in regulating a wall that a State erects on its own territory through an international legal system that squares sovereignty with territorial exclusivity.

Third, even if each court decision is locally sensible, the human-rights-leaning compromise that courts and quasi-judicial bodies have ultimately produced is senseless. To begin, the compromise does not serve the policy interests of either the individual or the State. Human rights enforcement bodies determine jurisdiction by the territorial location of the plaintiff: whether she was able to get into the State or close enough to establish contact with its agents. beyond the scope of the purely legal analysis provided in this Article. Instead my claim is narrower: I argue that walls are a logical answer to the compromise that courts worked out.

36. For a different answer to the question of why walls are being built now, see supra note 30. In a fascinating argument, Brown suggests that walls are built as the symbols of sovereignty at the time of its definitive waning. Walls, she explains, are built to assert identity and to establish the “us” (with purity and integrity) against the “them” on the outside. While these walls are efficacious in drawing the “we”—who’s in, who’s out—they are not actually effective in re-establishing sovereignty in practice.


38. See infra text accompanying notes discussion in pages 32-35.

39. \textit{Id.}
Territory here is the only category that matters. Alas territory is a normatively arbitrary category; it is random from the perspective of both the individual non-national and the State. Territory does not prioritize the substantive needs of the individual, but instead privileges ability to get close to the State or its agents—proximity that is determined by capacity (or special circumstances like luck, resources, gender, or physical traits such as youth, strength, and stamina). And, at the same time, territory is equally arbitrary from the perspective of the State: States with borders that are more accessible, or with neighbors that happen to suffer economic, political, or environmental crises, are punished regardless of the State’s real constraints and efforts to deal with the inflow of immigration.

In addition, the compromise cannot be normatively justified. Walls highlight the moral intractability of this territorially-based settlement. The ability of walls to restrict movement—thus also access to human rights—depends on how courts regulate them. At the moment, border walls erected as an immigration control policy remain relatively unregulated in human rights law. But it seems clear that courts will have to address the problem of border walls—and in the not-too-distant future. So what are courts likely to do? Building on existing precedents from both human rights and international law, I map three possible approaches that a human rights court or quasi-judicial body can take in adjudicating such a border wall. Each of these methods works out a different compromise to the fundamental tension between putting the universalist (human rights) or exclusionist (statist) frame at the center of immigration control, and each correlates to a different vision of sovereignty and borders in international law.

First, *adopt the universalist tradition*: a State owes protective duties on both sides of the wall. In this approach, the wall acts as a bridge: establishing contact with the wall is tantamount to getting inside the State. Jurisdiction here is grounded in proximity to a wall. Second, *adopt the exclusionist tradition*: a State accrues protective duties only upon initial entrance to its territory. Now the wall acts as a final barrier: getting close to the wall does not entail rights. Jurisdiction is aligned with territory. Third, and also the existing status quo—*adopt the territorially-based compromise that courts institutionalized between universality and exclusion*: a State carries thin procedural duties on the external side of the wall. But after gaining entrance, it bears significant protective responsibilities outside its consent. This time, the wall, a physical barrier, becomes the essence of human rights protection. Proximity by itself no longer denotes rights.

Because these approaches to the regulation of a wall offer three different resolutions to the same problem (the tension between universality and exclusion), the choice between them highlights the values of human rights courts. But, I suggest, none of the three approaches, when taken to their logical conclusions, can be normatively defended. Which leaves us at a normative dead-end.

Furthermore, the existing status quo itself leads to a perverse side effect. Under the compromise approach, an individual’s location vis-à-vis the wall
makes all the difference in the allocation of rights and duties (or lack thereof). This compels States to continually reinforce their walls, and to build additional layers of walls, to prevent would-be immigrants and asylum seekers from getting close enough to trigger territory-based human rights protections. At the same time, it also invites individuals to resort to ever more hazardous behavior to scale the walls that States construct. Consequently, the regime ends up protecting disproportionately those individuals who are mentally willing to assume serious risks and whose bodies are physically able to make the arduous attempt. Meaning, it protects only young men. But they receive this protection only if they risk themselves and are lucky enough to survive the ordeal. And, at the same time, this order also leaves too many non-nationals that human rights courts are committed to protect with no mechanism to access asylum rights. Ironically, this nonsensical result is due to the insistent actions of human rights courts and other quasi-judicial bodies to expand access to human rights. The path out of the desert and into the kingdom may be paved with good intentions, but is also barred with formidable walls.

I.

BACK DOOR STRATEGIES OF IMMIGRATION CONTROL

Over the past ten years, when adjudicating cases that bear on immigration, human rights courts and quasi-judicial institutions have worked out a compromise between universality and exclusion that brings down the gavel in favor of universality. They restrict States’ prerogatives to expel non-nationals out of what I refer to as the ‘back door,’ i.e. deporting them after they have already arrived inside the country illegally. At times, human rights enforcement institutions categorically ban deportation; at other times, they make it more difficult for the host State to deport non-nationals. Courts do so by conflating access and territory: they simultaneously expand substantive standards of protection and bootstrap protection on the fact of territorial presence. A non-national, therefore, has to reach the territory of the host State in order to trigger protection. Because courts condition human rights jurisdiction on physicality

40. Traditionally, immigration was outside the scope of human rights law. For example, the Universal Declaration of Human Rights (the Declaration) grants every individual the right to leave any country, including the immigrant’s native country. Universal Declaration of Human Rights, G.A. Res. 217 (III) A, art. 13, U.N. Doc. A/RES/3/217 A (Dec. 10, 1948). The Declaration only guarantees the right to enter one’s own country. Similarly, the International Covenant on Civil and Political Rights also guarantees every individual the right only to “enter his own country” but not the right to enter one’s country of choice. International Covenant on Civil and Political Rights art. 12, Mar. 23, 1976, 999 U.N.T.S. 171 [hereinafter ICCPR]. For a historical review of the application of a human rights framework to immigration, see Ruth Gavison, Immigration and the Human Rights Discourse: The Universality of Human Rights and the Relevance of States and of Numbers, 43 ISR. L. Rev. 26–28 (2010).

41. An inside/outside distinction is also familiar from the United States: non-nationals who are deemed to have entered U.S. territory are entitled to procedural due process, while aliens outside (or deemed to be outside) are not so entitled. See, e.g., Shaughnessy v. United States ex rel. Mezei, 345 U.S. 206 (1953); United States ex rel. Knauff v. Shaughnessy, 338 U.S. 537 (1950); but see
grounded in territory, greater access means greater application of territoriality. Increased access to territory-based rights, therefore, paradoxically, also reinforces the exclusionist tradition.

This finding emerges out of my systematic examination of cases before the United Nations Human Rights Committee (UNHRC), and the European Court of Human Rights (ECtHR). I analyzed the way in which these institutions disposed of cases bearing on two types of rights: (i) the right for family unity and private life, and (ii) the right not to be subject to torture or to cruel, inhuman punishment.

I selected these two adjudicative institutions because they are the most significant international human rights enforcement bodies operating today. Both also create entitlements that give private rights of action to the individuals claiming them, and, through individual case adjudication, produce decisions that are of general application. In addition, the UNHRC is the only active human rights complaints body with a “potentially universal reach,” and provides a window into the working of the United Nations in matters of immigration. The ECtHR, in turn, not only developed the most extensive case law on the rights of non-nationals, but also enjoys compulsory jurisdiction such that its case law is informally binding on all the parties that have signed and ratified the ECtHR.

I selected these two rights because they are the rights most commonly considered in the immigration setting and in particular the context of


42. The jurisdiction of the UNHRC has become “a key component in the human rights movement.” RUTH MACKENZIE ET AL., THE MANUAL ON INTERNATIONAL COURTS AND TRIBUNALS 427 (2d ed. 2010). The ECtHR, in turn, is considered “a success story,” id. at 356, and “has become a source of authoritative pronouncements on human rights law for national courts that are not directly subject to its authority.” ANNE-MARIE SLAUGHTER, A NEW WORLD ORDER 80 (2004).

43. The UNHRC adjudications are not binding on States, but are highly significant recommendations. In addition, the UNHRC is empowered to entertain individual complaints only under the Optional Protocol (which means that the State must consent to its jurisdiction). This Protocol has 114 States-parties, and the United States and Israel, which are discussed in more detail later, are not part of them. For more on the working on the UNHRC, see MACKENZIE ET AL., supra note 42, at 415–31.

44. As of 2010, the number of State-parties to the Optional Protocol that have accepted the jurisdiction of the UNHRC to receive individual communications was more than double the number subject to the jurisdiction of any regional courts. MACKENZIE ET AL., supra note 42, at 426–27.

45. But note that decisions of the UNHRC are more expressively political and have a weaker compliance pull as compared to those of the ECtHR, whose jurisdiction is binding.

expulsion.\textsuperscript{47} In addition, these two rights cover the range of rights at issue, from the ‘lighter’ right of an individual to have a family to the ‘heavier’ entitlement not to be tortured. The cases also document the breadth of applicants’ plights, from the applicant who broke the law in entering the host State with a hope of improving her life, to the one who fled her home country at gunpoint. I examined all of the communications and cases dealing with these two rights in the context of immigration control that reached the UNHRC and the ECtHR from the inception of the institutions until January 2014. In total, I surveyed a little short of 150 communications and cases.

Let us start with the right to family unity and private life found in Articles 17 and 23 of the International Covenant on Civil and Political Rights (ICCPR) and Article 8 of the Convention for the Protection of Human Rights and Fundamental Freedoms (ECHR).\textsuperscript{48}

First, under prior decisions of both the UNHRC and the ECtHR, it had been settled that while the right to a family and private life can constrain the back door option of deporting non-nationals, the State’s interest in security and public order outweighed the interest of the individual in family life.\textsuperscript{49} Case law of the last ten years, however, has brought this into question. Specifically, the cases address whether the right to family life for non-nationals who were convicted of crimes trumps the State’s right to security and public order in situations where reunion abroad between the applicant and his family is either not possible or could not be reasonably expected.

Thus, in \textit{Francesco Madafferi v. Australia},\textsuperscript{50} the UNHRC told Australia that the decision to deny a permanent visa for an author without a lawful status\textsuperscript{51} and who was of “bad character”\textsuperscript{52} (a judgment stemming from criminal acts committed in the home country) constituted arbitrary interference with family life. The reasons for removal, the decision read, were not sufficiently pressing, and the removal would have imposed “considerable hardship” on the author’s family (Madafferi had been married for fourteen years to his wife, an Australian

\textsuperscript{47} For discussion on family life, see for example, Immigration Act, 2014, c. 22 (U.K.).

\textsuperscript{48} For the UNHRC, see ICCPR art. 17, Mar. 23, 1976, 999 U.N.T.S. 171 (“No one shall be subjected to arbitrary or unlawful interference with his privacy, family, home or correspondence . . . .”) and id. art. 23 (“1. The family is the natural and fundamental group unit of society and is entitled to protection by society and the State. 2. The right of men and women of marriageable age to marry and to found a family shall be recognized.”). For the ECtHR, see Convention for the Protection of Human Rights and Fundamental Freedoms art. 8, Nov. 4, 1950, 213 U.N.T.S. 221 (“Everyone has the right to respect for his private and family life, his home and his correspondence.”) [hereinafter ECHR].


\textsuperscript{51} Id. ¶ 2.2. Madafferi arrived in Australia on a tourist visa and stayed there after his visa expired. Id. ¶ 2.1. He later applied for a spouse visa but his application was denied because of his prior conviction and outstanding prison sentence in Italy. Id. ¶¶ 2.3–2.4.

\textsuperscript{52} Id. ¶ 2.4.
national, and had four children). Instead the UNHRC ordered Australia to process a spouse visa for Madafferi with an eye to the State’s obligation to protect his minor children.

The ECtHR goes even further than the UNHRC. Madafferi had mitigating circumstances: he had a viable path to a lawful immigration status in Australia, his sentences in Italy had been extinguished, and there was no outstanding warrant for his arrest. The Strasbourg Court, however, was willing to reverse the expulsion order of applicants in a different case who had participated in serious crimes, even when the court acknowledged that it was not, in fact, “impossible for the spouse and the applicant’s children to live” in the applicant’s country of citizenship, but merely that doing so would “cause them obvious and serious difficulties.”

---

53. Id. ¶ 9.8.

54. Id. ¶ 11. For contrast, in several cases before Madafferi, the UNHRC found no violation of the right to family life in deporting lawful permanent residents who had criminal convictions. See Jaya Ramji-Nogales, Undocumented Migrants and the Failures of Universal Individualism, 47 VANDERBILT J. TRANSNAT’L L. 699, 736 n.161 (2014). More recently, in Fernandes v. Netherlands, Communication No. 1513/2006, U.N. Hum. Rts. Comm., 93d Sess., July 7–25, 2008, ¶ 1, U.N. Doc. CCPR/C/93/D/1513/2006 (Aug. 6, 2008), the UNHRC considered an application of two undocumented immigrants who were the parents of four children (three of them Dutch citizens). The case reached the court after the father’s application for a residence permit was rejected due to a criminal record. The Committee found their claim insufficiently substantiated and therefore inadmissible. Id. ¶¶ 2.3–2.5, 6.3.


56. Amrollahi v. Denmark, App. No. 56811/00, ¶ 41 (Eur. Ct. H.R. 2002), http://hudoc.echr.coe.int/eng/?i=001-60605. This case concerned an applicant convicted of drug trafficking, a serious crime with “devastating effects” on the society. Id. ¶¶ 15, 37. Note, in contrast to the UNHRC, most of the cases that come before the ECtHR involve applicants with lawful status who were ordered deportation based on criminal convictions. The ECtHR is less likely to find a violation of the right to family unity when dealing with applicants that were convicted on drug-related or other serious charges and more likely to find a violation when the applicants had a citizen spouse and children or had resided in the country since early childhood. Compare Keles v. Germany, App. No. 32231/02 (Eur. Ct. H.R. 2005), http://hudoc.echr.coe.int/eng/?i=001-70824, with Baghli v. France, 1999-VIII Eur. Ct. H.R. 169. In a series of cases, the ECtHR established the test for determining violation of the right to family unity: whether the deportation order was “necessary in a democratic society.” See, e.g., Dalia v. France, App. No. 26102/95, 33 Eur. H.R. Rep. 625, ¶¶ 49–55 (1998) (Eur. Ct. H.R.). Factors the ECtHR tends to weigh heavily include: the “nature and seriousness” of the offenses, the ability of the applicant to maintain contact with his or her family even if deported (i.e., whether the interference with the right to family is total or partial), and whether the claim is made on behalf of the individual being deported alone or additional family members as well (in particular children). See, e.g., Boultif v. Switzerland, 2001–IX Eur. Ct. H.R. 119, ¶ 48. For a useful discussion of how the ECtHR applies the right, see Ramji-Nogales, supra note 54, at 737–38. The ECtHR also made the deportation of foreigners who have committed serious crimes more difficult for the host State if the foreigner concerned is a person of a so-called “second generation.” See, e.g., Moastaquim case v Belgium 1991, App. No. 12313/86, ¶ 13, 44 (concerning a Moroccan national who arrived to Belgium at the age of two and he and his family and relatives all lived in Belgium); Beldjoudi v. France 1992, App. No. 12083/86 (concerning a plaintiff who was born in France of parents who originated from Algeria, a territory which was French at the time, and how was deemed to have lost his French nationality as his parents did not make a declaration of recognition).
Second, under prior case law, the UNHRC and the ECtHR accepted that claimants without a lawful status cannot present the State with a “fait accompli,” i.e., establishing residence does not lead to rights. However, more recent case law, discussed below, adds uncertainty over whether presenting the host State with the birth of a child can sway the balance in favor of the individual’s interest in family life over the State’s right to control its immigration policy.

For the UNHRC, the birth of a child does not categorically prevent the deportation of the parent who is in the State in breach of its immigration law. But it does make the deportation procedurally more difficult. Winata v. Australia concerns two Stateless individuals who overstayed their visa terms and gave birth to a son in Australia. The day after the son was granted Australian citizenship—because he was born in the country and had resided there for ten years—his parents asked for a protection visa, which Australia denied. In light of the length of time the parents and their son had spent in Australia, however, the Committee determined that Australia was under a duty to demonstrate “additional factors justifying the removal of both parents that go beyond a simple enforcement of its immigration law in order to avoid a characterization of arbitrariness.” This, in effect, required Australia to grant the parents a status.

For the ECtHR, in turn, the birth of a child might ban the deportation of the parent. An example is Nunez v. Norway. In that case, an applicant entered the country with a forged passport and, once there, received a residence permit and had children with whom she developed “long lasting and close bonds.” In this case, the Court ruling was based on the best interests of the children, and held

---


59. Id. ¶ 2.1. The couple were formerly Indonesian nationals. They arrived on valid visas that subsequently expired. Id. §§ 1.1.

60. Id. §§ 2.2–2.4. First, the couple applied for asylum, but after their application was denied, they appealed their asylum claim and applied for a parent visa. Id. Once their asylum appeal was denied, they asked that the government exercise humanitarian discretion based on hardship to their son of removal to Indonesia. Id.

61. Id. ¶ 4. For contrast, see Stewart v. Canada, Communication No. 538/1993, U.N. Human Rights Comm., 58th Sess., Oct. 21–Nov. 8, 1996, U.N. Doc. CCPR/C/58/D/538/1993 (Dec. 16, 1996), where, when dealing with Articles 12 of the ICCPR (the freedom of movement or the right to enter one’s country) and Article 17 (the right to a family life) of a permanent resident, the majority rejected the result of Canada’s immigration law.

62. While technically the couple was undocumented, they had a viable route to lawful status (parent visa). For a detailed discussion of the case, see Ramji-Nogales, supra note 54, at 734–35, and Gavison, supra note 40, at 36–37.


64. Id. ¶ 84.
that expulsion of the mother violated the right of the children to a family life. 65

A year later, however, in Antwi v. Norway,66 the Court reached the opposite conclusion with regard to an applicant who also came into the country using forged documents, received a work and residence permit, and had a daughter of whom he was the main caretaker.67 In that case, the judges held that even if an applicant established a family, there is no Article 8 violation if “[a]t no stage from when he entered [the country] . . . could he reasonably have entertained any expectation of being able to remain in the country.”68 While the majority in Antwi considered that there were “fundamental differences” between Antwi and Nunez,69 the strong dissenting opinion was adamant that the two cases were “very similar” and “the solution in Nunez should have been applied in the present case a fortiori.”70 This leaves a State uncertain as to how the Court will hold in the next case that deals with the expulsion order of a parent who entered the State in breach of its immigration laws and had a child who is still young at the time of the order.

Third, prior established jurisprudence of both the UNHRC and the ECtHR prioritizes the right to family life as a basis to restrict a State’s discretion to expel nonnationals, but such protection considered only immediate family members. More recently, however, the ECtHR (though not the UNHRC) began recognizing a free-standing right to private life, thereby protecting the totality of the social relationships that an alien’s presence spawns in the host country as a grounds to bar deportation. In Slivenko v. Latvia,71 the Court, sitting as the Grand Chamber, reversed the deportation order of a former Soviet army officer.

65. Id. ¶¶ 79–82, 84 (The applicant was “the children’s primary care person from their birth,” the children “lived all their lives in Norway,” and had already suffered “disruption and stress” due to the decision in the custody proceedings that moved them to the father after the deportation order was issued. “In these circumstances,” the Court concluded that, “the children were vulnerable” and that deporting the mother would violate Article 8).


67. Id. ¶¶ 6, 9, 72.

68. Id. ¶ 91.

69. Id. ¶ 100. In Nunez, the daughters developed “long lasting and close bonds to their mother,” and the decision in the custody proceedings to move the children to the father had already led the children to experience significant “disruption and stress,” and a “long period” elapsed “before the immigration authorities took their decision to order the applicant’s expulsion with a re-entry ban.” Id. But “[i]n light of the situation of the children of Mrs. Nunez, [Antwi’s daughter] had not been made vulnerable by previous disruptions and distress in her care situation . . . .” Also, the duration of the immigration authorities’ processing of the matter was not so long as to give reason to question whether the impugned measure fulfilled the interests of swiftness and efficiency of immigration control that was the intended purpose of such administrative measures . . . . [Therefore] the Court is satisfied that sufficient weight was attached to the best interests of the child in ordering the first applicant’s expulsion.” Id. ¶¶ 101–02.

70. Id. ¶¶ 9–10 (“Contrary to the opinion of the majority, the present case is very similar to Nunez, . . . . [I]f there is indeed a difference between Nunez and the present case, this lies in the fact that the latter is even more striking than the former. Consequently, the solution in Nunez should have been applied in the present case a fortiori.”).

and his family from Latvia following the withdrawal of Russian troops.\textsuperscript{72} As the deportation order concerned all members of the family unit, it did not amount to an interference with the Slivenkos’ right to family.\textsuperscript{73} Yet the judges concluded that the family’s right under Article 8 had, nonetheless, been violated because they were “removed from the country where they had developed, uninterruptedly since birth, the network of personal, social and economic relations that make up the private life of every human being. Furthermore, they lost the flat in which they had lived.”\textsuperscript{74}

The growing bias of both the UNHRC and the ECtHR in favor of the human rights tradition is possibly even more evident in cases bearing on the right not to be subject to “torture or to cruel, inhuman or degrading treatment or punishment.”\textsuperscript{75} The best examples are cases that deal with non-national applicants who are charged with involvement in terrorism.

Neither the ICCPR nor the ECHR contains a right to political asylum. But both the UNHRC and the ECtHR read a non-refoulement obligation,\textsuperscript{76} or the “cardinal principle of international refugee law,”\textsuperscript{77} into Articles 6 and 7 of the ICCPR and Article 3 of the ECHR. This reading of language from two treaty instruments leaves the host State in the worst situation. Under the Refugee Convention, the non-refoulement right is restricted in cases that involve criminal and security threats to the State.\textsuperscript{78} Under both the ICCPR and the ECHR, in turn, the right not to be subject to “torture or to cruel, inhuman or degrading treatment

\begin{thebibliography}{99}
\bibitem{72} Id. \S 16–18, 128–29.
\bibitem{73} Id. \S 97.
\bibitem{74} Id. at 232; see also Maslov v. Austria, 2008-III Eur. Ct. H.R. 301 \S 63. Further, in \textit{Kuric v. Slovenia}, in his partly concurring opinion, Judge Vučinić, observed that the right to private life required protection of the ability of an individual to have relationships in a “public context”—he described this aspect of the right as follows: “Article 8 protects . . . the right to personal development and the right to establish and develop relationships with other human beings as well as the outside world, even in the public context, which may also fall within the scope of ‘private life.’” \textit{Kurić v. Slovenia}, 2012-IV Eur. Ct. H.R. 1, 84 (partly concurring, partly dissenting opinion of Judge Vučinić).
\bibitem{75} ICCPR arts. 6–7, Mar. 23, 1976, 999 U.N.T.S. 171; ECHR art. 3, Nov. 4, 1950, 213 U.N.T.S. 221.
\bibitem{76} For a definition, see Guy S. Goodwin-Gill, \textit{The Right to Seek Asylum: Interception at Sea and the Principle of Non-Refoulement}, 23 INT’L J. REFUGEE L. 443, 444 (2011) (“The obligation on [S]tates not to send individuals to territories in which they may be persecuted, or in which they are at risk of torture or other serious harm . . . .”).
\bibitem{78} Refugee Convention art. 33(2), Apr. 22, 1954, 189 U.N.T.S. 150 (“The benefit of the present provision may not, however, be claimed by a refugee whom there are reasonable grounds for regarding as a danger to the security of the country in which he is, or who, having been convicted by a final judgment of a particularly serious crime, constitutes a danger to the community of that country.”). For a detailed discussion of this exception, see \textit{HATHAWAY, THE RIGHTS OF REFUGEES UNDER INTERNATIONAL LAW}, supra note 30, at 342–55. Hathaway writes: “In cases that fall under Art. 33(2), the asylum country is authorized to expel or return even refugees who face the risk of extremely serious forms of persecution.” Id. at 344.
\end{thebibliography}
or punishment” is absolute.\textsuperscript{79} The UNHRC and the ECtHR, however, have imported from the Refugee Convention only the right of non-refoulement without the qualification and, in addition, they attached it to the non-derogatory nature of the right not to be subject to “torture or to cruel, inhuman or degrading treatment or punishment.” This means that both enforcement institutions forbid the deportation of non-nationals charged with terrorism if they face degrading treatment upon return to their home country on account of their involvement in terrorism.

And so, in \textit{Ahani v. Canada},\textsuperscript{80} the UNHRC reviewed a communication dealing with an author who, after he was accepted as a refugee, was identified by Canada as a trained assassin and was put on deportation proceedings, even though he claimed that if sent back he would face torture and execution.\textsuperscript{81} Canada deported the refugee before the UNHRC reached its determination. But the Committee held that “the prohibition on torture . . . is an absolute one that is not subject to countervailing considerations.”\textsuperscript{82} Similarly, in \textit{Othman (Abu Qatada) v. United Kingdom},\textsuperscript{83} the ECtHR reversed the deportation order of a radical Islamic preacher regarded as one of Al Qaeda’s main inspirational leaders in Europe because of the risk that he would be tortured to obtain evidence.\textsuperscript{84} The decision held: “Article 3 is absolute and it is not possible to weigh the risk of ill-treatment against the reasons put forward for the expulsion.”\textsuperscript{85}

In addition, the ECtHR, but not the UNHRC,\textsuperscript{86} has gone even further in expanding the scope of the right not to be subject to “torture or to cruel, inhuman or degrading treatment or punishment.” The European Court has made three separate moves.

\textsuperscript{79} For the non-derogable nature of Article 3 of ECHR, see the landmark case \textit{Soering v. United Kingdom}, 161 Eur. Ct. H.R. (ser. A) ¶ 88 (1989) (“Article 3 . . . makes no provision for exceptions and no derogation from it is permissible . . . .”).


\textsuperscript{81} \textit{Id.} ¶¶ 2.1–2.5.

\textsuperscript{82} \textit{Id.} ¶ 10.10.


\textsuperscript{84} \textit{Id.} ¶ 25.

\textsuperscript{85} \textit{Id.} ¶ 185; see also \textit{Saadi v. Italy}, 2008-II Eur. Ct. H.R. 144, ¶ 139 (“The Court considers that the argument based on the balancing of the risk of harm if the person is sent back against the dangerousness he or she represents to the community if not sent back is misconceived.”); \textit{Chahal v. United Kingdom}, 1996-V Eur. Ct. H.R. 413.

\textsuperscript{86} In general, the jurisprudence of the ECtHR is more liberal than that of the UNHRC. In fact, the Court has been called “the crown jewel of the world’s most advanced international system for protecting civil and political liberties.” Laurence R. Helfer, \textit{Redesigning the European Court of Human Rights: Embeddedness as a Deep Structural Principle of the European Human Rights Regime}, 19 EUR. J. INT’L L. 125, 159 (2008). In cases that bear on immigration, applicants to the UNHRC often ask that the Committee take guidance from the jurisprudence of the ECtHR. See, e.g., \textit{Winata, supra note 58, ¶ 3.5; Ahani, supra note 80, ¶ 3.5.}
First, the ECtHR has drastically expanded the substantive grounds of non-refoulement. Traditionally, the European Court had carefully capped the scope of Article 3 non-refoulement obligations at ill treatment that resulted from persecution in situations where the alien faces a well-founded fear of harm in her home State on account of any of the five familiar grounds.\(^87\) But case law excluded “ill treatment” that derived either from widespread violence due to an “unsettled situation”\(^88\) or from an “acute pertinence of socio-economic” deprivation in the receiving country.\(^89\) In the span of three years, however, the ECtHR extended non-refoulement protections to include cases involving general, widespread violence or potential socio-economic deficiency.

In \textit{NA. v. United Kingdom},\(^90\) the ECtHR, sitting as the Grand Chamber, held that the Court will not discount “the possibility that a general situation of violence in a country of destination will be of a sufficient level of intensity as to entail that any removal to it would necessarily breach Article 3 of the Convention.”\(^91\) Three years later, in \textit{MSS v. Belgium and Greece},\(^92\) a case that dealt with an asylum seeker who reached the territory of the host State, the

---

\(^87\) See supra p. 7 (discussing five grounds from Refugee Convention).

\(^88\) See, \textit{inter alia}, Saadi, \textit{supra} note 85, § 131 (remarking that “the mere possibility of ill-treatment on account of an unsettled situation in the receiving country does not in itself give rise to a breach of Article 3”); Fatgan Katani v. Germany, App. No. 67679/01, (Eur. Ct. H.R. 2001); H.L.R. v. France, 1997-III Eur. Ct. H.R. § 41 (noting a “general situation of violence existing in the country of destination . . . . would not in itself entail, in the event of deportation, a violation of Article 3”); Vilvarajah v. United Kingdom, 215 Eur. Ct. H.R. (ser. A) § 111 (1991) (“Since the situation was still unsettled there existed the possibility that they might be detained and ill-treated as appears to have occurred previously in the cases of some of the applicants . . . . A mere possibility of ill-treatment, however, in such circumstances, is not in itself sufficient to give rise to a breach of Article 3 . . . .”); Press Release No. 228(2005), Council of Europe, European Court of Human Rights – Chamber Judgments Concerning France, Poland, Turkey and Ukraine, Council of Eur., https://wcd.coe.int/ViewDoc.jsp?id=893751&Site=COE (summarizing that in Muslim v. Turkey, App. No. 53566/99 (Eur. Ct. H.R. 2005), the ECtHR “reaffirmed that a mere possibility of ill-treatment as a result of temporary instability in the country did not in itself entail a breach of Article 3”).

\(^89\) N. v. United Kingdom, 2008-III Eur. Ct. H.R. 227, § 44 (“Although many of the rights it contains have implications of a social or economic nature, the Convention is essentially directed at the protection of civil and political rights.”); id. § 42 (“Aliens who are subject to expulsion cannot in principle claim any entitlement to remain in the territory of a Contracting State in order to continue to benefit from . . . . social or other forms of assistance and services provided by the expelling State. The fact that the applicant’s circumstances . . . . would be significantly reduced if he were to be removed from the Contracting State is not sufficient in itself to give rise to a breach of Article 3.”); Sheetkh v. Netherlands, App. No. 1948/04, § 141 (Eur. Ct. H.R. 2007), http://hudoc.echr.coe.int/eng?i=001-78986 (“While the Court by no means wishes to detract from the acute pertinence of socio-economic . . . considerations to the issue of forced returns of rejected asylum seekers to a particular part of their country or origin, such considerations do not necessarily have a bearing, and certainly not a decisive one, on the question whether the persons concerned would face a real risk of ill-treatment within the meaning of Article 3 of the Convention in those areas.”).


\(^91\) Id. § 115.

ECHR, again sitting as the Grand Chamber, held that acute financial deprivation, or a condition where an asylum seeker is “wholly dependent on State support” and finds herself “in a situation of serious deprivation or want incompatible with human dignity,” may likewise fall within the reach of Article 3 protection.93

Second, the ECHR has significantly liberalized the procedural threshold required to demonstrate an Article 3 non-refoulement violation.94 In the early 1990s, the Court applied a narrow assessment of risk: an applicant had to produce “substantial grounds” that he “faces a real risk”95 on account of one of the five grounds.96 However, by the early 2000s, the Court tolerated a more lax standard: “concerns as to the risks [the applicant] faced,” for example, were sufficient to trigger Article 3 non-refoulement duty.97 Similarly, the Court moved from requiring a fairly high level of individualization (an applicant’s personal “situation” must be “worse than the generality of other members” of his community “who were returning to the country”)98 to accepting a more general risk (for instance, possibility of ill-treatment on account of “a general situation of the non-observance of human rights in the applicant’s home country.”).99

Third, in dealing with those classified as asylum seekers, the ECHR enlarged the right of non-refoulement from a minimal negative obligation not to deport (non-removal)100 to a positive obligation to protect. The key case here is M.S.S v. Belgium & Greece, mentioned previously. In deciding the case, the ECHR’s Grand Chamber held that the failure to process asylum applications “within a reasonably short time and with utmost care” 101 in circumstances where the applicant is “wholly dependent on State support” and in “a situation of

93. Id. ¶¶ 252–53, 263 (deprivation must be serious enough to reach levels of “extreme material poverty”).
94. For a detailed discussion of this point, see VAN DIJK ET AL., supra note 49, at 433–34.
96. See supra p. 7 (discussing five grounds from Refugee Convention).
98. Vilvarajah, supra note 88, at 111–12.
99. N.A v United Kingdom, App. No. 25904/07, 2008 at 115 (“the Court has never excluded the possibility that a general situation of violence in a country of destination will be of a sufficient level of intensity as to entail that any removal to it would necessarily breach Article 3 of the Convention.”).
100. See, e.g., FOUNDATIONS OF INTERNATIONAL MIGRATION LAW 193 (Brian Opeskin et al. eds., 2012) (“[T]he duty of non-refoulement only prohibits measures that cause refugees to ‘be pushed back into the arms of their persecutors’; it does not establish an affirmative duty to receive refugees.”); Gregor Noll, Seeking Asylum at Embassies: A Right to Entry Under International Law?, 17 INT’L J. REFUGEE L. 542, 548 (2005) (“Non-refoulement is about being admitted to the [S]tate community, although in a minimalist form of non-removal.”).
101. M.S.S. v. Belgium & Greece, supra note 92, at 103 (Judge Sajó, partly dissenting)
serious deprivation" 102 engages the State’s responsibility to provide asylum seekers with affirmative support and, in particular, adequate housing. 103

The combined result of these three moves is that the Strasbourg Court is growing the regime of refugee law from one that is grounded in narrow exceptions relevant essentially to first-world concerns 104 into one that can deal with mass atrocities (economic, environmental, and political) across the world. In some circumstances, moreover, this Court also attaches positive obligations. And so the ECtHR holds the State owing significant protection to an undefined number of individuals it never intended to let into the country in the first place.

To be sure, none of the rulings analyzed above provide precise parameters for when the State can and cannot deport non-nationals. Many questions remain open. For example, when dealing with the right to a family life, the precise scope of protection is contested. The vast majority of available case law deals with non-nationals in one of two situations: (i) an individual in a permanent lawful status who broke his or her terms of entrance (like Amrollahi) or (ii) those without a status but with a viable path to lawful status prior to deportation proceedings (like Winata). This leaves unresolved whether enforcement bodies would be willing to prioritize the family right of an individual over the State’s prerogative to exclude in cases that involve less sympathetic undocumented migrants who push harder on the immigration policy of the host State. 105

Similarly, with the right to non-degrading treatment, it is still undefined how bad the violence or poverty must be to bar deportation.

What is certain is that when it comes to immigration, in the past ten years the UNHRC and the ECtHR have changed their bias in favor of the universalist tradition. They increased the access of non-nationals to human rights. In particular, they read human rights norms more strictly and more absolutely, and developed substantive standards of protection beyond the five traditional grounds in the Refugee Convention—especially with regard to the ECHR. Importantly, however, the way in which they moved in the direction of the human rights tradition interlocks with, rather than opposes, the statist dedication to territory: human rights courts and quasi-judicial bodies correlate jurisdiction with physicality grounded in territory. 106

Beneficiaries have access to more

102. Id. ¶ 253.

103. Id. ¶ 263 ("The Greek authorities have not had due regard to the applicant’s vulnerability as an asylum-seeker and must be held responsible, because of their inaction, for the situation in which he has found himself for several months, living on the street, with no resources or access to sanitary facilities, and without any means of providing for his essential needs.").

104. Harold Koh refers to this as the “good aliens.” See, e.g., Harold Hongju Koh, Who Are the Archetypal “Good” Aliens? 451 (Jan. 1, 1994) (unpublished manuscript) (on file with the Yale Law School Legal Scholarship Repository) ("T]he archetypal ‘good’ alien . . . is a white, healthy, law-abiding, self-sufficient, anti-communist, heterosexual, male political refugee, who arrives by himself at the U.S. Embassy in Moscow and seeks political asylum; Rostropovich and Baryshnikov are two obvious examples."). For more on a comparison between the “good” alien of the Cold War and the “bad” alien of the 1990s onwards, see Chimni, supra note 77, at 355–60.

105. For this point, see Ramji-Nogales, supra note 54, at 733–38.

106. Human rights protection is also triggered if the plaintiff reaches under the effective
rights, but they can only trigger the State’s protection of those rights if they are able to reach the State’s shores. Once inside the State, the rights are inherent in the individual and external to the State’s interests. The State, in turn, is held accountable for these rights, even if meeting this expectation is politically or financially costly. Outside the State’s jurisdiction, however, the plight of the non-national is of no legal concern to the State. And so, as human rights adjudicatory bodies moved in the direction of the universalist legal tradition, they have also, in effect, further produced territoriality.

II. FRONT DOOR STRATEGIES OF IMMIGRATION CONTROL

States have moved to tighten up immigration from what I term the “front door”: stopping would-be immigrants or asylum seekers ex ante before they reach any direct contact with the territory of the receiving State and can activate protective duties. Here, I only look at two strategies of “front door” control: maritime migrant interdiction on the high seas and the building of border walls as an immigration control policy. These two strategies are similar. States utilize defined physical boundaries to stop immigrants from getting in, either by land or sea, so that their entry does not activate State obligations for their protection. Indeed, Professor Harold Koh referred to interdiction as a “floating Berlin Wall.” But while interdiction is extensively researched, walls remain

control of the state, even if she is not physically present on the state’s territory proper. See infra p. 28 for discussion of the extraterritorial application of human rights law.

107. In the context of non-refoulement, this idea was nicely summed up by the House of Lords: the legal protection “is concerned only with where a person must not be sent, not with where he is trying to escape from.” European Roma Rights Centre v. Immigration Officer at Prague Airport, [2003] EWCA (Civ.) 666, [37] (Eng.), aff’d, R v. Immigration Officer at Prague Airport, [2004] UKHL 55 (appeal taken from Eng.). James Hathaway adds that a weakness of non-refoulement is that it traps “would-be refugees . . . inside their own countri[es].” HATHAWAY, THE LAW OF REFUGEE STATUS, supra note 30, at 19.

108. Other scholars use the term “non-entrée.” See James C. Hathaway & Thomas Gammeltoft-Hansen, Non-Refoulement in a World of Cooperative Deterrence 6 n.12 (Mich. Law Sch. Law & Economics Working Paper, 2014) (noting the term “was first employed by James Hathaway” in 1992, and that “[i]n essence, it suggests that whereas refugee law is predicated on the duty of non-refoulement (that refugees shall not be turned away), the politics of non-entrée is based on a commitment to ensuring that refugees shall not be allowed to arrive.”). I, however, employ the phrase “to close the front door” to refer to restrictions that take place at the actual border and therefore involve the specificity of the border itself. In this way, my phrase “front door” is different from and narrower than “non-entrée”: “front door” only applies to restrictions that take place on the actual, physical territorial border of the State; “non-entrée” applies more broadly to all restrictions on entrance wherever they take place.

109. Harold Hongju Koh, Closed Door Policy for Refugees, LEGAL TIMES S36, S37 (July 26, 1993) (“The Kennebunkport Order effectively erected a floating Berlin Wall around Haiti, preventing Haitians from fleeing not just to the United States, but to any of the scores of islands between the United States and Haiti.”); see also Linda Greenhouse, Court is Asked to Back Haitians’ Return, N.Y. TIMES, Mar. 3, 1993, at A16 (“Mr. Koh said his position would not require the United States to accept all Haitian immigrants. He said there were other islands the Haitians might reach if they were not prevented from leaving by a ‘floating Berlin wall.’”); Harold Hongju Koh, The ‘Haiti
Yet walls make concrete the morally unsatisfactory nature of the compromise that courts worked out between universality and exclusion.

Let me begin with interdiction. Starting in the 1980s, highly developed nations increasingly turned to maritime interdiction on the high seas. By the early 2000s, the practice was consolidated into a key border enforcement tool for coastal States, and, in particular, for the United States, the European Union, and Australia. The U.S. Supreme Court, called to review the practice in Sale v. Haitian Ctrs. Council, provided what later developed into the model justification for interdicting States. The Supreme Court held that human rights
obligations are strictly territorial, and international treaties “cannot impose . . . extraterritorial obligations on those who ratify it through no more than its general humanitarian intent.” Thus, human rights do not apply on the high seas and are only triggered “on the threshold of initial entry.”

And the response of human rights courts and quasi-judicial bodies? Both the UNHRC and the ECtHR have increasingly constrained States’ ex ante strategies of interdiction. In *Jamaa v. Italy*, the ECtHR, sitting as the Grand Chamber, provided the landmark ruling on interdiction. The judges explained that human rights jurisdiction “is essentially territorial.” But it is also engaged “[w]henever the State through its agents operating outside its territory exercises control and authority over an individual, and thus jurisdiction.” When a non-national is affected by those acting on behalf of the State, “the State is under an obligation . . . to secure to that individual the [human rights] that are relevant to the situation of that individual.” In interdiction cases, this means that an intercepting State must provide the passengers on the boat with procedural guarantees (an individual refugee-status determination procedure) and a

116. *Sale*, 509 U.S. at 183. In contrast, see Harold Hongju Koh arguing that foreign courts were bound by “principles of comity, sanctity of treaty, and respect for human rights that must form the bedrock of any new world order[,]” *Reflections on Repatriation and Haitians Centers Council, 35 HARV. INT’L L.J.* 1, 20 (1994).


119. *Id.* ¶ 71 (and is “presumed to be exercised normally throughout the State’s territory”).

120. *Id.* ¶ 74.

121. *Id.; see also id.* at 173 (Pinto de Albuquerque, J., concurring) (“The prohibition of *refoulement* is not limited to the territory of a State, but also applies to extraterritorial State action, including action occurring on the high seas.”). Note that in this holding, the ECtHR is merely repeating what it said many times prior. See, e.g., Ilașcu v. Moldova, 2004-VII Eur. Ct. H.R. 179; Banković v. Belgium, 2001-XII Eur. Ct. H.R. 333. But, in contrast, Chimène I. Keitner confirms that even in a context in which courts increasingly face claims regarding the rights of people located beyond their countries’ borders, they largely remain bound to territorial adjudication. *Rights Beyond Borders*, 36 YALE J. INT’L L. 55, 57–58 (2011). The ECtHR has been reluctant to apply the ECHR outside the territory of the Convention States (notably Banković v. Belgium, 2001-XII Eur. Ct. H.R. 333). However, even in this case, the Court concluded that “the ECtHR has consistently held that the obligations under the ECHR apply extraterritorially in situations where a “State, through the effective control of the relevant territory and its inhabitants abroad . . . exercises all or some of the public powers normally to be exercised by that government.” Banković, 2001-XII Eur. Ct. H.R., ¶ 71. In more recent cases, the ECtHR based the decisions in which it declined jurisdiction for acts outside the territory of a Member State not on territorial grounds, but on other considerations. See Ilașcu v. Moldova, 2004-VII Eur. Ct. H.R. 179, ¶ 310–31 (holding Moldova responsible even in the absence of effective control over the Transnistrian region within Moldova).

122. *Jamaa v. Italy*, 2012-II Eur. Ct. H.R. 97, ¶ 185 (the host State must provide an “examination of each applicant’s individual situation” by personnel that is “trained to conduct individual interviews” and “assisted by interpreters or legal advisers.” This means that collective expulsion is in breach of Article 4 of Protocol No. 4 to the Convention).
substantive obligation (non-refoulement, or not to send back an individual who faces harm).

Similar to the ECtHR, the UNHRC has also held that the State is responsible for the human rights (including non-refoulement) of “all persons in their territory and all persons under their control.”

And the extraterritorial application of human rights was likewise supported by multiple other international human rights bodies as well as national courts.

In Part II, this Article showed that the UNHRC and the ECtHR deploy jurisdiction based on territory: they require territorial presence in order to activate human rights protective obligations. Here, with interdiction on the high seas, the UNHRC and the ECtHR apply jurisdiction extraterritorially: they correlate jurisdiction with physicality grounded in contact, such that jurisdiction follows the State on the high seas wherever it establishes contact with or exercises effective control over the non-national, whether that is inside or


outside of the State’s borders. This extraterritorial application of human rights law is not an abandonment of the logic of territoriality; it does not break out of the conceptual bind between access and territory. Instead it is an extension, even a hypertrophy, of territoriality: it only expands the limits of the existing territorial scheme.

The act of exercising border constraint—the State’s interception or contact with a non-national—is legally crucial. The plaintiff (the individual right bearer) is not only where she is physically found, but also where she might have been without the coercion of the State (the interception practice), so that her intended destination State becomes a defendant. A State, then, can coerce, or repel an asylum seeker or a would-be immigrant from entering its territory, but that act of coercion—the act of border control—itself triggers human rights protective responsibilities because the State exercises effective control over the individual. And so, on the high seas, individuals are always protected by human rights law, or, in the words of the ECtHR Grand Chamber, “the maritime environment cannot justify an area outside the law where individuals are covered by no legal system capable of affording them enjoyment of the rights and guarantees protected by [the ECHR].”

A State that seeks to maintain control over the inflow of would-be-immigrants and asylum seekers coming in by boat, without accruing protective duties over an undefined number of people, must therefore move its immigration control to the shadows. It ought to deter at sea without looking like it is deterring, and develop soft deterrents in lieu of, or in advance of, hard deterrents that would trigger human rights protections. Indeed some European States are already outsourcing interdiction practices to source or transit countries, thereby avoiding any direct fingerprint that would trigger jurisdiction and broad protective obligations.

What about walls? Around 2005, States also began building border walls and other physical or technological constraints as part of their immigration control policy. Today, we see such walls across the North America, the EU, the Middle East, and Africa. The legal justification for such walls, as stated by the United States District Court for the District of Columbia, “pertains to

both foreign affairs and immigration control” and is “inherent in the executive
department of the sovereign.”129

As for the human rights response? At the moment, border walls are
relatively unregulated under the law. Yet immigrants are coming to these walls
in ever larger and increasingly coordinated numbers.130 It is thus only a question
of time until a human rights court will be called on to regulate such a wall.

A human rights court or a quasi-judicial institution that is asked to
adjudicate a border wall built as part of an immigration control strategy will
have to choose between three competing approaches to regulating such a wall.
These choices correspond to the tensions, outlined at the beginning of this paper,
between the universalist and exclusionist frameworks. Each approach correlates
to a radically different vision of borders and sovereignty in the international
order, and, at the same time, none of these three methods for regulating a border
wall can be defended normatively and continuously.

A. First Method—Adopt the Universalist Tradition.

A court that chooses this tradition builds upon the similarities between
liquid and solid barriers, interdiction and wall-building, to apply the interdiction
precedent to a wall scenario: if human rights apply on the high seas before non-
nationals enter the territory of the State, human rights are also guaranteed to
non-nationals approaching a wall, before they cross to the other side. This would
mean that a host State could build a wall, but still owe procedural duties
(individual assessment of refugee claim) and substantive duties (non-
refoulement) to anyone who comes close to the wall. In doing so, the court takes
the Jamaa v. Italy ruling to its ultimate conclusion: jurisdiction aligned with
physicality and grounded in proximity.

To support this approach, a court could defer to traditional interpretations
of the duty of non-refoulement under the Refugee Convention, which are
normally understood to constrain both ejection from within a State’s territory
and non-admittance at its frontiers.131 Such a court could even go a step further.
A recent report by the Inter-American Commission of Human Rights that deals,
inter alia, with the “terrible effects”132 of the U.S.-Mexico wall, explains:

130. See discussion infra Conclusion.
131. See, e.g., Exec. Comm. of the High Comm’r’s Programme, Non-Refoulement, Conclusion
No. 6(c) (XXVIII) (Oct. 12, 1977), http://www.unhcr.org/3ae66c43ac.html (acknowledging “the
fundamental importance of the observance of the principle of non-refoulement—both at the border
and within the territory of a State”); GREGOR NOLL ET AL., STUDY ON THE FEASIBILITY OF
PROCESSING ASYLUM CLAIMS OUTSIDE THE EU AGAINST THE BACKGROUND OF THE COMMON
EUROPEAN ASYLUM SYSTEM AND THE GOAL OF A COMMON ASYLUM PROCEDURE 36 (“Today, there
appears to be ample support for the conclusion that Article 33(1) of the Refugee Convention is
applicable to rejection at the frontier of a potential host [S]state.”). For a detailed discussion, see
HATHAWAY, THE RIGHTS OF REFUGEES UNDER INTERNATIONAL LAW, supra note 30, at 315–18.
132. INTER-AM. COMM’N ON HUMAN RIGHTS, REPORT ON IMMIGRATION IN THE UNITED
STATES: DETENTION AND DUE PROCESS §§ 107–08 (2010),
One of the most harmful effects of the physical barriers erected along the border is that . . . they merely steer immigrants in the direction of those border areas where no physical barriers have been erected and where conditions tend to be so extreme as to make the crossing highly dangerous. Summing up, this type of measure increases the death rate among undocumented migrants . . . .

This report is only an observation and a caution on the effects of the U.S.-Mexico barrier. But it does suggest that the wall, by impeding immigration flow at certain crossings and channeling it instead to more dangerous routes, may itself trigger human rights jurisdiction because of the foreseeable harm to would-be migrants and asylum seekers.

Earlier, in the discussion of the interdiction cases, this Article highlighted that the UNHRC and the ECtHR aligned jurisdiction with physicality grounded in contact (the non-national’s coming within the effective control of the State). The State’s act of coercion—turning away the boat—confers jurisdiction regardless of whether the border control practices take place on the State’s territory or on the high seas. However, here contact is no longer required to trigger responsibilities. Getting close to the wall (even if there is no actual contact with the State or its agents), or jurisdiction grounded in proximity, would be equal to establishing territorial presence inside the State. A host State would owe protective duties not only to anyone it actively forced away, but also to a potentially unlimited number of plaintiffs who reach the vicinity of the wall.

In this approach, a border wall acts as a bridge: being on the other side of the wall is as good as being inside the State’s territory. But a wall is simply a physical manifestation of the border. It reinforces a border that was always there and is not disputed. If getting close to the wall triggers human rights protection (jurisdiction grounded in proximity), then the State loses effective control over its borders: it accrues obligations to individuals on both sides of the border. Indeed, Professor Guy Goodwin-Gill, one of the leading scholars of refugee law and a legal adviser in the Office of the UNHCR from 1976 to 1988, explains that borders “do not mark the limit of [international] law.” And so, just as much as on the high seas individuals are always protected by human rights law (recall Jamaa’s statement that even in “the maritime environment” there is no “area outside the law”), there is no place on land that is not covered by human rights law. Again from Professor Goodwin-Gill: “[T]here is no physical space


133. Id. ¶ 107.

134. The United States could respond to the concern of the Commission by sealing off the more dangerous crossings (i.e., increasing the wall) as well as by taking down the wall. Alternatively, it could do nothing.


137. It should be noted that more open borders do not mean automatic protection. Rather it means that more people would come under human rights jurisdiction and would, therefore, be able to invoke protection. However, whether they would actually be covered by the law would still depend

http://scholarship.law.berkeley.edu/bjil/vol34/iss1/1
and no realm of human activity that is beyond the rule of law."138 Sovereignty in this approach no longer denotes a space that is outside human rights law. Pushed to its extreme, the result is open borders: a universal application of human rights that is divorced from territorial limitations (one’s place of birth carries no legal significance in the operation of human rights).

An “open borders” international regime is, however, at this point, utopian and disconnected from reality.139 It would provide absolute rights while largely ignoring consequentialist concerns about implementation or remediation.140 Furthermore, if States came to view as unsustainable the number of immigrants that human rights courts might press them to accept, they might even choose to withdraw altogether from the jurisdiction of international human rights courts and other quasi-judicial institutions altogether.141 This is not an empty threat—in fact, former British Prime Minister David Cameron put this exact loaded gun on the table, vowing that he is ready to lead Great Britain outside the ECtHR if it is the only way to send back foreign criminals.142

Even if a scenario of open borders were feasible, in a reality of finite resources, boundaries must be permitted somewhere in order for welfare rights to be economically and politically tolerable. Without boundaries, citizenship rights (or membership that guarantees some form of an insider preference) would become meaningless, and all that would be left would be individual property rights (or the ability to provide for oneself). Citizens and non-citizens alike would be eligible to receive precisely nothing from any public entity. This leaves those citizens without property in the host State worse off than before borders were open, and non-nationals that come into the State no better off.

on whether they met the definition of protection.


139. Without borders, the existing notion of a state-based polity (including citizenship) disappears.

140. This resembles a Dworkinian top-down process of constitutional adjudication that is not concerned with considerations of efficacy, consequences, empirical data, or political pressures. But Dworkin writes in the context of a closed philosophical system of abstract legal principles with no real world consequences. See, e.g., RONALD DWORKIN, LAW’S EMPIRE (1986). For more on the difficulty of separating rights from remedies in the context of constitutional law, see Daryl J. Levinson, Rights Essentialism and Remedial Equilibration, 99 COLUM. L. REV. 857 (1999).

141. For more on judicial backlash and human rights law in a different context, see Andrew T. Guzman & Katerina Linos, Human Rights Backsliding, 102 CALIF. L. REV. 603 (2014).

142. Matt Chorley & James Slack, Britain Could Leave European Convention on Human Rights if it is the Only Way to Kick Out Foreign Criminals, Cameron Vows, DAILY MAIL, June 3, 2015 (“Our plans set out in our manifesto do not involve us leaving the European Convention on Human Rights. But if we can’t achieve what we need . . . when we’ve got these foreign criminals committing offence after offence and we can’t send them home because of their right to a family life, that needs to change. I rule out absolutely nothing in getting that done.”). This is not the first time that Cameron has made this threat. Only a few days after the ECtHR banned the deportation of Abu Qatada, in Othman v. United Kingdom, 2012-I Eur. Ct. H.R. 159, Cameron called for the ECtHR to restrict its power to overrule national judgments on immigration matters. Stephen Castle, Cameron Calls for European Court to Limit Its Reach, N.Y. TIMES, Jan. 25, 2012.
B. Second Method—Adopt the Exclusionist Tradition.

Since, as a practical matter, walls and interdiction are the same, a court that elects this tradition will use a case that bears on a wall to revisit and to pressure the interdiction precedent: if a State’s border is final and the wall simply concretizes the border, then the wall is also a final obstacle to entry. Getting close to the wall does not trigger human rights jurisdiction. Instead jurisdiction is softened back from the interdiction precedent (jurisdiction initiated through physicality grounded in contact) to requiring territorial presence (jurisdiction associated with geography). To substantiate this approach, the court could refer to precedents coming out of international courts adjudicating the two most notorious walls in international law: the Israeli Security Fence and the Berlin Wall. These precedents grant the State the power to build a wall on its own territory, defining to which persons it owes obligations and to whom it does not.

In its Advisory Opinion, The Legal Consequences of the Construction of a Wall in the Occupied Palestinian Territory, the International Court of Justice (ICJ) was called to review the legality of the wall built by Israel. The Court held the Israeli Security Fence illegal per se. Yet it expressly limited its analysis to those parts of the wall constructed outside the territory of Israel. By implication, the ICJ considered the parts of the wall built within the State to be necessarily lawful and without limitations vis-à-vis human rights jurisdiction.

Further, in a series of cases that involved shootings on the Berlin Wall, both the ECtHR and the UNHRC suggested that even if a wall built on a State’s

---

143. Legal Consequences of the Construction of a Wall in the Occupied Palestinian Territory, Advisory Opinion, 2004 I.C.J. 136, ¶ 121 (July 9, 2004).
144. Id. at 141 (the ICJ was asked to render an opinion on the question: “What are the legal consequences arising from the construction of the wall being built by Israel, the occupying Power, in the Occupied Palestinian territory?”).
145. Id. ¶ 121 (the very construction of the barrier on occupied territory violated international law because it was erected to “to create a ‘fait accompli’ on the ground that could well become permanent, in which case . . . it would be tantamount to de facto annexation [of Palestinian land].”). As such, Israel was under an obligation to cease construction works, to dismantle the structure already built, to repeal or render ineffective all legislative and regulatory acts relating hitherto, and to make reparations for all damages caused by the construction of the barrier.
146. Id. ¶ 67 (explaining that “some parts of the complex are being built, or are planned to be built, on the territory of Israel itself,” but not considering that “it [wa]s called upon to examine the legal consequences arising from the construction of those parts of the wall”).
147. Indeed in examining the evolution of the jurisprudence challenging the barrier both before the ICJ and the Israeli High Court, Yishai Blank found that “no legal argument was made against a barrier which would have been erected on the internationally recognized border of Israel.” Blank, Legalizing the Barrier, supra note 30 at 311. Michael Sfara argues that “[i]n fact, it would have been possible to build a wall or a fence, even a tech with crocodiles, without raising any legal difficulty, especially not an international one. The simple and legal way would have been to construct the ‘separation barrier’ right on the Green Line,” SHAUL ARIELI & MICHAEL SFARD, HOMAH U’MEHDAL [The Wall of Folly] 145 (2008) (Isr.).
own border does infringe on an important human right, the wall may still withstand a legal challenge if it serves a legitimate aim “to protect the border” and the aim is “limited” and “respect[s] the need to preserve human life.”

In deferring to a State’s power to erect a wall on its own territory, the decisions on both the Israeli Security Fence and the Berlin Wall are in line with larger international legal orthodoxy. With some important exceptions, the international order is still centered on the geography of the State. To avoid questioning Statehood, and thereby the larger legal and political order, all key players of the regime (international courts, treaties and doctrines, and


149. Streletz v. Germany, 2001-II Eur. Ct. H.R. 409, ¶ 71 (noting that the aim of the Berlin Wall was “to protect the border between the two German States ‘at all costs’ in order to preserve the GDR’s existence, which was threatened by the massive exodus of its own population.”). This aim, the judges held, “must be limited.” Id. ¶ 72. Above all it must “respect the need to preserve human life,” such that it cannot have an “indiscriminate effect” or a categorical nature to “annihilate border violators . . . and protect the border at all costs.” Id. ¶¶ 72–73.

150. The traditional example is Island of Palmas (U.S. v. Neth.), 2 R.I.A.A. 829, 838 (Perm. Ct. Arb. 1928) (territorial sovereignty is “the point of departure in settling most questions that concern international relations”). For a more recent example, see Military and Paramilitary Activities in and Against Nicaragua (Nicar. v. U.S.), 1986 I.C.J. 14, ¶ 263 (June 27) (affirming “the fundamental principle of State sovereignty, on which the whole of international law rests, and the freedom of choice of the political, social, economic and cultural system of a State.”).

151. See, e.g., Treaty of Lausanne (Frontier Between Turkey and Iraq), Advisory Opinion, 1925 P.C.I.J. (ser. B) No. 12, ¶ 53 (Nov. 21) (“The very nature of a frontier and of any convention designed to establish frontiers between two countries imports that a frontier must constitute a definite boundary line throughout its length.”); Island of Palmas, 2 R.I.A.A. at 870 (“International law . . . has the object of assuring the coexistence of different interests which are worthy of legal protection. If . . . only one of two conflicting interests is to prevail [the case involves a territorial conflict] . . . the interest which involves the maintenance of a state of things having offered at the critical time to the inhabitants of the disputed territory and to other States . . . ought, in doubt, to prevail . . . .”); Temple of Preah Vihear (Cambodia v. Thai.), 1962 I.C.J. 6 (June 15) (ICJ held that when two countries establish a frontier between themselves one of the primary objectives is to achieve finality and stability); Beagle Channel (Arg. v. Chile), 11 R.I.A.A. 53, 89 (Ct. Arb. 1977) (the Arbitration Tribunal observed in respect of the Argentina-Chile Boundary Treaty of 1881 that “the regime set up by the Treaty . . . was meant thenceforth to govern the question of boundaries and title to territory, and that it was meant to be definitive, final and complete”); Nicaragua v. United States, 1986 I.C.J. ¶ 55 (“No State or group of States has the right to intervene, directly or indirectly, for any reason whatever, in the internal or external affairs of any other State.”) (quoting Article 18 of the Organization of American States Charter).

152. See, e.g., U.N. Charter art. 2, para. 1 (sovereign equality of the UN members); id. art. 2, para. 4 (the prohibition on the threat or use of force “against the territorial integrity or political independence of any state”); id. art. 2, para. 7 (the reserve domain of domestic jurisdiction into which intervention is not permitted); Vienna Convention on Succession of States in Respect of Treaties art. 11, Aug. 23, 1978, 1946 U.N.T.S. 3 (“A succession of States does not as such affect: (a) a boundary established by a treaty; or (b) obligations and rights established by a treaty and relating to the regime of a boundary.”); Vienna Convention on the Law of Treaties art. 62, para. 2, May 23, 1969, 1155 U.N.T.S. 331 (“A fundamental change of circumstances may not be invoked as a ground
prominent scholars) prioritize the stability of borders and disfavor the creation of new territorial and boundary difficulties.

In this second approach, and in contrast to the first approach, the wall acts as a wall—a barrier rather than a bridge. The border is sacrosanct: it marks the precise area over which a State may exercise absolute dominion. A State can freely choose to build a wall at its borders. Such a wall only serves to literalize the border—from a legal perspective, nothing has happened as a result of building a fence. If the State’s territory is not disputed, then its border is final and complete, and the wall is a definitive block to entrance. Jurisdiction is squashed back to geography rooted in territory (human rights obligations are strictly territorial). Merely getting close to the State is not the same as getting into the State. Sovereignty now means a space outside human rights law: a State is only bound by obligations to which it consented via positive law making.

But, much like the first approach, this approach also cannot be normatively justified if taken to its ultimate conclusion. If courts remain deferential to border walls as part of an immigration control policy, more States may gravitate toward building physical walls as their preferred strategy to control immigration. And there is no reason that States will restrict themselves to building walls only on the very border itself rather than also expanding walls to more creative locations. In fact, just recently the United Kingdom offered to give France an eleven-foot steel fence that had been used to protect world leaders at the NATO summit. The United Kingdom suggested that the fence could be used to stop hundreds of migrants from countries such as Afghanistan, Eritrea, and Ethiopia for terminating or withdrawing from a treaty: (a) if the treaty establishes a boundary . . . .”). For doctrines, see for example Frontier Dispute (Burk. Faso/Mali), 1986 I.C.J. 554 §§ 24–25 (Dec. 22) (discussing the importance of uti possidetis juris (“as you possess under law”) that transforms the former boundary of a colony into an international frontier at the moment it becomes independent, because “the maintenance of the territorial status quo in Africa is . . . the wisest course . . . . The essential requirement of stability in order to survive”).


154. See, e.g., Customs Régime Between Germany and Austria, Advisory Opinion, 1931 P.C.I.J. (ser. A/B) No. 41, at 45 (Sept. 5) (State’s independence means that it has the “sole right of decision in all matters economic, political, financial or other”); Military and Paramilitary Activities in and Against Nicaragua (Nicar. v. U.S.), 1986 I.C.J. 14, ¶ 55 (June 27) (“No State or group of States has the right to intervene, directly or indirectly, for any reason whatever, in the internal or external affairs of any other State.”).


Taken to the extreme, the alignment of norms with the supremacy of State sovereignty means, in effect, deferring to developed States’ interests and sacrificing our evolving norms concerning the universal and fundamental dignity of every individual. All that will be left from the universality of human rights will be many smaller spheres of specific rights regimes. Different people will be subject to different rights, depending on their geography.

C. Third Method—Adopt the Compromise Approach that Human Rights Courts Tailored Between Universality and Exclusion and that is Structured Through Territory.

As opposed to the earlier two approaches, a court that adopts this compromise will differentiate between the two strategies of front-end immigration exclusion, interdiction and walls, and use the distinction to limit the procedural rules that emerge out of interdiction. This keeps intact the normative force of *Jamaa*, which guarantees access to individuals at sea, but, at the same time, maintains the State’s right to exclude non-nationals on land and next to a wall. A court would do so by using the wall to balance between these two conflicting policy interests of the individual (universality) and the State (exclusion): on the external side of the wall, jurisdiction is softened back to geography. On the internal side the wall, individuals may have expansive rights independent of State consent. At least when it comes to walls, jurisdiction retreats back to physicality grounded in geography. Proximity would no longer denote rights.

Such a court would differentiate between liquid and solid barriers by drawing on formally available legal distinctions. For example, the court could defer to legal precedents emerging from the law of the sea that prioritize land over the seas, so that what applies on the sea does not extend to the land.\footnote{Existing jurisdiction that deals with law of the sea repeatedly affirms that “the land dominates the sea.” Territorial and Maritime Dispute Between Nicaragua and Honduras in the Caribbean Sea (Nicar. v. Hond.), 2007 I.C.J. 659, ¶ 113 (Oct. 8); see also Maritime Delimitation in the Black Sea (Romania v. Ukraine), Judgment of 3 February 2009, ICJ Reports 2009, 61, para. 77; Aegean Sea Continental Shelf (Greece v. Turk.), 1978 I.C.J. 3, ¶ 86 (Dec. 19); North Sea Continental Shelf (Ger./Den.; Ger./Neth.), 1969 I.C.J. 51, ¶ 96 (Feb. 20).} Alternatively, the court could also cite the intersection of two sovereign States at an international border, as opposed to the lack of any State authority upon the high seas. In the case of migration by land, the claim that a host State is not responsible for protection duties is also a claim that another State is responsible.
This is not the case in an interdiction scenario (migration by sea), which usually occurs in international waters where there is no other responsible party.

Finally, the court might also rely on the relative passivity of border walls as a method of exclusion, as compared to interdiction. Border walls are passive elements in two ways. First, a wall prevents a would-be immigrant from doing a specific act (getting in) but leaves her other options open, while an interdiction coerces a would-be immigrant to do a specific act (turn around).\textsuperscript{159} Second, once the wall is constructed, exclusion no longer requires a new exercise of agency on the part of the State: a wall can restrain entrance even years after it was built.\textsuperscript{160}

In this scenario, like the second approach but in contrast to the first approach, the wall truly acts as a wall—a barrier, rather than a bridge. Jurisdiction is aligned with territory (human rights obligations are strictly territorial) and proximity is not the same as getting into the State. Sovereignty again means a space outside human rights law as a State is only bound to what it has consented to.

This third approach, however, also fails the normative test. Translating this approach into actual practice results in too many distinctions that do not make sense and rulings that may lead to perverse effects. For example, a host State would not owe obligations to an individual that starves while waiting on the other side of a wall, but if she climbs up and sits on top of the fence or attaches herself to the fence in some hazardous manner and refuses to leave, then a destination State that removes her will owe such duties. The incentive structure, therefore, would be for an individual to risk herself precisely so that the host State would be forced into action and such action would trigger jurisdiction.\textsuperscript{161}

In addition, the regime would tempt (or perhaps even require) individuals to take steps that are dangerous before they can access rights. At least when it comes to asylum seekers, this means that the protective regime itself adds on an actual life-threatening danger for those who are at least allegedly already fleeing persecution, before their claim can even be heard.

This Article does not suggest how a human rights court will choose between these three approaches to the regulation of a wall. But this choice brings us to the unresolved end-point of the arrangement that courts worked out between universality and exclusion and that is conditioned on territory: each of the three different approaches to regulation correlates to a drastically different vision of borders and sovereignty in the world, and none of them can be

\textsuperscript{159} My discussion here builds on David Miller, \textit{Why Immigration Controls are Not Coercive: A Reply to Arash Abizadeh}, 38 \textit{POL. THEORY} 111 (2010), http://cesem.ku.dk/papers/whyimmigrationcontrolsarenotcoercivedavidmiller.pdf.

\textsuperscript{160} In fact, the UNHCR in a different context (in-country interception at airports) suggested already that “there is a distinction . . . between ‘the active interdiction or interception of persons seeking refuge from persecution’ . . . and ‘passive regimes, such as visa and carrier sanctions.’” European Roma Rights Centre v. Immigration Officer at Prague Airport, [2003] EWCA (Civ) 666, [48] (Eng.).

\textsuperscript{161} For how this incentive system operates in the context of interdiction, see Mann, \textit{Dialectic of Transnationalism}, supra note 110.
continuously defended. Thus, the story of walls presents a court with a choice between utopia disconnected from reality, a walled world, or a nonsensical order that incentivizes dangerous behavior. In other words, it is a story of disappointment.

III.

A CASE STUDY: THE ISRAEL-Egypt WALL

We still do not know how an international human rights court will adjudicate a wall erected as an immigration control strategy. But a case study from a national court may be illuminating—the Israeli Supreme Court’s response to the fence Israel built on its border with Egypt. I chose Israel because it is “the only Western country that has a relatively long land border with Africa” and that has built a wall that runs all throughout the length of the border. I examined all four cases that came before the Israeli Supreme Court in regards to this fence. Using these cases, I demonstrate how, when faced with a decision regarding how to regulate the Israel-Egypt Fence, the Israeli Supreme Court adopted the statist tradition and structured the border fence as the point of equilibrium between universality and exclusion.

The only time that the Israeli Supreme Court was called to directly review the Israel-Egypt fence was in Anu Plitim v. Ehud Barak -Minister of Defense. The case concerned the first group of people from the African Continent—eighteen men, two women, and a child—who brought a case on the Israel-Egypt fence after its completion. It offers a rare judicial review of the fence in real time: the Court was asked to rule on the situation as the plaintiffs were begging for their lives under the unforgiving desert sun on the Egyptian side of the wall.

Both sides agreed that Israel, as a sovereign State, had a right to build a wall on its territory. Their dispute was over the function of the wall in


163. For a description of the wall, see Gidon Ben-zvi, Israel Completes 245 Mile, NIS 1.6 Billion Security Fence Along Sinai Border with Egypt, ALGEMEINER, Dec. 4, 2013, (Isr.), and Amos Harel, On Israel-Egypt Border, Best Defense is a Good Fence, HAARETZ, Nov. 13, 2011 (Isr.).


165. Id. Before the fence was completed, and during the Mubarak regime in Egypt, Israel’s policy was to intercept asylum seekers after their entry to Israel and immediately expel them back to Egypt without any guarantee as to the safety of the returnees, known as the “Hot Return Procedure.” Following a petition to the Supreme Court the government eventually announced that due to the change of regime in Egypt, the use of this practice had ceased. See HCJ 7302/07 Hotline for Migrant Workers v. Minister of Defence (July 7, 2011) (Isr.).


relationship to the application of non-refoulement. The petitioners adopted the human rights (universalist) tradition: they referred to the Jamaa ruling and the UNHCR to argue that non-refoulement is engaged also on the external side of the fence.168 Otherwise, the petitioners explained, the fence transforms Israel’s legal obligations under the Refugee Convention “into a legal ‘dead letter.’”169

The State, in turn, adopted the exclusionist (statist) tradition to the wall: Israel, a sovereign State, had the right to decide who was entitled to enter its territory, and conversely, it was not obliged to act with respect to aliens who were located outside its actual territory, effectively marked by the new fence.170 Here, the fence was precisely designed to prevent infiltration into Israel: it is a final and complete constraint on entrance. In fact, the attorney general added, the fence does not have gates, which means that admittance of the group is not even physically possible.171

The Israeli Supreme Court never decided between the universalist and the exclusionist traditions. Instead, the judges waited for three days and in that time, the government resolved the matter. Israeli soldiers cut the fence, crossed to the Egyptian side and admitted into Israel the two women and the child as a humanitarian gesture.172 They put the eighteen men on an Egyptian van that drove away.173 The soldiers then stitched the fence back together again.174 Thereafter, the eighteen African men were never heard from again.

Following the acts of the Israeli government, the judges, in a unanimous decision, dismissed the case: “3 members of the group were allowed to enter Israel . . . . The rest of the group members, 18 persons, left their whereabouts near the fence and turned back . . . . the petition became redundant.”175 And so by waiting, and without choosing between universality and exclusion, the Israeli Supreme Court allowed the fence to act as a final barrier to entry as a matter of fact. The fence was broken—and then immediately sealed—for reasons that have to do with compassion and that exist outside the normative realm of the law.

While not reviewing the legality of the Israel-Egypt fence directly, the Supreme Court also dealt with the fence in two more decisions: Adam v. Knesset176 (Adam) and Gebreselaissie v. Israeli Government177 (Gebreselaissie).

168. Id.
169. Id.
170. Id.
171. Id.
172. Id.
173. For a discussion of this incident, see Shatz, supra note 166.
174. Id.
177. HCJ 8425/13 Gebreselaissie v. Israeli Government (Sept. 28, 2014), RefWorld (unofficial
Both cases examined the constitutionality of two consecutive Amendments and Temporary Provisions to the Prevention of Infiltration Law (Offences and Jurisdiction), which authorized the State to hold in detention illegal immigrants (statutorily termed “infiltrators”), whom it cannot expel, in order to prevent settlement in Israel and to deter future arrivals.178

The Third Amendment to the Prevention of Infiltration Law, reviewed by the Adam court, allowed the imprisonment of infiltrators for a period of up to three years without trial. It was struck down because it ran contrary to the Basic Law on Human Dignity and Freedom.179 The Fourth Amendment to the Prevention of Infiltration Law, in turn, was passed by the government soon after the Supreme Court found the Third Amendment unconstitutional. It limited the maximum extent of detention to one year, and applied this sanction only to new “infiltrators” who would enter the country from then on. It also established a new “infiltrator staying facility,” where the State could compel undocumented immigrants, not liable for deportation, to live indefinitely.180 Then, in Gebreselaissie, the Court also struck down this newer amendment, triggering another political and legal earthquake—the Supreme Court had never overturned a law twice—because the new legislation had failed to comply with the constitutional guidelines set out in its first opinion.181 The government changed the name of the facility, but it did not alter the reality of imprisonment.182

In both decisions, the Israel-Egypt fence provided the Court with an alternative to detention. In the words of Justice Edna Arbel, writing the main opinion in Adam: “there is a fair probability that it would have been possible to manage with a less injurious means in the form of the border fence between Israel and Egypt.”183
For Justice Arbel, the fence on the border with Egypt is efficacious. While she consents that proving causality between construction of the fence and reduction of entrance to Israel is difficult,\(^\text{184}\) she writes:

It was not without good reason that the government decided to invest enormous resources in the construction of the fence. . . . \(\text{[I]}\)t may be assumed that the border fence may help significantly to reduce the phenomenon of infiltration, whether because of the physical barrier or because of the need to invest greater resources in order to enter Israel unlawfully, in such manner that the investment will not be worthwhile for the infiltrator or for his smugglers. . . . To this it should be added that there are additional means that [a] state can employ in order to enhance the efficiency of the physical barrier, such as electronic means and so forth.\(^\text{185}\)

And, at the same time, the fence also carries no legal limitations. The determinative question, in Justice Arbel’s analysis, is empirical, and the normative inquiry follows the quantitative data: so long as the fence successfully blocks “infiltrators” and the numbers of those who do manage to get into Israel is small, Justice Arbel explains, the “detention of asylum seekers for the purpose of deterring additional asylum seekers from arriving in the state” is not constitutional.\(^\text{186}\) If numbers are low, Justice Arbel continues, a detention that deprives a person of her liberty “makes a moral stain on the network of human values espoused by Israeli society.”\(^\text{187}\) But, if the numbers increase, then an administrative detention for purposes of deterrence could in fact become constitutional: “in an extreme situation in which the purpose becomes extremely vital for the survival of the state and . . . [to] maintain its most basic interests, it may be possible to justify this purpose [detention], notwithstanding the grave and forceful injury to the infiltrator’s liberty.”\(^\text{188}\)

\(^\text{184}\). HCJ 7146/12 Adam v. Knesset, ¶¶ 98–101, 103, 108 (Arbel, J.) (noting “it is unclear whether” the normative framework, the Amendment, or the wall was the “dominant factor in the dramatic reduction in the number of infiltrators entering Israel.”); see id. ¶¶ 1–3, 5–6 (discussing more on the argument of causality in the case); but see id. ¶¶ 25, 38 (Justice Vogelman doubting Arbel’s causality). For more on the argument of causality in the case, see id. ¶ 1 (Amit, J.); id. ¶ 6 (Hendel, J.); id. (Grunis, J.) ¶ 2–3, 5.

\(^\text{185}\). Id. ¶ 103.

\(^\text{186}\). Id. ¶ 92.

\(^\text{187}\). Id. ¶ 114 (internal quotation marks omitted).

\(^\text{188}\). Id. ¶ 93; see also id. ¶ 2 (Justice Amit stating “As emerges from the figures before us, as of today, the number of infiltrators who have penetrated Israel in recent years totals some 65,000, close to one percent of the population in Israel. . . . [I]t could be argued that one percent of the population is a number that an enlightened and economically strong country such as the State of Israel can and should bear . . . . Such is the situation today . . . . But what of the future? . . . . What is the numerical ‘red line’ that a country can bear without concern of tangible injury to its sovereignty, its character, its national identity, its cultural and social profile, the structure of its population and its diverse features, and without fear for its resilience and fear of reaching [a] breaking point in terms of congestion, welfare, and the economy, internal security and public order? Naturally, the State of Israel, like any other enlightened country, cannot absorb all the unfortunates, the oppressed and the persecuted throughout the world and in Africa. . . . In balancing basic rights with other basic rights, or with vital state interests, therefore, we must be cognizant of the figures, estimates, and forecasts. There are situations in which ‘quantity means quality’ . . . . As noted, this is not currently the
The Israel-Egypt fence emerges in Justice Arbel’s opinion as an equilibrium point. It stabilizes the two legal traditions, universalist and exclusionist, by dividing them geographically. On the Israeli side, “the stranger, infiltrator, or refugee who has entered Israel” has rights because a “person’s liberty is a right that accompanies him wherever he goes . . . whether he is present in a place where he has been permitted to be present or has entered a place he has been forbidden to enter.”\(^{189}\) Therefore, even an economic migrant who breached the State’s immigration law but is already inside the country—or, is standing ‘in front of our eyes’—is guaranteed protection. On the Egyptian side, by contrast, Israel owes no protective duties. A person who was unable to cross the fence has no face and no rights.\(^{190}\) And what differentiates the State’s protective duties and lack thereof? One’s location vis-à-vis the fence. Justice Arbel concludes that a barrier requires considerable financial resources, but “the protection of human rights costs money, and a society that respects human rights must be willing to bear the financial burden.”\(^{191}\) The fence, which itself lacks any normative significance, becomes the essence of human rights protection.

This is where the analysis ends: the Israeli Supreme Court adopted the exclusionist tradition and permitted Israel to build a fence on its own border, and then used this fence to square the circle. The fence stabilized the conflict between universality and exclusion by dividing them geographically. The long-term stability of this equilibrium is unclear, and it may be more or less stable across different States depending on both empirical factors (for example, territorial location of the State, level of migration into the State, etc.) and legal obligations (such as whether the State has ratified the Optional Protocol to the ICCPR or is subject to ECtHR jurisdiction.)\(^{192}\) Nonetheless my mapping is a warning; if, as the Israeli Supreme Court said, immigration walls are both an effective immigration tool and unencumbered by constraints vis-à-vis human rights, then States will build walls. They will build them on the State’s borders and around ports and maybe in even more locations that we can predict at

\(^{189.}\) Id. ¶ 113.

\(^{190.}\) The fence also acts as the point of equilibrium and becomes the essence of human rights in the second case, HCJ 8425/13 Gebreselaissie v. Israeli Government (Sept. 28, 2014), RefWorld (unofficial translation) (Isr.), http://www.refworld.org/cgi-bin/texis/vtx/rwmain/opendocpdf.pdf?redoc=y&docid=54e607184. See, inter alia, id. ¶ 9 (Amit, J.) (“The State’s responsibility to these individuals that entered its territory is not the same responsibility to who are not in its borders . . . . even when we are dealing with uninvited guests.”). But, importantly, “the State is permitted to act . . . to prevent . . . the arrival of additional uninvited guests.” Id. How? Amit answers: “by placing . . . the physical barrier of the fence.” Id.

\(^{191.}\) HCJ 7146/12 Adam v. Knesset, ¶ 103.

\(^{192.}\) For example, the United States and Israel have not ratified the Optional Protocol and thus they did not consent to the jurisdiction of the UNHRC to entertain individual complaints. This means that an author in the United States or Israel cannot challenge the U.S.-Mexico wall or the Israel-Egypt fence before the UNHRC.
present. What will then be left of the universality of human rights law will be minimized into one’s location vis-à-vis the fence: on one side, the kingdom is given; on the other side, the desert sun. This is possibly best summed up in two statements by Israeli Prime Minister Benjamin Netanyahu in regards to the Israel-Egypt fence: “We do not intend to stop refugees fleeing for their lives,” he said. “[W]e allow them in and will continue to do so.” But, Netanyahu also added elsewhere: “It is important that everyone understand that Israel is no longer a destination for infiltrators.” “We are determined to stop the flow of infiltration. We built a fence for this purpose.” And so asylum seekers who can satisfy the access criteria to trigger jurisdiction have expansive protections in Israel, but most cannot. In other words, they have rights but not protection.

CONCLUSION

Let me now go back to where I began: the tension between universality and exclusion. In the past ten years, working piecemeal, moving from one decision to another, the UNHRC and the ECtHR have worked out a path-dependent compromise between the two legal traditions that leans in favor of universality. They reached this compromise by using territory, or the location of the individual plaintiff, to map the individual’s access: a non-national that establishes physicality—associated with territory, contact, and maybe even proximity—activates norms of protection. These norms are increasingly absolute and inflexible, go beyond the traditional five grounds of protection, and exist outside a State’s interests or constraints (universality). In contrast, a non-national who fails to establish physical presence has no rights to which the State has not willingly consented to (exclusion).

The resulting compromise makes judiciable a process that does not easily lend itself to international regulation: it substitutes complex multi-party political negotiation about who deserves asylum (who is most vulnerable) and from what


196. My point here is that who gets to be considered for human rights protection is arbitrarily decided. The test for protection is only invoked for an individual who successfully established physicality—associated with territory, contact, and maybe even proximity. An individual who scaled a border wall, for example, may still be denied protection if she does not meet the criteria for protection (for instance, faces persecution on account of the five grounds). But it is worth remembering that many individuals who scale the wall disappear and stay without entitlements, or else root themselves socially in the new host state and in time could secure entitlements through these social ties.
State (who is most capable) with a set of arbitrary rules that ask a court only to locate the plaintiff and to answer relatively simple questions.\textsuperscript{197} In doing so, the compromise allows international courts to decide human rights obligations that were never resolved politically.\textsuperscript{198}

But the outcome of this compromise—who benefits and who is hurt—is arbitrary, thus making the regime radically unstable. There are two key interests involved in cases that bear on immigration. First, for the individual: what is the nature of the misery that should be alleviated, and how and by whom should such misery be assessed?\textsuperscript{199} Second, for the State, how to distribute protective duties, and how and by whom such duties should be determined? However, territory, I suggested in this article, is an arbitrary legal category from the perspective of both these two interests.

From the perspective of the non-national, the compromise collapses the whole account of the individual’s interests into a single question: whether she is able to meet the regime’s condition of access. But territory is a poor proxy for who is most needy: it does not take into consideration the substantive interests of the individual or the nature of her predicament. From the perspective of the host State, in turn, the regime privileges a single normatively random category: territorial location of the plaintiff vis-à-vis the State or its agents. Alas, territory is also a bad proxy for who has a lower cost of absorption of nonnationals—it leaves out of the protective equation the State’s real constraints (such as size, Gross Domestic Product, numbers of non-nationals coming in, etc.) and aggregated efforts (procedurally and substantively) to deal with non-nationals at a particular moment. Territory, in other words, says nothing about who is most vulnerable and who is most capable of helping.

Further, the compromise is also fundamentally unjust. Protection is conditioned upon establishing physical presence. This dynamic disproportionally favors those individuals with capacity—defined in terms of luck, resources or physical abilities—who can get close enough to the State or its agents. They are protected because they are strong, fortunate, or both, not necessarily because of the substantive causes of their misery.

\textsuperscript{197} For more on the way in which human rights obscure political inequality, see, for example, Richard Thompson Ford, \textit{Rights Gone Wrong: How Law Corrupts the Struggle for Equality} 21 (2012); David Kennedy, \textit{The Dark Sides of Virtue: Reassessing International Humanitarianism} 13 (2004); Cass R. Sunstein, \textit{Rights and Their Critics}, 70 Notre Dame L. Rev. 727, 743–44 (1995).

\textsuperscript{198} Indeed it is hard to imagine a situation where the UNHRC carries the authority and power to hold, for example, that Canada ought to be responsible for the protection of twelve percent of women who suffer sex offenses and flee Central America. This is not to say that international courts will not play any role in a new and revised regime. But their role is likely to be narrower: they may only be called into action in cases that bear on the most extreme forms of torture and degrading behavior. For a similar argument, see James C. Hathaway, Leveraging Asylum, 45 Tex. Int’l L.J. 503 (2010).

\textsuperscript{199} Or, should the Refugee Convention be changed to reflect the changing circumstances in the world? And, if yes, then how? And who decides?
Walls take this compromise to its perverse extreme and so make concrete its intractability. The correlation between protection (access) and territoriality invites States that seek to maintain their exclusionary powers to erect additional layers of walls to prevent would-be immigrants from getting close enough to the actual border to trigger proximity-based human rights protections. And so the question of “who can establish physical presence” becomes “who can scale walls that are almost impassable.” The answer is often strong, fast individuals with an aptitude for risky behavior; in other words, young men. But they are rewarded by the regime only after they have risked themselves in traversing an ever-growing numbers of barriers; and, if they endure. The result is reminiscent of a gladiatorial fight: those savage and bloody combats of the slave against other men, tigers, and armed chariots in old Rome, which, if victory were achieved, could free the slave. Today those who survive the terror of the fight—the traversing of the fence—are welcome to enter the kingdom. And we, those who are lucky enough to be in the kingdom, are watching.

This dynamic is perhaps most readily visible around the two fences that Spain built in North Morocco. On a single day in May 2014, between one and two thousand Sub-Saharan migrants rushed the razor-wire fences in Melilla—"actually three fences, two 20 feet high and a middle one that is slightly lower."200 About 450 of the migrants managed to make it over the towering fence: only two of them were women.201 Those who made it to the other side "kissed the ground" and yelled "with joy as they touched Spanish soil."202 They were jubilant because at the very moment that their legs left one side of the fence and touched the ground on the other side, Spanish protective rights were triggered.203 That moment alarmed Spain, as the contact activated expansive duties, including, at a minimum, providing each of these migrants with an individual status determination before deportation, and arranging much more substantive accommodations for those who qualify as refugees.204 But Spain was already “at its limit” in terms of capacity to absorb new arrivals, according to its minister of the interior.205

200. Gall, supra note 3.
201. Id.
203. Id.; Suzanne Daley, As Africans Surge to Europe’s Door, Spain Locks Down, N.Y. TIMES, Feb. 27, 2014. (Most of the people who made it into Spanish territory “will probably spend a year or more in the immigration center as their applications for asylum are processed. Few will get such status. But most will end up transferred to the mainland before being handed an order to leave Spain. Most cannot be deported because Spain does not have treaties with many of the countries they come from. . . . [M]any of those who make it to Melilla and Ceuta will be largely free to remain in Spain or other European nations.”).
204. Under the Schengen agreement and the Dublin regulation (a building block of Schengen), migrants that enter Europe must be processed by the country through which they enter. See Illegal Immigration Europe’s Huddled Masses, ECONOMIST, Aug. 16, 2014.
In this episode, what made the difference between those non-nationals that benefited and those that were hurt by the legal regime was physical ability (the power to climb up the first fence, jump from one fence to the other without a crashing fall, and then climb down the third fence, all in “one minute [thirty]”) and luck (whether the individual happened to stand next to the one chunk of the fence that crashed down that day, or whether he survived the jumps more or less intact). There was no consideration of an individual’s worthiness for protection or of preferences that could or should be shown toward particular groups. For example, are men more worthy of protection than women? Young athletes more than the elderly? Or, from the other direction, are other EU states better equipped to handle asylum requests from 400 individuals than Spain—a country undergoing a dramatic economic crisis, and that has already absorbed many waves of migrants?

Because it is all about the wall, the migrants and asylum-seekers risk all they have into scaling the physical barrier. They attack the fences again and again until they either successfully cross over or fatally injure their bodies. “I was thinking that I was finally in Spain,” explained a sixteen-year-old boy from Niger who had been “violently thrown back” after successfully scaling all three fences but failing to pass the last line of police. Nevertheless, he will try again: “I’m not going to go back now to Niger, where there is nothing to do and no work, when every time I now wake up I can at least already see Europe.” At the same time, the Spanish government is taking increasingly elaborate steps to fortify the fence, erecting a growing numbers of concentric barriers. In 1998, Spain built the first fence. Then in 2005, it enhanced this single barrier with two more fences. In 2013, Spain permanently reintroduced razor-sharp barbed wire to the top of the border fences (it had been installed in the past but was removed because it inflicted serious bodily harm). A year later, the state added what it calls an “operational border” to the fixed border—set wherever the last line of police security stands—arguing that even if individuals crossed the three fences they are still not on Spanish territory until

206. “‘You have to get over in one minute 30,’ said Nili Onana, a basketball player from Cameroon, who made it over in a wave on May 28 and was interviewed in a short-stay center for immigrants in Melilla.” Gall, supra note 3.
209. Id.
211. Tremlett, supra note 15.
212. The Spanish government first introduced the razor-sharp barbed wire in 2005 but it had mostly been removed from the top of the fence after causing serious injuries to migrants as they tried to cross the border. See Paul Hamilos, Razor Wire on Fence Dividing Melilla from Morocco Condemned as Inhumane, GUARDIAN, Nov. 1, 2013.
they have crossed the “operational border.”213 And, more recently still, operating in cooperation with Spain, Morocco began building an extra ditch and fence, “crowned with concertina wire about 500 meters (almost 1,640 feet) from the existing Spanish fences, further extending the obstacle course for the migrants.”214

And the human rights community? They also focus on the fence. Important human rights groups are now preparing a case against Spain that challenges the location of the fence. Their central premise is that when Spain began erecting the first fence, Morocco insisted that no Spanish construction machinery operate on Moroccan soil. And so, they argue, even the first of the series of the three fences that Spain erected actually rests inside Spain.215 The implication is that just reaching the outer perimeter of the enclave may mean that the migrants have already entered Europe. The Spanish government’s delegate to Melilla aptly summarized this argument when he responded to the case by saying that if the human rights group is successful then “just by touching the first fence” a person would have “already reached Spain.”216 Or, in the terminology used in this Article, jurisdiction is attached to proximity to the fence: getting close to the fence is as good as crossing over. Spain will owe protective duties on both sides of its border. With an estimated 80,000 migrants and asylum seekers that have already headed for Spain’s two exclaves by the middle of 2014,217 this could exponentially expand the numbers entitled to legal counsel, asylum claims, or proper deportation proceedings from Spain.

Whether we will keep moving toward a world of walls or instead work out another uneasy compromise between universality (human rights) and exclusion (sovereignty) remains to be seen. One thing is sure: today there are more than fifty million people displaced.218 And the desert is getting even drier and the kingdom more lavish still. The international response, in the words of Ban Ki-moon, the UN Secretary-General, is to place human rights “at the centre”219 of the efforts to meet this mammoth challenge of displacement. But, as this Article argues, this approach is profoundly flawed: it is conditioned upon a compromise that is based on territory — where an individual is located — alas territory is an arbitrary legal category. And so the legal victory of the human rights tradition has resulted in a practical defeat for both individuals and States. It is a deeply unjust regime.

213. Minder, At Spanish Enclave, supra note 208.
214. Id. “Spain talks about having great cooperation with Morocco,” said a founder of a human rights organization that is challenging Spain. Id. “But this cooperation is really just about paying Morocco to do the dirty work for Spain . . . .” Id.
215. Id.
216. Id.
217. Minder, Spain Struggles, supra note 6.
218. Harriet Sherwood, Global Refugee Figure Passes 50m for First Time Since Second World War, GUARDIAN, June 20, 2014.
Chinese Reception and Transplantation of Western Contract Law

Wang Jingen* & Larry A. DiMatteo**

Introduction ..........................................................................................................................45
I. Borrowing: Reception and Transplantation in China ..........................................................46
   A. China’s Double Transplantation and Resulting Problems .........................................51
   B. Comparative Law as an Avenue to Law Reform .........................................................53
      1. Perils and Virtues of the Comparative Law Methodology .........................................53
      2. One Methodology, Two Approaches ........................................................................55
II. Case Studies: Gaps and Inconsistencies in the CCL .....................................................56
   A. Late Acceptance Rules ...............................................................................................57
      1. Late Acceptance: Counteroffer or Effective Acceptance? .......................................58
      2. Late Acceptance by Late Performance ......................................................................59
      3. Survey of National and International Rules of Late Acceptance ...............................60
         a. German Law ........................................................................................................60
         b. American Common Law ......................................................................................61
         c. Uniform Commercial Code ...............................................................................63
      4. Comparing Counteroffer and Effective Acceptance Approaches ...............................65
      5. Chinese Contract Law (CCL) ..................................................................................71
         a. CCL on Late Acceptance ......................................................................................71
         b. Acceptance by Performance ...............................................................................75
         c. CCL Late Acceptance in Practice ........................................................................77
         d. Reforming the CCL Late Acceptance Regime ......................................................79
   B. Anticipatory Breach .....................................................................................................83
      1. Seriousness of Breach ...............................................................................................84
      2. Express Repudiation and Reasonable Grounds for Implied Repudiation .....................84
      3. Anticipatory Breach under the CCL ........................................................................85

Published by Berkeley Law Scholarship Repository, 2016
INTRODUCTION

The transformation of the People’s Republic of China (China) into a market economy and its ascendancy into a global economic power increases the importance of studying its private laws (contract, torts, property, and unjust enrichment). The twin pillars of a market economy are private property and contract law. This Article will focus on the latter of the two pillars. The evolution of Chinese contract law provides an opportunity to study the influences of foreign laws and the formal transplantation of foreign and international law into a different cultural and legal tradition. China’s formation of private contract law, beginning in the mid-1980s, is particularly interesting because of the breadth of foreign law influences involved in its development. However, the use and partial transplantation of a variety of sources can have unintended consequences. In the case of the Chinese Contract Law (CCL), it has led to a number of gaps and inconsistencies.

Part II of this Article provides the context for the more in-depth analysis of Part III. First, it provides a brief history of the evolution of modern Chinese contract law, including the variety of foreign laws used in its development. Second, it reviews the notion of “double transplantation,” which in China’s case involved the adoption of the United Nations Convention on Contracts for the International Sale of Goods (CISG) and the reuse of the CISG in drafting the CCL. Third, it briefly illustrates the benefits of comparative law methodology as a tool for understanding and reforming the CCL.

Part III provides case studies focusing on three inconsistent and gap-ridden areas in the CCL: late acceptance rules, anticipatory breach, and the right to cure. These case studies analyze the CCL and the multiple interpretations applied to these three areas. Part III then uses comparative law sources to recommend how the CCL can be reformed to become a more consistent, rational, and comprehensive contract law.

Finally, Part IV provides some concluding remarks.
I. BORROWING: RECEPTION AND TRANSPLANTATION IN CHINA

The People’s Republic of China has gone through a series of receptions and legal transplantations1 from foreign and international private laws since it began in earnest to transition from a planned economy to a market economy during the 1980s.2 In order to facilitate trade, it adopted the Foreign Economic Contract Law of 1985 (FECL) to assure foreign parties a more modern Chinese contract law would apply to their transactions. The FECL was a comprehensible contract law stylized after modern Western civil codes. In 1988, China also became an original signatory to the CISG. This demonstrated again China’s willingness to follow Western-style contract law, as well as its foresight in seeing the benefits of a uniform international sales law to the emerging economic power that it was fast becoming. The westernization of Chinese contract law was also found in the Economic Contract Law (1981), General Principles of Civil Law (1986), and the Technology Contract Law (1987). In 1999, China elected to harmonize its domestic and foreign contract laws. The FECL was repealed and a uniform national contract law was enacted—the CCL.3

DOI: http://dx.doi.org/10.15779/Z38756T

*Associate Professor, Quanzhou Normal University, China. This research was supported by Fujian Social Sciences Funding Program (No. FJ2015C036).

**Huber Hurst Professor of Contract Law, University of Florida, Warrington College of Business.

1. The term “legal transplants” was coined in Alan Watson’s seminal work Legal Transplants. He defined legal transplants as “the moving of a rule or system of law from one country to another.” ALAN WATSON, LEGAL TRANSPLANTS: AN APPROACH TO COMPARATIVE LAW 22 (1974). Another term used for the transference of entire legal systems is “reception.” For example, scholars speak of the reception of Roman law by the emerging countries of Europe, as well as the reception of French (or German) Civil Law by other countries, including certain countries in Latin and South America. See K. ZWEIGERT & H. KÖTZ, AN INTRODUCTION TO COMPARATIVE LAW 98-119 (Tony Weir trans., 3d ed. 1998). The words “transplant” and “reception” can be used interchangeably. This Article uses “transplant” to refer specifically to the transfer or expression of rules and “reception” as transfer in the broader sense of the adoption of foreign law for an entire area of law (contract, criminal, civil procedure, and so forth). “Reception” can describe foreign “influences” on the entire legal system of the receiving country. For example, German law is highly respected in China and has had a strong influence on the development of Chinese private law. Even though this Article will primarily use the narrower Watsonian term “transplant,” it is also referring to China’s broader reception of Western legal concepts and forms of legal reasoning. This Article’s focus on the transfer of written rules does not mean to discount the significance of studying the broader reception of legal ideas, which, although more abstract, can be an even more powerful force in changing a legal culture or tradition. See Jörg Fedtke, Legal Transplants, in ELGAR ENCYCLOPEDIA OF COMPARATIVE LAW 434 (Jan M. Smits ed., 2006). Jörg Fedtke notes that the two concepts are in fact closely related: “In many cases, borrowing will not result in the copying of a specific text but rather in the transplantation of an idea.” Id. at 436 (emphasis added). The word borrowing is a better, more encompassing term that is broad enough to capture both formal transplantation and various other forms of influence. This terminology is especially useful in China’s case, given it is a civil law country by nature that has also been influenced by common law and international private law instruments.

2. See, e.g., CHINA’S GREAT ECONOMIC TRANSFORMATION (Loren Brandt & Thomas G. Rawski eds., 2008) (documenting rise as a globally influential market economy).

3. The analysis in this Article is restricted to the contract law of mainland China as
Thus, in a short period of time China received or enacted a series of Western-style contract laws—the FECL, CISG, and CCL. This Article will look at a few of the inevitable complications of such a transformation of national law. Inevitable problems in adopting foreign law may arise from: (1) translating foreign legal concepts, principles, and rules from one language to another; (2) introducing a new foreign legal regime into existing domestic legal and cultural traditions; (3) interpreting the words of a new law, which may already have particular meaning in the foreign traditions from which they came; and (4) introducing a foreign text into a country without an existing body of jurisprudence or expertise to properly and consistently apply the new law.

China’s adoption of the CISG was one of the more successful receptions. One reason for this success is the wealth of international case law and commentaries Chinese legal bodies have been able to rely on in applying CISG’s provisions. Currently, there is no hard evidence whether or not the Chinese courts have been adept at applying the CISG in a consistent way, rendering well-reasoned and autonomous interpretations in accordance with CISG’s mandate. However, there is strong evidence that Chinese arbitral bodies have successfully done so. There are currently 432 published decisions, in English, of Chinese courts and arbitral bodies applying the CISG, including 336 from China’s premier arbitral body, the China International Economic and Trade Arbitration Commission (CIETAC). The CIETAC awards have generally been well-reasoned and of high quality, showing the Commission’s ability to understand and properly apply the CISG in an unbiased manner.

The larger issue and focus of this Article is whether Chinese courts have successfully interpreted and applied the CCL. There is no simple answer to this question. There are a myriad of reasons why it is difficult to assess Chinese courts’ ability to consistently apply the CCL. First, China is generically classified as a civil law country. As such, case law is not as important as it represented by the CCL. It will not discuss the law of the semi-autonomous regions of Hong Kong (English common law), Macau (Macau Civil Code), or Taiwan (Civil Code of Republic of China and Portuguese civil law).


7. The civil law nature of Chinese law may be traced back to the Qin criminal laws (221-206 BC) and the subsequent laws of the Song, Yuan, Ming, and Qing dynasties. “Compiled in 1740, the 436 statutes and 1900 sub-statutes of the Great Qing Code was (“was” is grammatically incorrect here, maybe check source) the last dynastic legal code of Imperial China and, like its predecessors, was chiefly a criminal code.” DANIEL C. K. CHOW, THE LEGAL SYSTEM OF THE PEOPLE’S REPUBLIC OF CHINA 41-42 (2d ed. 2009). The civil law nature of Chinese law has been recently reaffirmed by its decision to continue the process of enacting a Chinese Civil Code (Draft CCC). See Wang Liming, Historic Characteristics of Modern Civil Code and its Codification Process, 8 TSINGHUA L. REV. 6-16 (2014).
would be in a common law country. Second, the depth of the jurisprudence surrounding the CCL is relatively limited given that the law has been on the books for a short period of time. Third, the text of the CCL has a number of gaps and inconsistencies that have made it difficult for the courts to understand and uniformly apply its rules. This Article will highlight some of these gaps and inconsistencies and suggest a number of solutions that would make the CCL a more holistic and rational law.

The transplantation of law has been a common occurrence in world history. Alan Watson in his seminal book *Legal Transplants* states “legal transplants—the moving of a rule or a system of law from one country to another—have been common since the earliest recorded history.” 8 Roscoe Pound noted: “History of a system of law is largely a history of borrowings of legal materials from other legal systems and of assimilation of materials from outside of the law.” 9 Modern Western legal systems, and many non-Western ones, have evolved through the transplantation and assimilation of either Roman civil law or English common law. 10

The theory of transference or transplantation of law is not without its critics. Pierre Legrand regarded transplantation of laws as an illusion, arguing the “impossibility of legal transplant . . . what can be displaced from one jurisdiction to another is, literally, a meaningless form of words.” 11 Despite the debate over the normative power or degree of success that legal transplants may have, a historical accounting, as noted by Alan Watson, shows that receptions of transplants and foreign law influences have been common in the evolution of legal systems. 12 Successful transplantations should be measured by a relative standard, whether transplantation leads to improvements in the law of the transplanting country, and not by an absolute standard, such as whether meanings attached to the words and concepts of the transplanted law have acquired the same meaning in the country of transplantation as in the country of origin. The likelihood of relative success is largely dependent on the transplanted law, those who apply it, and the level of sensitivity afforded to the legal and cultural context of the transplanting country. 13 In the end, that

9.  *Id.* at 22.
10.  *Id.*
13.  Professor Chen Lei states that in the case of legal transplantation in China and Hong Kong:

    [O]ne can conclude that as long as legal ideas are sensitive to the cultural and political context, they can move freely across the continent and influence legislation and developing legal reform—realizing that the concept of law we use as our perception of law does not prevent us from establishing a universal legal theory.
sensitivity will generate meanings and applications that vary from the law of the country from which the law was transplanted. In this context, success should be judged relative to the law prior to transplantation: is the transplanted law as interpreted and applied more consistent, more rational, and more attuned to modern commercial dealings than the prior law?

Law, whether found in a code or in case law, especially in the area of commercial law, has always had a binary relationship with the context in which it is interpreted and applied. This relationship, writ large, places the role of law in society as both a receptive and a proactive element. Commercial law, for example, generally reflects the usages, customs, and norms of commercial practice. At the same time, the law can influence the development of good practices and deter the development of exploitative behavior through what Karl Llewellyn referred to as the role of “marking out the limits of the permissible.”

In the case of China, this binary relationship heavily favors the importance of historical but evolving customs over the strict application of formal law. In contract law, the reception of foreign and international law influences can be seen as the first step in the development of a new Chinese legal culture that combines the uniqueness of Chinese customary practice with the new formalized rules of the CCL. The transplanted law acts as a catalyst bringing about an interpretive debate as to what the transplanted law should mean, and how the legal culture should change to make it work.

Current Chinese commercial laws reflect the influence of European civil laws, especially German law, and common law to a certain extent. The problem of legal transplants, as noted above, is that the text of law is easily movable from one country to the next, but legal tradition, reasoning, and theory are not so easily transplanted. Thus, legal text is taken out of the legal tradition and culture that gives it meaning and placed within (in the case of China) the

---


15. See Lei, Contextualizing Legal Transplant, supra note 13, at 194 (citing Gunther Teubner, Legal Irritants: Good Faith in British Law or How Unifying Law Ends up in New Divergences, MOD. L. REV. 11 (1998)).

16. Chen Lei notes, “Europe’s civil law tradition shares many similar values with China’s legal tradition.” He explains the civil law and traditional Chinese law prefer the “generalization of principles” and more of a communitarian perspective, as opposed to the more individualistic spirit of the common law. Lei, Contextualizing Legal Transplant, supra note 13, at 197. See also Liang Huixing, The Reception of Foreign Civil Law in China, 1 Shandong U. L. REV. 5 (2003) [hereinafter Foreign Civil Law]; Percy R. Luney Jr., Traditional and Foreign Influences: Systems of Law in China and Japan, 52 Law & Contemp. Probs. 129 (1989); Xiangmin Xu et al., The Similarities Between Civil Law Legal Family and Chinese Legal Family, 5 J. Ocean U. of China 48 (2005). But see Mary Ip, The Revised Contract Law and Its Implications on Consumerism in China, 9 Int’l J. Bus. 42, 45 (2004) (stating that the CCL “adopted and modified certain basic elements from the common law system, such as offer and acceptance.”).
context of a different legal tradition with its own distinct legal thought and view of the role of law in society. However, through a gradual process, the original legal culture, reasoning, and theory that first animated the transplanted law can be used later to understand and nurture that law in the country of transplantation.

Professor Han Shiyuan has made such a case in the area of pre-contractual liability. German law has had the greatest influence on modern Chinese law. An example of this influence is China’s adoption of *culpa in contrahendo* or bad faith negotiation, which is found in civil law but not in common law. Professor Han notes the concept of pre-contractual liability (*culpa in contrahendo*) was first introduced into Chinese law with the adoption of the FECL in 1985, and a notion of bad faith negotiation that is similar to *culpa in contrahendo* was subsequently incorporated into Articles 41 and 42 of the CCL.18 However, Han argues the ability of the Chinese courts to understand and apply such a concept depends on what he calls “theory reception.”19 A law of pre-contractual liability is more than a set of fixed rules; it is based on a broad theory of good faith. He notes CCL Articles 41 and 42 make numerous “references to foreign civil law theories and provisions,”20 including Articles 2.1.15 and 2.1.16 of the UNIDROIT Principles of International Commercial Contracts (PICC),21 and Articles 2:301 and 2:302 of the Principles of European Contract Law (PECL).22 However, simply referencing other legal instruments on the principle of good faith is unlikely to effectuate a transplant of legal theory or lead to a more complete understanding of the legal concept; only education in the civil law can achieve such a level of understanding.

The CCL was intended to harmonize China’s domestic and foreign contract laws. In doing so, the drafters relied on Western-style laws such as the CISG and the PICC. Since the CISG was adopted by China in 1988, it was a natural source for “modernizing” or “westernizing” China’s domestic contract law. However, the amalgamation of rules from different sources and legal traditions in creating the CCL resulted in unavoidable problems. Two problems relating to the CCL can be described as the “comprehension problem” and the “comprehensiveness problem.” The comprehension problem relates to the inherent difficulty of transplanting foreign laws from one legal system to another. The severing of rules and principles from the social, economic, and political context of their development undercuts the clarity of their meaning. Hugh Collins noted an “objection to transplants of legal rules insists that legal concepts fit into clusters of concepts, which together comprise a coherent and

18. Id. at 158-59.
19. Id. at 158.
20. Id.
22. COMMISSION ON EUROPEAN CONTRACT LAW, PRINCIPLES OF EUROPEAN CONTRACT LAW, pts. 1 & 2 (Ole Lando & Hugh Beale eds., combined and revised ed. 2000) [hereinafter PECL].
consistent set of rules and principles for the regulation of some aspect of social life.”

Therefore, transplanting a subset of a cluster of rules, or in China’s case, taking rules from numerous sources, has had a fundamental impact on the CCL’s comprehensibility. This use of a patchwork of different sources has led to omissions or gaps in the CCL, making it less comprehensive than it could have been.

Thus, the CCL has suffered a crisis of meaning because it was the product of partial transplants of rules uprooted from their overall conceptual schemes. China, like the countries of the former Soviet Union, adopted Western-style contract and commercial codes but has struggled to develop court systems that could place the new codes into their societal contexts. This is largely due to the courts’ failure to understand the conceptual scheme behind those rules and to adapt that scheme appropriately to a new context. This type of legal know-how takes generations of legal education and practice to develop. It is beyond the scope of this Article to determine how far along the Chinese courts have moved in interpreting and applying the CCL in a consistent way.

The second problem of legal transplants—the comprehensiveness problem—is related to the subject of the current undertaking. No code, especially not one created by using numerous foreign sources, provides a complete set of rules that covers every possible real-life scenario. There are interstitial gaps and inconsistencies within the web of rules that make up contract law. These gaps are eventually worked out by the courts. However, the separation of the formal legal text from its surrounding jurisprudence makes it more difficult for the courts in the transplanting country to resolve the resulting interpretive problems. The next section will discuss in more detail these problems of transplantation in relationship to the creation of the CCL.

A. China’s Double Transplantation and Resulting Problems

As noted above, China was one of the original eleven countries, along with the United States, to adopt the CISG. As an international convention, the CISG is not strictly an example of transplantation. Countries often adopt conventions in order to harmonize law internationally. Classic examples include the carriage of goods by sea conventions (Hague Rules and Hague-Visby rules) and the Convention on the Recognition and Enforcement of Foreign Arbitral Awards (New York Convention) (over 160 signatory countries).24 The CISG is the most successful attempt at harmonizing international private substantive law with eighty-four signatory countries and climbing.25 While it is not a traditional


example of transplantation, the CISG can nonetheless be seen as a pseudo-
transplant because it was primarily crafted from European civil law and Anglo-
American common law. As such, existing jurisprudence in the longstanding
European free market legal systems equipped European countries with a high
degree of judicial expertise for implementing the CISG. Also, substantial
numbers of commentaries were written on the new law in a short period of time.
These resources were not readily available in China. However, as noted above,
Chinese arbitral tribunals have shown a surprising adeptness in applying the
CISG.

The idea of double transplantation refers to the transplantation of foreign
law into another legal system and then the subsequent transplantation of that law
by the transplanting country to another area of its law. In the present case, the
double transplantation involves the adoption of the CISG as China’s
international sales law and China’s subsequent use of the CISG as a major
source in drafting the CCL. The drafters of the CCL were heavily influenced
by academic research, including studies of contract laws in the “United States,
Canada, Germany, the United Kingdom, Europe and Australia.”26 Thus, the re-
transplantation of the CISG and the use of a variety of foreign law sources in
drafting the CCL help explain the existence of inconsistencies and gaps within
the CCL, as well as the difficulty of its interpretation and application by Chinese
courts. The gaps and inconsistencies in the CCL will need to be resolved
through an interpretive process. This process has begun with scholarly
commentaries offering different interpretive “solutions” to the problems posed
by the CCL. The Supreme People’s Court, the highest court in China, has also
issued interpretive guidelines to help guide the lower courts in developing
uniform interpretations.27

This Article will enter these academic discussions by highlighting three
problematic areas of the CCL—late acceptance, anticipatory breach, and the
right to cure. It will also make recommendations on how best to solve these
shortcomings. The broader point of the Article is to study the issues, problems,
and solutions transplantation and reception of foreign law pose when combined
and introduced into a foreign legal system. The next section will briefly review
comparative law methodology as a way to understand and apply the CCL.

26. CHOW, supra note 7, at 345.
27. See, e.g., Interpretation I of the Supreme People’s Court of Several Issues Concerning the
Application of the Contract Law of the People’s Republic of China (promulgated by Sup. People’s
Ct., Dec. 19, 1999, effective Dec. 29, 1999), CLI.3.23702(EN) (Lawinfochina) (China);
Interpretation of the Supreme People’s Court on Issues Concerning the Application of Law for the
Trial of Cases of Disputes over Sales Contracts (Sup. People’s Ct., May 10, 2012, effective July 1,
2012), CLI.3.176318 (EN) (Lawinfochina) (China).
B. Comparative Law as an Avenue to Law Reform

Comparative contract law research has been conducted for a number of reasons. First, it has been used as a teaching device to educate students on different legal systems, typically by focusing on the differences between the civil and common law systems. A more dense literature can be found in the comparative analysis of different civil law systems, especially between the Germanic and Franco legal traditions.28 Less literature is found comparing differences among common law systems.29 Second, comparative contract law has been used as a source in the drafting of international law instruments. The most important example of this is the drafting of the CISG,30 which drew heavily from common and civil law systems. In addition to providing a degree of supranational harmonization, the CISG has also been used as a comparative law instrument in the reformation or modernization of national laws.31 Third, comparative law can be used in legal reform at the national level. This includes various degrees of use, ranging from mere influence to “legal transplant.”32 It is this third use of comparative law that will be the focus of this Article.

1. Perils and Virtues of the Comparative Law Methodology

Professor Watson lists a number of “perils” and “virtues” of comparative law methodology. Under perils he lists superficiality, incompetency, unsystematic study, and temporality.33

Watson’s list of perils is supported by intuition. First, unless a researcher is fully acculturated in both of the legal systems being compared, a degree of superficiality is inherent in such research. Second, there is a risk that a researcher from one legal system comparing its law to a foreign legal system’s law may misinterpret the foreign law being compared, a kind of incompetency problem. Third, a selectivity problem exists because a systematic comparison of entire legal systems is beyond most researchers’ abilities or scope. Thus, the researcher will analyze specific legal rules, the selection of which will be at least


29. But see, e.g., COMMERCIAL CONTRACT LAW: TRANSATLANTIC PERSPECTIVES (Larry DiMatteo et al. eds., 2013) (comparing the common law systems of the United Kingdom and United States); COMPARATIVE CONTRACT LAW: BRITISH AND AMERICAN PERSPECTIVES (Larry DiMatteo & Martin Hogg eds., 2015) (same).


31. The CISG has heavily influenced the modernization of contract and sales law in China, Germany, The Netherlands, and is likely to have similar influences in the revisions of the French, Japanese, and Spanish Civil Codes.

32. MATHIAS SIEMS, COMPARATIVE LAW 191-220 (2014). Professor Siems says that there are positive and negative views of the integrity of comparative law. The positive view is exemplified in the work of Alan Watson. Siems refers to Watson as the “father of legal transplants.” Id. at 195.

33. WATSON, supra note 1, at chs. 2-3.
This can lead to a degree of arbitrariness in the conclusions reached and a danger of generalizing from those conclusions to characterize the greater body of law. Fourth, disparity in economic and legal development between the two countries being compared may lead to a bias toward the more highly developed law or country. This may prove troublesome because a given law may be efficient at one stage of development but become inefficient at a higher stage of development. For example, a strict product liability law may not make much sense in a poor and underdeveloped country but may make sense in a highly industrialized and developed one. The two legal systems are simply at different points on the evolutionary path.

Additional perils also exist. Homeward trend is an issue, which occurs when a researcher from a different legal system examines the rules or lack of rules of another legal system and is subjectively prejudiced by the legal concepts of law found in the researcher’s own legal system. This does not have to be an issue of temporality since the two countries being compared might be at the same level of development. The comparatist must also fully recognize the multiple interlocking systems that make up a society (economic, cultural, legal, religious, and so forth). Some societies may allocate certain issues to the legal realm, while others may deal with such issues through non-legal systems. The separation of legal rules from these interlocking systems commonly results in misunderstandings regarding the meanings of those rules and how they should be applied.

The authors believe the virtues of comparative law are numerous. Comparative law may be used as a method to better understand the evolution of law. It may also be used as a powerful tool for reforming law, as it provides a survey of options used in other systems. Finally, comparative law may be used to analyze the transplantation of laws from one system to another, and the subsequent application of the laws in the receiving system.

Watson notes the formal rules being transplanted are subject to interpretation by the courts of the receiving country. This sudden disconnect between text and context means the rules “may equally operate to different effect in the two societies, even though [they are] expressed in apparently similar terms.” This is especially the case when introducing a Western, highly formalized law into a country with non-Western economic, social, and cultural

34. The notion of “homeward trend bias” has been used in relation to the interpretation of the CISG by different national court systems. See Ingeborg Schwenzer, Divergent Interpretations: Reasons and Solutions, in INTERNATIONAL SALES LAW: A GLOBAL CHALLENGE 102, 103 (Larry A. DiMatteo ed., 2014) (identifying homeward trend as “interpreting the provisions of the CISG according to existing or merely presumed domestic counterparts”).

35. See Daniel Berkowitz et al., The Transplant Effect, 51 AM. J. COMP. L. 163 (2003) (noting that the success of transplants is dependent on conforming to existing or merely presumed domestic counterparts).

36. WATSON, supra note 1, at 20.
norms. In China’s case, a tradition exists that is heavily based on Confucianism. In such a system, formalized private law has historically played a far lesser role than cultural norms found in business relationships and resorting to the courts as a means of dispute resolution is disfavored. One example of this tradition is the Chinese concept of guanxi in which business transactions are regulated by informal social and status-based relational norms. Guanxi places a great deal of importance on respect, reputation, and relational networks and not on the enforcement of formalized institutional support systems, such as contractual rights. Thus, a comparative law analysis should also seek to study the effects of a transplanted law on existing social and cultural systems.

2. One Methodology, Two Approaches

There are two traditional approaches in comparative law studies—the common core approach and the “better rules” approach. The first approach, championed by Continental European scholars such as Rodolfo Sacco at the University of Turin and Rudolf Schlesinger at Cornell University in the 1950s and 1960s, looks at the commonalities among different legal systems. The second approach analyzes the differences between legal systems and assesses which of their different rules are “better.” Oxford Professor Hugh Collins describes this comparative law methodology as a “utilitarian approach to comparative law . . . [which] seeks through a comparison of the legal rules and techniques of different jurisdictions the best solutions to legal problems. The aim is to identify better solutions in foreign legal systems and then to recommend their incorporation into domestic law.” This Article will use both approaches. A comparative analysis will be performed on the rules for late acceptance by comparing the rules of major legal systems and international law instruments. A less in-depth use of comparative law will be used in the sections on anticipatory breach and the right to cure.


40. Rodolfo Sacco is one of Europe’s most famous comparative law scholars. See id.


42. Collins, supra note 23, at 397.
II. CASE STUDIES: GAPS AND INCONSISTENCIES IN THE CCL

Formation of a contract is generally determined under an offer-acceptance paradigm. The CCL adopts this widely held model of contract formation. Before proceeding with the analysis of the CCL’s late acceptance rules or lack thereof, it is important to state the obvious: contract law is a rules-based system. One area of contract law where this rules-based system is extensive is in the offer-acceptance rules of contract formation. The rule density in this area includes primary rules, exceptions to those rules, and exceptions to the exceptions. A comprehensive set of such rules answers the core questions of whether a contract has been formed, when it has been formed, and the content of the concluded contract. Weaknesses in these rules prevent the law from efficiently answering these questions. Such a weakness can be found in the CCL’s late acceptance rules.

Contracts are formed in a variety of ways, including a bilateral contract (exchange of promises), a unilateral contract (offer promise followed by acceptance by conduct or performance), or an implied-in-fact contract (conduct followed by conduct). The traditional contract model involves an exchange of promises either orally or in written form. The common contract formation paradigm involves the exchange of offer and acceptance, and common and civil laws have developed precise offer-acceptance rules relating to the formation of contracts. Under these rules, acceptance is the key communication that creates binding obligations. Generally, the civil and common laws have similar rules for the conclusion of a contract. However, the two systems have notably different rules relating to the time when an acceptance becomes effective. According to civil law, a contract is formed when an acceptance is received by the offeror. By contrast, a contract is concluded under common law upon the sending or dispatch of the acceptance (as long as the transmission of the acceptance is by reasonable means). Thus, a contract under common law is formed at an earlier point in time, which limits the time during which the offeror may revoke the offer. Despite the difference in when an acceptance is

43. Parviz Owsia, The Notion and Function of Offer and Acceptance under French and English Law, 66 Tul. L. Rev. 871, 872 (1992) (“The most common mechanism of contract formation, offer and acceptance, is used as a standard tool under both [common and civil law] systems.”).
44. CHOW, supra note 7, at 350 (stating that provisions of the CCL “borrow heavily from foreign law and [are] based upon an offer and acceptance model”).
45. See discussion infra Part III.A.
46. E. ALLAN FARNSWORTH, FARNSWORTH ON CONTRACTS 205-06, 253-54 (3d ed. 2004) [hereinafter FARNSWORTH ON CONTRACTS]
47. But see PETER HUBER & ALASTAIR MULLIS, THE CISG: A NEW TEXTBOOK FOR STUDENTS AND PRACTITIONERS 100-02 (2007) (“Conclusion of contract otherwise than by offer and acceptance”).
48. PECL, supra note 22, art. 2:205 n.2.
49. FARNSWORTH ON CONTRACTS, supra note 46, at 337-38.
considered effective, the offer-acceptance rules of common and civil law systems, including rules dealing with late acceptance, show a high degree of consistency.

The next section examines the rules and rationales relating to acceptance in the context of the CCL. It will discuss and analyze the late acceptance rules in the civil and common law systems, along with the late acceptance rules found in the CISG. It then questions the lack of similar rules in the CCL. It concludes that the CCL should be reformed to more fully address scenarios involving late acceptance, including late dispatch of acceptance and belated delivery of acceptance after a timely dispatch. Such reform would not only fill a gap in the CCL, but would also make the CCL consistent with the CISG.

A. Late Acceptance Rules

The lack of an adequate set of late acceptance rules in the CCL is puzzling given that the CISG incorporates generally recognized late acceptance rules. Again, the CISG is the law of China in international sales and was used as a primary source in writing the CCL. The CCL represents an exception to most contract law regimes in terms of not possessing a complete set of rules dealing with the issue of late acceptance. Late acceptance rules are important because they directly impact if and when a contract is formed. Late acceptance and how the law responds to it raise a number of important questions. Do the reasons for late acceptance—belated dispatch or delayed transmission—require different rules? Is late acceptance itself something that can bind a contract? Or, is late acceptance a rejection of the offer, becoming instead a counteroffer? Legal systems answer these questions differently, leading to different real-world outcomes.

This part of the Article provides a comparative analysis of late acceptance rules in German law, American law (UCC and common law of contracts), and the CISG. The Article divides the existing rules into two types—"counteroffer theory" rules and "effective acceptance theory" rules. It then reviews Chinese law and recommends the adoption of new default rules to guide the reformation of the CCL.

Both the civil and common law systems base contractual obligations on the parties’ agreement to enter into a legally binding contract. As a general matter, a contract becomes binding when an acceptance reaches the offeror (except under the common law). Most international sales and contract instruments have adopted the civil law’s receipt rule.50

Common law’s dispatch rule limits the problem of late acceptance because the contract is binding, even if acceptance is lost or delayed in transmission, as long as it was properly sent within a reasonable period of time. However, if the dispatch is not proper—not transmitted in compliance with requirements set out in the offer, not sent by a reasonable means of transmission, not properly

50. See, e.g., CISG, supra note 4; PICC, supra note 21.
addressed or posted, or not sent within a reasonable time—the contract is not binding until received by the offeree. If acceptance is not received due to one of the above reasons within the time stated in the offer or within a reasonable time, then it is considered a late acceptance.51 Another way of understanding late acceptance is to understand that offers self-terminate after a period of time. If acceptance is not sent under common law or received under civil law within a reasonable period of time, there cannot be a contract because the offer has lapsed. However, most legal systems provide special rules in cases of late acceptance. The following sections will address these rules.

1. Late Acceptance: Counteroffer or Effective Acceptance?

The first rule of effective acceptance is that it must be unconditional and unequivocal. In short, the terms and conditions of the acceptance must mimic those of the offer. Under common law, acceptance must be a “mirror image” of the offer.52 The “mirror image” rule is associated with the common law doctrine that considers differences in the terms and conditions of the acceptance, relative to the offer, as a rejection of the offer. Only when acceptance meets all the conditions of the offer can it constitute an effective acceptance. Therefore, if the offeror has fixed a specific time or period for acceptance, the offeree must accept within that period of time.53 If a time or period has not been fixed in the offer, contract law implies the offeree must accept within a reasonable period of time. If the offeree does not respond to the offer within a reasonable period of time, the response will be considered a late acceptance.54

As a general rule, late acceptance, whether due to late dispatch or due to delay in transmission, is treated as a counteroffer that gives or returns the power to make a contract to the original offeror. The American Law Institute’s Restatement (Second) of Contracts (Restatement),55 the American Uniform Commercial Code (UCC), and the CISG provide similar rules in cases of late acceptance. If it is obvious to the offeror the acceptance was timely and properly dispatched but was delayed in transmission, the offeror must notify the offeree of the lateness of its receipt in order to prevent the formation of a contract.56 If late acceptance is due to a belated dispatch and not a problem in transmission, then the offeror can treat the offer as lapsed, and the late dispatch constitutes a

52. The “mirror image” rule is associated with the common law doctrine that considers differences in the terms and conditions of the acceptance, relative to the offer, as a rejection of the offer.
53. See, e.g., Bürgerliches Gesetzbuch [BGB][Civil Code], Jan. 2, 2002, Bundesgesetzblatt Teil 1 [BGB. I] 42, last amended Oct. 1, 2013, § 148, translation at http://www.gesetze-im-internet.de/englisch_bgb/ (Ger.) (“If the offeror has determined a period of time for the acceptance of the offer, the acceptance may only take place within this period.”).
rejection of the offer.\footnote{CALAMARI & PERILLO, \textit{supra} note 56, at 89 (“If an offer lapses before an acceptance becomes effective, it would seem to follow that the late acceptance is an offer.”).} However, the offeror can accept the late acceptance by sending a notice to the offeree of the offeror’s intention to recognize the acceptance as binding the contract. The question then becomes whether the contract is formed when the late acceptance was sent (common law’s dispatch rule), when the late acceptance was received (civil law and CISG’s receipt rule), when the offeror dispatches a notification of effective late acceptance (in the case of a belated dispatch), or when the offeree receives the notification sent by the offeror. This issue will be discussed later in the Article. These general rules relating to late acceptance provide the context in which this Article reviews the CCL.

2. \textit{Late Acceptance by Late Performance}

In a unilateral contract, the offeror invites the offeree to accept by conduct or performance. This “invitation” may be express or implied. Implied acceptance by performance may be based on prior dealings, trade usage, or business customs.\footnote{RESTATEMENT (SECOND) OF CONTRACTS §§ 32, 53 (1981); U.C.C. §2-206 (AM. LAW INST. & UNIF. LAW COMM’N 2014).} Three questions must be answered: (1) What type of conduct or performance is needed to bind the contract—beginning performance (such as beginning the manufacture of the goods) or completing performance (such as sending existing goods), (2) what happens if the performance is delayed, or delayed after it has begun, and (3) will either scenario be considered equivalent to late acceptance in a bilateral contract? The UCC only requires the offeree to begin performance for the conduct to be considered a binding acceptance.\footnote{See also LA, CIV. CODE ANN. art. 1939 (2015) (“When an offeror invites an offeree to accept by performance and, according to usage or the nature or the terms of the contract, it is contemplated that the performance will be completed if commenced, a contract is formed when the offeree begins the requested performance.”). See also U.C.C. § 2-206(2) (AM. LAW INST. & UNIF. LAW COMM’N 2014) (“The beginning of a requested performance is a reasonable mode of acceptance.”).} The CISG indicates that complete or near complete performance is required.\footnote{CISG, \textit{supra} note 4, art. 25 (fundamental breach).}

\textit{Restatement} sections 45 and 50(2) state acceptance by performance “requires that at least part of what the offer requests be performed.”\footnote{RESTATEMENT (SECOND) OF CONTRACTS (1981), §§ 45, 50(2).} This is the case when the offer is one for a unilateral contract in which acceptance can only be effectuated by performance and not by promise.\footnote{\textit{Id.} § 45 cmt. a.} \textit{Restatement} section 62 is more explicit by noting that the tendering or beginning of performance is an acceptance by performance when the offer provides the offeree the choice of accepting by promise or performance.\footnote{\textit{Id.} § 63.} The rationale given in both cases is that
the beginning of performance “operates as a promise to render complete performance.”\(^{64}\) Acceptance by performance can be especially important under the UCC, which requires a writing (most notably under the statute of frauds) to create an enforceable contract in almost all cases.\(^{65}\) A notable exception to the writing requirement exists for the purchase of “specially manufactured goods”\(^{66}\) where there is “either a substantial beginning of their manufacture or commitments for their procurement.”\(^{67}\) The question remains whether late acceptance rules apply to unilateral contracts as they do in bilateral contracts if an offeree unreasonably delays the beginning of performance. The most rational answer is that promise and conduct are both methods of acceptance and, therefore, late acceptance rules are applicable. If aware of the belated performance, the offeror should be able to treat it as a counteroffer that he is free to reject. A caveat would be the case where, despite the delay in performance, the offeree still has the ability to perform an on-time delivery. In practice, the parties would likely communicate regarding the progress of performance. If progress is unduly delayed and the parties assume a contract was formed, an alternative would be for the non-breaching party to declare an anticipatory repudiation and sue for damages.

3. **Survey of National and International Rules of Late Acceptance**

The German Civil Code (BGB) and American common law will be used as representatives of their respective legal systems. This review will also consider the rules found in the UCC and the CISG.

a. **German Law**

In German law, the reason for the belated acceptance—be it belated dispatch or a delay in transmission—is important to the application of late acceptance rules. The BGB’s late acceptance rules are found in Articles 149 and 150.\(^{68}\) BGB Article 150, entitled “Law and Altered Acceptance,” provides the general rule that “late acceptance of an offer is considered to be a new offer.”\(^{69}\) Traditionally, civil law provides that late acceptance due to belated dispatch is a counteroffer, which the original offeror is free to accept, reject, or ignore.\(^{70}\)

---

64. A caveat to the acceptance by performance rule is when the offeree sends non-conforming goods not as an acceptance but as an accommodation. In that case, the offeror is free to accept or reject the goods. See U.C.C. § 2-206(1)(b) (AM, LAW INST. & UNIF. LAW COMM’N 2014) (“[A] shipment of nonconforming goods does not constitute an acceptance if the seller seasonably notifies the buyer that the shipment is offered only as an accommodation to the buyer.”).

65. Id. § 2-201.

66. Specially manufactured goods of the type that are specifically made for the buyer and “are not suitable for the sale to others in the ordinary course of seller’s business.” Id. §2-201(3)(a).


68. BGB art. 149-50 (Ger.).

69. Id. art. 150 (The late acceptance of an offer is considered to be a new offer.).

70. See COMMENTARY ON THE UNIDROIT PRINCIPLES OF INTERNATIONAL COMMERCIAL
However, BGB Article 149 (“Late Receipt of a Declaration of Acceptance”) places an obligation on the offeror to respond to a late acceptance if the acceptance “was sent in such a way that it would have reached him in time if it had been forwarded in the usual way” and “if the offeror ought to have recognized” the acceptance had been properly sent. In such cases, the offeror is required to “notify the acceptor [offeree] of the delay” within a reasonable period of time. If the offeror unduly delays in sending the notification, “the acceptance is deemed not to be late” and a contract is concluded.

b. American Common Law

The American common law of contracts is the law of the individual state court systems. The basic principles and concepts of the common law are similar across the states, but their interpretation and application may vary. This creates majority and minority views and, in some cases, a series of minority views without a mainstream or majority view. In sum, the same terminology and rules are applied with different outcomes. The Restatement may be considered a more stable representation of American common law because it provides concise descriptions of common law rules while providing normative insights into what the law should be. American courts often reference Restatement provisions, and in some cases, adopt Restatement rules. This phenomenon confirms the prescriptive role the Restatement plays in American law. Therefore, the Restatement will be used as representative of American common law.

The Restatement fails to provide late acceptance rules. This is predominantly due to the fact that late acceptances—whether due to a belated dispatch or a delay in transmission due to the fault of the offeree—are treated as counteroffers. However, if acceptance is properly dispatched in a timely fashion but is delayed or lost in transmission, a contract is formed at the time of dispatch under the common law’s dispatch or “mailbox” rule. Restatement

CONTRACTS (PICC) 27273 (Stefan Vogenauer & Jan Kleinheisterkamp eds., 2009) [hereinafter COMMENTARY ON THE PICC].

71.  BGB art. 149 (Ger.)
If a declaration of acceptance received late by the offeror was sent in such a way that it would have reached him in time if it had been forwarded in the usual way, and if the offeror ought to have recognized this, he must notify the acceptor of the delay after receipt of the declaration without undue delay, unless this has already been done. If he delays the sending of the notification, the acceptance is deemed not to be late.

72.  Id.

73.  A classic example is the serial referencing of Section 90 of the Restatement in court decisions. Section 90 deals with liability predicated upon detrimental reliance, commonly referred to as promissory estoppel. Thus, liability premised solely on the breach of a promise has been supplemented by liability based upon reliance. Reliance damages may be awarded in cases of non-contractual promises (one-way promise). See RESTATEMENT (SECOND) OF CONTRACTS § 90 (1981). E. Allan Farnsworth notes that: “Restatement Second [§ 90] states that recovery ‘may be limited as justice requires,’ language that is generally invoked in limiting recovery to damages based on the reliance interest.” See FARNSWORTH ON CONTRACTS, supra note 46, at 180.

section 63 states that an acceptance “is operative and completes the manifestation of intent as soon as put out of the offeree’s possession, without regard to whether it ever reaches the offeror.” Ancillary to the dispatch rule is that the acceptance must be sent by means dictated by the offer. If the offer does not provide for the means of transmission, then the offeree must send the acceptance using a reasonable means of transmission under the circumstances, and the acceptance has to be “properly” dispatched (for example, correct mailing address and postage). However, Restatement section 57 makes an exception where an acceptance is improperly dispatched but is received “within the time in which a properly dispatched acceptance would normally have arrived” so that it is deemed to have been “operative upon dispatch.” Thus, an improper dispatch of the acceptance will delay the application of the “mailbox” rule pending receipt, but if the communication is nonetheless reasonably received, then the “mailbox” rule goes into effect. The time of contract formation is then the time when the acceptance was dispatched. In the case of a delay in transmission, the fact that an acceptance is received unreasonably late is irrelevant since the contract has already been formed. The remaining issue involving the late sending of the acceptance is whether the offeror may accept the late acceptance. Restatement section 70 answers this in the negative, stating a belatedly sent acceptance is a counteroffer.

Despite the lack of necessity for late acceptance rules, a density of offer-acceptance rules can be seen in Restatement sections 49 and 54. In the case of a unilateral contract where the offeror invites acceptance by performance, the offeree is not required to provide notification of the commencement of performance. However, Restatement section 54 places such an obligation to give notice of commencement when the offeree “has reason to know that the offeror has no adequate means of learning of the performance with reasonable promptness and certainty.” The offeree must exercise due diligence to notify the offeror of acceptance unless “the offer indicates that notification of acceptance is not required.” This issue of the offeree providing notice of performance in a unilateral contract is an unsettled issue in the CCL.

Another issue relates to delays in the transmission of an offer. Is the time provided for acceptance of the offer extended for a period of time equivalent to the delay in the transmission of the offer? The Restatement answers the question in the negative: “the period within which a contract can be created by

---

75. Id. § 63.
76. Id. § 63(a).
77. Id. § 65.
78. Id. § 66.
79. Id. § 57.
80. Id. § 70.
81. Id. § 54.
82. Id.
83. Id. § 54(2)(a-c).
acceptance is not thereby extended if the offeree knows or has reason to know of the delay.\textsuperscript{84} However, if the delay is due to the fault of the offeror and the “offeree neither knows nor has reason to know that there has been delay, a contract can be created by acceptance within the period which would have been permissible if the offer had been dispatched at the time that its arrival seems to indicate.”\textsuperscript{85} A reasonable interpretation of this phrase is that the time for acceptance is extended for a time equivalent to the delay.

c. Uniform Commercial Code

The UCC does not provide a full regime of offer-acceptance rules and does not contain late acceptance rules, except in cases where the offeror invites acceptance by performance (notice must still be given) and where silence can be a means of acceptance.\textsuperscript{86} Part of this absence is alleviated by the written confirmation rule,\textsuperscript{87} the writing exception for specially manufactured goods,\textsuperscript{88} and the recognition of unilateral contracts.\textsuperscript{89} The informality of transactions involving the sale of goods is evident in the core section on contract formation. Section 2-204 states a contract may be formed “in any manner sufficient to show agreement, including [by] conduct.”\textsuperscript{90} It further states an agreement to enter into a contract may be recognized “even though the moment of its making is undetermined.”\textsuperscript{91} Additionally, unlike the common law of contracts, a contract may be formed even if not all the material terms have been agreed to, as long as the parties intended to form a contract and “there is a reasonably certain basis for giving an appropriate remedy.”\textsuperscript{92} Since the UCC is not considered a comprehensive or a complete preemption of the common law, the common law is used to fill in the gaps in the UCC.

Regarding acceptance by performance, a unilateral contract can be formed by the conduct of the offeree in two instances. If the offer invites acceptance by conduct or performance, the type of conduct needed to bind the contract depends on the circumstances. If the goods exist, then the offeree may accept by prompt shipment of the goods. If the goods are not in existence or are not in hand, then the “beginning of performance” is a reasonable method of acceptance.\textsuperscript{93} However, the UCC qualifies this method by requiring that the offeror receive

\textsuperscript{84} Id. § 49.
\textsuperscript{85} Id.
\textsuperscript{86} U.C.C. § 2-206 (AM. LAW INST. & UNIF. LAW COMM’N 2014).
\textsuperscript{87} Id. § 2-201(2).
\textsuperscript{88} Id. § 2-201(3) (a).
\textsuperscript{89} Id. (“[U]nder circumstances which reasonably indicate that the goods are for the buyer, has made either a substantial beginning of their manufacture or commitments for their procurement”).
\textsuperscript{90} Id. § 2-204(1).
\textsuperscript{91} Id. § 2-204(2).
\textsuperscript{92} Id. § 2-204(3).
\textsuperscript{93} Id. § 2-201(a).
notice of the beginning of performance. It states “an offeror who is not notified of acceptance within a reasonable time may treat the offer as having lapsed before acceptance.”\textsuperscript{94} Thus, the contract is created by conduct but the offeree must still notify the offeror of the beginning of performance.

The requirement that the offeree notify the offeror of the beginning of performance in the case of acceptance by performance raises the same issues seen in the late acceptance by promise scenario. What if the notification is delayed in transmission and is received by the offeror belatedly? If the notice is sent belatedly, can a contract still be formed? The fact that Section 2-206(2) states that the offeror of a late notice “may” treat the offer as lapsed implies the offeror may also consider the contract as formed. Does the offeror have any responsibility to notify the offeree of the late notice in the case of a delay in transmission? The UCC does not provide answers to these questions. However, in the case where the offeror treats the offer as lapsed while knowing the offeree is continuing in its performance, it seems the offeror would be obligated to notify the offeree of the nonexistence of the contract. Again, the common law’s offer-acceptance rules would apply to UCC transactions. For example, UCC Section 2-206 does not expressly preclude the offeror from revoking the offer after the beginning of performance but before receiving notice. A comment to Section 2-206 notes the importance of the common law in such situations: “Nothing in this section bars the possibility that under the common law performance begun may have an intermediate effect of temporarily barring revocation of the offer.”\textsuperscript{95} This is supported by analogy to the dispatch or “mailbox” rule that dictates a contract is formed at the time the acceptance is dispatched. In the acceptance by performance scenario, the beginning of performance is the equivalent to a dispatch. Thus, the parties have entered into a binding contract, conditional on the offeree sending the required notice. Under German law, if there is any doubt as to whether a contract has come into existence, even if there has been a commencement of performance, the contract will be judged as having not come into existence.\textsuperscript{96}


The American common law and UCC approach view a late acceptance as a counteroffer. Some civil law countries also take this approach. For example, Article 1393 of the Civil Code of Québec states “[a]n acceptance which does not correspond substantially to the offer or which is received by the offeror after the offer has lapsed does not constitute acceptance.”\textsuperscript{97} In contrast, the Italian Civil Code holds out the possibility that a late acceptance may still be an effective

\textsuperscript{94} Id. § 2-206(2).
\textsuperscript{95} Id. § 2-206 cmt. 3.
\textsuperscript{96} BGB § 154(1) (Ger.).
\textsuperscript{97} Civil Code of Québec, S.Q. 1991, c 64, art 1393 (Can.).
acceptance. The CISG adopted the Italian law approach as stated in CISG Article 21, “A late acceptance is nevertheless effective as an acceptance if without delay the offeror orally so informs the offeree or dispatches a notice to that effect.”

If a letter or other writing containing a late acceptance shows that it has been sent in such circumstances that if its transmission had been normal it would have reached the offeror in due time, the late acceptance is effective as an acceptance unless, without delay, the offeror orally informs the offeree that he considers his offer as having lapsed or dispatches a notice to that effect.99

Accordingly, if the offeree is informed of the effectiveness of his late acceptance, the “lapsed” offer remains in force and leads to the formation of a contract through the late acceptance.100 This rule allows a late acceptance to be effective retroactively, so that the originally proposed contract is concluded at the time the late acceptance reached the offeror. The CISG’s late acceptance rule is classified in this Article as the “effective acceptance theory” approach, which is explained further below.

4. Comparing Counteroffer and Effective Acceptance Approaches

What are the differences between counteroffer and effective acceptance theories of late acceptance? First, the nature of the original offeror’s reply to the late acceptance is different. Under counteroffer theory, the late acceptance cannot be an effective acceptance. The offeror considers such late acceptance as a counteroffer and may or may not “accept” it. The Restatement provides “[a] late acceptance may be an offer which can be accepted by the original offeror.” For example, “A” offers to sell “B” a tractor for $12,000 and states he needs to have B’s answer within three days. On the fourth day, B telephones A to accept. B’s response constitutes a late acceptance and a rejection of A’s offer. However, A may choose to accept B’s response as a counteroffer and respond, in a timely manner, indicating his intent to accept B’s offer. In this case, the original offeror’s reply is an “acceptance” to the counteroffer (late acceptance). However, the original offeror cannot at his election regard or render the late acceptance an effective acceptance. The contract is formed when the original offeror dispatches a notice of acceptance of the original offeree’s counteroffer.

In effective acceptance theory, the original offeror may consider the late acceptance as an effective acceptance if the offeror, without delay, notifies the offeree to that effect. The offeror’s reply is in essence a declaratory notice

98. Art. 1326 Codice civile [C.c.] (It.).
99. CISG, supra note 4, art. 21.
100. See HUBER & MULLIS, supra note 47, at 97-98.
102. See id. § 70 cmt. a (“Nor can the original offeror ‘waive’ his right to reject, or at his election regard the counter-offer as an acceptance.”).
instead of an acceptance.\textsuperscript{103} Such a declaratory notice can be seen as activating the effectiveness of the late acceptance and reactivating the lapsed “offer” at the same time. Schlechtriem and Schwenzer assert: “[t]he offeror’s declaration of approval therefore cures a late acceptance, even if his declaration is lost or arrives late.”\textsuperscript{104} This interpretation implies the contract is formed retroactively to the time of receipt of the late acceptance upon the dispatch of the offeror’s notice.

Second, the risks of the original offeror’s reply are different. As mentioned above, according to the counteroffer theory the original offeror’s reply to the late acceptance is considered an “acceptance.” The general rule (civil law and CISG) is that a declaration of intent (offer or acceptance) becomes effective at the point when this declaration reaches the other party.\textsuperscript{105} Therefore, if the notice fails to reach the original offeree, there is no receipt and no contract. By contrast, the common law’s dispatch rule states an acceptance, with a few exceptions, is effective at the time it is sent. Thus, the offeror’s ability to revoke its offer comes to an end. Under the BGB, if a revocation of the counteroffer reaches the original offeror before or at the same time that the original offeror’s acceptance reaches the original offeree, no contract is formed.\textsuperscript{106} Under counteroffer theory, the risks of loss or delay of the acceptance to the counteroffer are borne by the original offeror. Therefore, if a revocation of the counteroffer were received by the original offeror during the period in which the acceptance is delayed, the receipt of the delayed acceptance would be of no consequence. According to the common law, there would be a contract since the acceptance of the counteroffer was effective upon dispatch. The fact that the revocation was received prior to the acceptance is of no consequence.

Article 21(1) of the CISG contains an exception to the CISG’s general receipt rule. A declaratory notice becomes effective as long as it has been “dispatched.”\textsuperscript{107} Therefore, the risk of loss or delay of the declaratory notice is borne by the original offeree.\textsuperscript{108} This rule is in line with the common law’s dispatch theory in which the effectiveness of an acceptance is triggered by its dispatch. According to CISG Article 27, such dispatch shall be made by means appropriate under the circumstances.\textsuperscript{109} Under American common law, there is

\begin{itemize}
\item \textsuperscript{104} Id. at 253. See also Fritz Enderlein & Dietrich Maskow, International Sales Law 104 (1992).
\item \textsuperscript{105} CISG, supra note 4, art. 15(1), 18(2).
\item \textsuperscript{106} BGB art. 130(1) (Ger.) (effectiveness of a declaration of intent to absent parties). See also CISG, supra note 4, art. 18 (2) (an acceptance of an offer becomes effective at the moment the indication of assent reaches the offeror).
\item \textsuperscript{107} Schlechtriem & Schwenzer, supra note 103, at 307, 314.
\item \textsuperscript{108} See id. at 253.
\item \textsuperscript{109} Id. at 307 (“However, in so far as the ideas underlying Article 27 are also relevant to communications provided for in Part II and the need for them to be dispatched or to ‘reach’ the addressee has been left open, it will be possible, on the basis of Article 7 (2), to apply the principle

Published by Berkeley Law Scholarship Repository, 2016
not much difference between counteroffer theory and effective acceptance theory in this area since the common law generally treats acceptance as effective when it is sent, not when it arrives.\textsuperscript{110}

Under the effective acceptance theory, the offeree is prevented from revoking because the contract is formed when the late acceptance is received, pending the sending of a notice by the offeror informing the offeree of the effectiveness of the late acceptance (delay in transmission scenario). In effective acceptance theory, after late acceptance has been received, the offeree does not have a right to revoke if the offeror dispatches the required declaratory notice. This is similar to the common law’s rule relating to an acceptance overtaking a rejection. If an acceptance is sent but is overtaken by a subsequently sent rejection, a contract is formed upon the dispatch of the acceptance, removing the right of rejection.\textsuperscript{111}

The common law makes a number of exceptions to the dispatch rule for effective acceptance. First, if the offeree uses an improper form of transmitting the acceptance, improperly addresses the communication, or does not take “reasonable precautions to ensure safe transmission,”\textsuperscript{112} the acceptance becomes effective on receipt.\textsuperscript{113} Second, in cases where a rejection is sent but not yet received before the offeree changes his or her mind and sends an acceptance, the dispatch rule would work an injustice on an offeror who first receives and relies upon the rejection. In such cases, the common law subjects the acceptance to a receipt rule.\textsuperscript{114} Therefore, whichever communication, rejection or acceptance, reaches the offeror first becomes effective.

To repeat, if the acceptance is sent first, followed by a rejection, a contract is formed under the dispatch rule, but if the rejection is sent first and is overtaken by an acceptance, a contract is formed if the acceptance is received before the rejection. The rationale for the second rule is that it would be unjust not to allow the offeror to rely on the rejection if it is the first-received instrument. This seems nonsensical since one would have the same injustice if the rejection were received prior to the acceptance in the event the rejection

\footnotesize{underlying Article 27.”).}

\textsuperscript{110} See \textsc{Restatement (Second) of Contracts} § 63 (1981):

\begin{quote}
Unless the offer provides otherwise,
\begin{enumerate}
\item an acceptance made in a manner and by a medium invited by an offer is operative and completes the manifestation of mutual assent as soon as put out of the offeree’s possession, without regard to whether it ever reaches the offeror; but
\item an acceptance under an option contract is not operative until received by the offeror.
\end{enumerate}
\end{quote}

\textsuperscript{111} See \textsc{id}; \textsc{Farnsworth on Contracts}, \textit{supra} note 46, at 337-38.

\textsuperscript{112} \textsc{Joseph M. Perillo & John D. Calamari, Calamari and Perillo on Contracts} 112 (West Grp. 5th ed. 2003).

\textsuperscript{113} See \textsc{Restatement (Second) of Contracts} §§ 66-67 (1981); \textsc{Farnsworth on Contracts}, \textit{supra} note 46, at 339.

\textsuperscript{114} \textsc{Restatement (Second) of Contracts} § 63 cmt. c, illus. 7 (1981); \textsc{Farnsworth on Contracts}, \textit{supra} note 46, at 340-41.
overtakes the acceptance. The rationale for not protecting the offeror in the first instance is the offeror is more likely to suffer harm if the offeree is able to opportunistically play the market. For example, the offeree sends an acceptance by regular mail, deemed to be a proper means of transmission, which takes two to three days to deliver. While the acceptance is in transmission, the offeree monitors the market and determines he can now obtain a contract at a better price. He then sends an e-mail rejection that is received prior to the acceptance in order to take advantage of the market change. This would be unfair to the offeror and, therefore, the dispatch rule remains in place.\textsuperscript{115} This is a weak argument since the same rationale can be applied to the “exception rule” (acceptance overtaking a rejection), wherein the offeree sends a rejection by regular mail, monitors the market, and then expeditiously sends an acceptance. The rational approach would be to apply to both situations (rejection overtaken by acceptance and acceptance overtaken by rejection) the rule that whichever instrument is received first should control.

Third, under counteroffer theory, unless the late acceptance (counteroffer) specifies the period allowed for the original offeror’s reply, the original offeror must accept the counteroffer by dispatch within a reasonable time.\textsuperscript{116} However, under the effective acceptance approach, the offeror shall notify “without delay” of the effectiveness of the acceptance.\textsuperscript{117} The question is whether there is a difference between acceptance of a counteroffer within a reasonable time and notification of the effectiveness of a late acceptance without delay. Lando and Beale indicate the time set for acceptance of a counteroffer within a reasonable time is generally longer, under most circumstances, than the time provided for effective notice without delay in effective acceptance theory.\textsuperscript{118} Therefore, the offeror is provided additional time under counteroffer theory to speculate on market movements before accepting. However, the additional period of time provided for acceptance of a counteroffer can be terminated at any time by the original offeree’s revocation of the counteroffer.\textsuperscript{119}

\textsuperscript{115} Id. at 113.

\textsuperscript{116} See Contract Law of the People’s Republic of China (adopted and promulgated by the 2d Sess. of the 9th Nat’l People’s Cong. Mar. 15, 1999, effective Oct. 1, 1999) [hereinafter CCL], art. 23 (“An acceptance shall reach the offeror within the time limit fixed by the offer. If no time limit is fixed by the offer, the acceptance shall reach the offeror in accordance with the following provisions: (1) if an offer is made orally, acceptance shall be made promptly unless the parties stipulate otherwise; and (2) if an offer is not made orally, the acceptance shall reach the offeror within a reasonable period of time.”); BGB art. 146 (Ger.) (“An offer expires if a refusal is made to the offeror, or if no acceptance is made to this person in good time in accordance with sections 147 to 149.”); CISG, supra note 4, art. 18 (2) (“An acceptance is not effective if the indication of assent does not reach the offeror within the time he has fixed or, if no time is fixed, within a reasonable time, due account being taken of the circumstances of the transaction, including the rapidity of the means of communication employed by the offeror.”); \textit{Restatement (Second) of Contracts} § 41 (1981) (“An offeree’s power of acceptance is terminated at the time specified in the offer, or, if no time is specified, at the end of a reasonable time.”).

\textsuperscript{117} SCHLECHTRIEM & SCHWENZER, supra note 103, at 254.

\textsuperscript{118} See PECL, supra note 22, art. 2:207.

\textsuperscript{119} CCL, supra note 116, art. 18; \textit{Restatement (Second) of Contracts} § 42 (1981).
In some situations, the time of acceptance is so truncated there is little opportunity for a late acceptance. The CISG proffers “an oral offer must be accepted immediately unless the circumstances indicate otherwise.”\(^{120}\) In reality, this is no different than the general rule that an offeree must accept within a reasonable period of time. Reasonable time is determined by a contextual inquiry. The need to accept an oral offer immediately can be justified under the broader rule that an offer must be accepted within a reasonable period of time, such as prior to the termination of the telephone conversation in which the offer was made. For example, because stock and bond markets are volatile by nature, contracts are formed within seconds. A stockbroker telephones a client and states he can purchase a certain stock at a given price. The client replies by stating he needs a few minutes to look at his financials and hangs up the phone. A few moments later the client calls the stockbroker, accepts, and orders the stockbroker to purchase the stock. The problem is the offer reasonably lapsed at the time the offeree ended the first telephone conversation. By the time the client called back, the market price would have likely changed, and it would therefore have been unfair to preclude the stockbroker from selling the stock to another client after the termination of the first call.

Fourthly, differences arise concerning whether the original offeror’s reply (notice) can be withdrawn. In counteroffer theory, the original offeror’s reply is deemed an acceptance; however, such acceptance may be withdrawn if the withdrawal reaches the original offeree before or at the same time as the notice of acceptance.\(^{121}\) In effective acceptance theory, as discussed above, the original offeror’s notice of effective acceptance is a declaratory notice, which becomes effective once it has been dispatched. Therefore, it seems such notice cannot be withdrawn. However, it can be argued that CISG Article 22 can be applied by analogy to the offeror’s notice.\(^{122}\) The original offeror should be allowed to withdraw his or her notice of late acceptance if the withdrawal reaches the offeree prior to receipt of the notice. Schlechtriem and Schwenzer state: “the effectiveness of a declaration which only needs to be dispatched does not necessarily have anything to do with the declarer’s being bound by his declaration.”\(^{123}\) They reason the notice-withdrawal scenario is analogous to the offer-revocation situation where the offeror is allowed to revoke the offer prior to the offeree’s receipt of the offer. In fact, the offer may be revoked after receipt, unless the offer is determined to be an irrevocable offer.\(^{124}\) The issue

\(^{120}\) CISG, supra note 4, art. 18(2).

\(^{121}\) CCL, supra note 116, art. 27 (“An acceptance may be withdrawn. The withdrawal notice of the acceptance shall reach the offeror before or at the same time as the acceptance notice reaches the offeror.”); CISG, supra note 4, art. 22 (“An acceptance may be withdrawn if the withdrawal reaches the offeror before or at the same time, as the acceptance would have become effective.”).

\(^{122}\) See SCHLECHTRIEM & SCHWENZER, supra note 103, at 253, 314.

\(^{123}\) Id. at 314.

\(^{124}\) Article 16(2) creates a broad firm offer rule based upon the offeror fixing an amount of time upon which the offer will remain open, under Article 16(2)(a) or even if there is not such statement or assurance if it is “reasonable for the offeree to rely on the offer as being irrevocable.”
then becomes whether the notice of late acceptance is more like an offer or an acceptance. The principle of good faith in CISG Article 7(1) supports the idea that a rule that protects the addressee of such declarations is preferred. However, Schlechtriem and Schwenzer make a strong case that “where the addressee is not aware of a declaration on the ground that it has not yet reached him, it should be possible to withdraw it, since he has not yet acquired any position worthy of protection.”

The common law’s dispatch rule for effective acceptance results in a different answer. The authoritative opinion is the original offeror’s acceptance of a counteroffer binds the contract at the time of dispatch. Comment c to Restatement section 63 states that even if “the offeree has power to reclaim his acceptance from the post office or telegraph company [such an act] does not prevent the acceptance from taking effect on dispatch.” Therefore, it can also be argued that the offeror’s notice under the effective acceptance theory is akin to an acceptance, which binds the contract upon dispatch and makes it impossible for the offeror to withdraw the notice of effective acceptance.

Lastly, the time of the conclusion of a contract varies between the two theories. In counteroffer theory, according to the principle of “reach” under the BGB, the contract is formed when the original offeror’s “acceptance” of the late acceptance reaches the original offeree. According to effective acceptance theory, the dispatching of the offeror’s declaratory notice binds the contract retroactively to the time when the late acceptance was received. In contrast to the CISG and BGB, American common law treats the declaratory notice as binding the contract at the time of its dispatch. Therefore, depending on the applicable law a contract may be concluded at the time of the receipt of the late acceptance, at the time that the offeror dispatches a declaratory notice, or at the time that the declaratory notice reaches the offeree.

CISG, supra note 4, art. 16(2), 16(2)(b).

125. Id. art. 7(1).
126. SCHLECHTRIEM & SCHWENZER, supra note 103, at 315.
128. BGB art. 130 (Ger.). See also CCL, supra note 116, art. 26 (“The acceptance becomes effective when the acceptance notice reaches the offeror.”); CISG, supra note 4, art. 18 (2) (“An acceptance of an offer becomes effective at the moment the indications of assent reaches the offeror.”); CISG, supra note 4, art. 23 (“A contract is concluded at the moment when an acceptance of an offer becomes effective in accordance with the provisions of this Convention.”).
129. See SCHLECHTRIEM & SCHWENZER, supra note 103, at 254 (“The contract is formed not when the offeror gives written notice of approval to the offeree or orally informs him thereof, but retroactively at the time when the late declaration of acceptance reached the offeror...”); see also Commentary on the Draft Convention on Contracts for the International Sale of Goods prepared by the Secretariat, at 25, art. 19 cmt. 3, U.N. Doc. A/CONF. 97/5 (March 14, 1979) (“It is the late acceptance which becomes the effective acceptance as of the moment of its receipt, even though it requires the subsequent notice to validate it.”); HUBER & MULLIS, supra note 47, at 99.
5. Chinese Contract Law (CCL)

Generally, the gaps in the CCL need to be filled by the courts through analogical reasoning. This could be accomplished either through the use of other rules in the CCL (internal analogical reasoning), by using other similarly situated case decisions (common law), or, as is common in commercial law, recognizing an existing trade usage to fill in the gap (rule creation). The lack of a comprehensive set of late acceptance rules in the CCL is especially troubling, since the drafters of the CCL should have anticipated such a scenario in an offer-acceptance model of contract formation. A review of other national laws and international legal instruments show the BGB, CISG, and the PICC have adopted a complete set of late acceptance rules. This unnecessary gap in the CCL can be easily fixed through amendments.

a. CCL on Late Acceptance

Under CCL Article 26, acceptance becomes effective when the notice reaches the offeror. If the offer does not require a return notice or promise of acceptance, then it becomes effective when an act of acceptance is performed in light of trade practices or as indicated by the offer. The flaw in the CCL is it recognizes or can be interpreted as recognizing either the counteroffer or effective acceptance approaches. CCL Article 28 provides “if the offeree dispatches an acceptance beyond the time limit for acceptance, it shall constitute a new offer unless the offeror notifies the offeree in time that the acceptance is effective.” Despite the term “new offer,” the rule is a reflection of effective receipt theory.

Despite the wording of Article 28, most Chinese scholars prefer the counteroffer approach. One scholar argues:

The usefulness of the proviso [in Article 28] is questionable, since the late acceptance may be deemed as a new offer. The offeror’s notice of considering the late acceptance as effective, [if the law is to be re-drafted], shall be converted into ‘acceptance’ and the ‘without delay’ shall be replaced by ‘in a reasonable time.’

According to the above approach, Article 28 of the CCL should be interpreted to mean a late acceptance is always a counteroffer and can never be transformed into an effective acceptance by the original offeror. This includes the situation where the acceptance is properly dispatched by the offeree but is delayed in transmission at no fault of the offeree. Unfortunately, Article 28 does not distinguish a late acceptance due to a delay in transmission from one sent belatedly; therefore, all late acceptances are treated the same.

131. CCL, supra note 116, art. 28.
132. HAN SHIYUAN, CONTRACT LAW 100 (Law Press China 3d ed. 2011).
133. See A PROPOSITIONAL VERSION WITH REASONS FOR CIVIL CODE DRAFT OF CHINA 63, art. 886(1) (Liang Huixing ed., 2013) [hereinafter Liang, PROPOSITIONAL CCC].
Despite the wording of Article 28 and the scholarly commentary regarding it, this Article states that the effective acceptance theory is the preferable approach. First, compared with the counteroffer theory, effective acceptance theory protects the offeree’s reasonable expectations that a contract has been formed, especially in cases of a delay in transmission. In cases where the offeree knowingly sends a belated offer, then, depending on the situation, such expectations would be unreasonable. A pitfall of the counteroffer approach is it allows the offeror additional time to act opportunistically, since counteroffer theory affords the offeror a “reasonable time” (under CCL Article 23) to respond, while the late acceptance theory requires notice “without delay.” Therefore, the offeror has more time to monitor market price fluctuations before committing to the contract. At the other end, especially in the case of delayed transmission, the offeree is under the impression she has contracted at a fixed price and is unable to enter the market to stem her losses if her counteroffer is eventually rejected. The more time the law allows the offeror to decide, the greater the risk of speculation and uncertainty relative to the offeree. The effective acceptance theory provides greater certainty and is a more efficient rule that deters opportunistic behavior.

As a matter of efficiency, since contract formation has already been delayed due to the late arrival of the acceptance, contract law should provide an efficient rule that requires a faster, more certain way to contract formation. In the late acceptance situation, whether by belated dispatch or delayed transmission, the offeror is given the option to conclude or not conclude the contract. This is as it should be, but this option should not be subject to abuse. That is why effective acceptance theory requires a prompt response to the receipt of the late acceptance. Unlike an acceptance containing conflicting or additional terms (counteroffer), where the offeror may need time to consider the new or different terms, the late acceptance rules only apply if the late acceptance is an unequivocal acceptance of the offer (a valid acceptance but for its lateness). In the end, the offeror should be incentivized to make a prompt decision to contract or not to contract.

Second, under the effective acceptance approach, the conclusion of a contract is determined retroactively to the time of the arrival of the late acceptance. This provides security in the formation of the contract in that it fixes a certain time, as opposed to having to determine when the declaratory notice is sent or received. It also freezes the ability of the offeree to change her

134. Of course, the offeree may characterize its reply to offer as a counter offer either explicitly or by referring to the offer as already lapsed. If this is the case, the effective acceptance theory cannot be applied. See COMMON EUROPEAN SALES LAW (CESL): COMMENTARY 200 (Reiner Schulze ed., 2012) [hereinafter CESL].
137. See KÖTZ & FLESSNER, supra note 54, at 33.
mind by revoking her late acceptance prior to receiving the confirmatory notice from the offeror that a contract has been reached. In this regard, the effective acceptance approach benefits the offeror.

Third, since China adopted the CISG in 1988,\textsuperscript{138} any divergence between its rules and those in the CCL is inherently inefficient. The offer-acceptance rules in the CISG should have been uniformly transferred to the CCL. Unfortunately, the CISG acceptance rules transplanted to the CCL were not an exact copy, resulting in uncertainty as to their meaning in the CCL. Reforming the CCL to be consistent with the CISG would rectify this unfortunate divergence.

Lastly, the trend in the modernization of national sales laws and in international legal instruments such as the CISG has been toward adopting the effective acceptance approach.\textsuperscript{139} In the context of global efficiency, greater similarity among legal regimes will lead to more certainty, which will in turn tend to reduce impediments to international trade. However, the late acceptance rule as stated in the CISG is not without flaws. The primary flaw in Article 21 is the use of the word “dispatches,” which leads to a plausible interpretation that late acceptance becomes effective at the time of the dispatch of notice by the offeror, even if it is not subsequently received by the offeree. The better interpretation is that the general theory of receipt (adopted by the CISG) places the risk of transmission on the most efficient insurer of its receipt—the sending party (offeror). The purpose of Article 21 should be read as requiring a dispatch of notice “without delay,” but not as preempting the general theory of receipt adopted by the CISG. Therefore, the notice becomes effective at the time of the offeree’s receipt of the notice. A complete scheme of late acceptance rules would require the notice be dispatched “without delay” and the notice received within a reasonable time from dispatch. This is possible under the CISG, as noted above, with dispatch without delay being an express rule and the necessity of its receipt (within a reasonable time) being an implied general principle of the CISG.\textsuperscript{140}

Further support for the adoption or interpretation of the CCL as an effective acceptance regime, as well as a receipt rule for the purpose of sending notices to

\begin{footnotes}


\item[140] The implied general rule that a communication is only effective upon receipt is derived from the CISG’s offer and acceptance approach: an offer is good when it “reaches the offeree.” \textit{CISG, supra note 4}, art. 15(1). A withdrawal of an offer is valid “if the revocation reaches the offeree before he has dispatched an acceptance.” \textit{Id.} art. 16(1). An acceptance is effective “at the moment the indication of assent reaches the offeror.” \textit{Id.} art. 18(2). Thus, when the word “notice” is used, it is simply a substantive rule that notice must be given, but its effectiveness should be determined by the receipt rule.
\end{footnotes}
determine the effectiveness of the late acceptance, is found in the PICC. The PICC follows the effective acceptance theory of the CISG, but uses the word “reach” instead of “dispatch” when referring to the sending of notice by the offeror. PICC provides “[a] late acceptance is nevertheless effective as an acceptance if without undue delay the offeror so informs the offeree or gives notice to that effect.”141 As to when such notice becomes effective, it states: “A notice is effective when it reaches the person to whom it is given.”142 The PECL also requires the offeror to inform the offeree of the status or effectiveness of the late acceptance.143 A comment to the PECL’s late acceptance rules makes clear the notice must reach the offeree.144 Another comment expressly rejects the counteroffer theory:

Some legal systems treat a late acceptance as a new offer, which the offeror may accept within the time set for acceptance [within a reasonable period of time], which is often longer than the time provided in [Article 2:207(1): ‘without delay’]. The [PECL does not contain such a rule.145

Therefore, whether the lateness of the acceptance is due to a belated sending or a delay in transmission, the late acceptance can still be an effective acceptance.146 The need for the offeree to receive the offeror’s notice is the mainstream view supported by the receipt theory.

In sum, the CCL late acceptance rules are comprised of a variety of elements, some of which are not a good fit with others in the context of the national and international laws reviewed in this Article. First, CCL Articles 28 and 29 provide the standard bifurcated approach to late acceptance—for acceptances belatedly sent (Article 28) and cases of delayed transmission (Article 29). Article 28 adopts both the counteroffer and late acceptance approaches. Thus, a belatedly sent acceptance is a counteroffer that the original offeror has a reasonable time to accept unless the offeror “promptly” responds to accept the late acceptance (effective acceptance approach). Article 29 asserts in the delayed transmission scenario, the late acceptance is effective to conclude a contract unless the offeror promptly notifies the offeree otherwise.

141. PICC, supra note 21, art. 2.9(1).
142. Id. art. 1.9(2).
143. PECL, supra note 22, art. 2:207 (Late Acceptance).
144. Id. cmt. B (Assent to a Late Acceptance); id. illus. 1 (offeror’s notice “comes to [offeree’s] notice”).
145. Id. cmt. D.
146. Notes to the PECL, supra note 22, indicate which laws are in accordance with Article 2:207 and which ones are not. Regarding Article 2:207(1) (belated sending of acceptance), it lists the CISG (art. 21(1)), PICC (art. 2.9(1)), BURGERLIJK WETBEOK (BW) Article 6:223(1) (Neth.), Portuguese CC (art. 229), and Codice Civil (C.c.) Article 1326(3) (It.). Those laws that treat the late acceptance (belatedly sent) as a counteroffer include BGB Section 150 and ASTIKOS KODIKAS [A.K.] [CIVIL CODE] 19(1) (Greece). In the case of late acceptance due to a delay in transmission, the PECL is in accord with the CISG (art. 21(2)) and PICC (art. 2.9(2), as well as BGB Section 149, ASTIKOS KODIKAS [A.K.] [CIVIL CODE] 190 (Greece), and BURGERLIJK WETBEOK (BW) art. 6:223(1) (Neth.).
Unfortunately, Article 29 fails to include the common wording that the delay in transmission must have been obvious or known to the offeror. Unlike CISG Article 21(1)’s unfortunate use of the word “dispatch[es],” language in the CCL makes clearer notice provided by the offeror is good upon receipt—not upon dispatch. CCL Articles 28 and 29 require the offeror to promptly inform the offeree of the validity of said acceptance. The word “inform” indicates the offeree must actually receive the notice. This interpretation is consistent with the receipt rule, which is standardized throughout the CCL. Use of the effective receipt theory for both types of late acceptance would simplify these concepts by establishing a single rule—the offeror may activate the effectiveness of late acceptance if he informs the offeree to that effect “promptly” (without delay). Under this rule, notice becomes effective when it reaches the offeree, while the contract is concluded retroactively to the date when the late acceptance reaches the offeror. Furthermore, the offeror is free to withdraw his notice before or at the same time as when the notice reaches the offeree. Alternatively, if the counteroffer approach is retained in Article 28, the sub-optimal choice in such cases is that the offeror is allowed to accept the counteroffer (late acceptance) “in a reasonable time” instead of “without delay” so that he can speculate at the offeree’s expense.

b. Acceptance by Performance

The CCL recognizes unilateral contract formation in which a contract can be formed by the offeree “performing an act” in light of prior dealings between the parties, trade practices, or as otherwise indicated in the offer. Questions remain as to how one should define an “act,” and whether the offeree needs to provide notice it has performed the required act. CCL Article 26 addresses these questions by stating: “If an acceptance needn’t be notified, it becomes effective when an act of acceptance is performed in accordance with transaction practices

148. See CCL, supra note 116, art. 16 (An offer becomes effective when it reaches the offeree); id. art. 17 (An offer may be withdrawn, if the withdrawal notice reaches the offeree before or at the same time that the offer reaches the offeree); id. art. 18 (An offer can be cancelled if the revocation reaches the offeree before the offeree sends its acceptance); id. art. 20(1) (An offer can be extinguished if notice of rejection reaches the offeror); id. art. 23 (An acceptance shall reach the offoror before the notice of acceptance). The articles dealing with late acceptance do not use the words “dispatch” or “receipt.” CCL Article 28 (belated sending of acceptance) states: “[O]fferor informs the offeree of the effectiveness of the said acceptance promptly.” CCL Article 29 (delay in transmission) states: “[O]fferor informs the offeree promptly that it does not accept the acceptance.” Given the use of the receipt rule throughout the CCL formation rules, the only reasonable interpretation of Articles 28 and 29 is that the notice of effectiveness or non-effectiveness of the late acceptance must be received by the offeree.
149. See id. art. 22.
or as required in the offer.” 150 Thus, if the offer does not require notice and invites acceptance by performance, no notice is required. In determining what constitutes an “act” of acceptance, the court is to consider whether the offer provides a “definition” of the act needed to bind the contract. For example, the offer may state: “ship the goods immediately.” If the offer is silent as to what constitutes an act, the court is to look at “transaction practices.” 151 However, it is unclear what “transaction practices” actually entails. For example, do “transaction practices” entail prior dealings between the parties, trade usage, business custom, or some combination of these factors? The accepted interpretation of “transaction practices” includes practices developed between the parties (prior dealings). 152 Chinese commentary has asserted the act of acceptance may be based upon an “established long term relationship,” 153 as in the case of when silence may be a means of acceptance. 154

The key issue is whether the contract becomes binding at the beginning of the offeree’s performance or when the performance is completed. CCL Article 26 states acceptance occurs “when an act of acceptance is performed.” A strict interpretation of this provision would find that complete performance is required. However, this would be an inefficient interpretation, since it allows the offeror to revoke his or her offer prior to the offeree’s completion of performance, thereby producing wasted expenditures. Since Article 26 does not define “act of performance,” a more liberal and reasonable interpretation would be the contract is binding at the beginning of performance, unless the offer states otherwise. Professor Han argues for such an approach as a way of protecting the reasonable expectations of the offeree such that as long as he commences performance within a reasonable time, he is assured that a contract has been formed and is no longer subject to revocation. 155

Another issue is whether the offeree is required to notify the offeror of performance. Notice would not be required if the offeree did not provide notice in prior dealings with the offeror. As noted above, Article 26 uses the term

150. See id. art. 26.
151. Id.
152. See WANG LIMING, STUDY ON CONTRACT LAW 238 (2002) [hereinafter LIMING, STUDY OF CONTRACT LAW].
153. SHIYUAN, CONTRACT LAW, supra note 132, at 108-09.
154. The CCL does not contain a rule that provides for an exception whereby the offeree’s silence by the offeree can function as an acceptance. CCL Article 21 states “[a]n acceptance is a statement made by the offeree indicating assent to an offer.” However, given that Article 26 does not require notice for acceptance by performance “in accordance with transaction practices,” those same “transaction practices” should allow for acceptance by silence. See, e.g., RESTATEMENT (SECOND) OF CONTRACTS (1981) §69 (Acceptance by Silence); CISG Article 18(1) states: “Silence or inactivity does not in itself amount to acceptance.” The phrase “in itself” is interpreted to mean that usage, such as prior dealings, can make silence a means of acceptance. CISG COMMENTARY 268 (S. Kröll, L. Mistelis & P. Perales Viscasillas ed. 2011): “Circumstances which may indicate intent to accept [by silence] the offer include the agreement of the parties to that effect, the practices established between the parties, as well as usages that are binding upon the parties on the grounds of Art. 9 [international trade usage].”
155. SHIYUAN, supra note 132, at 108.
“transaction practices.” This term, however, is confusing due to the lack of a counterpart phrase in Western contract law. Usage may refer not only to practices developed between repeat players (prior dealings), but also to industry or business usage, customs, and practices. It seems Article 26 is referring to the former type of usage. Nonetheless, usage of the latter type, in which giving notice is not customary or expected, implies failure to give notice would not prevent a contract from becoming binding. Although not stated in the CCL, this interpretation is found in Chinese commentary. A further question is then presented: if the offeree sends a required notice in a unilateral contract situation when is the contract formed? Commentary suggests the contract is formed not by performance but upon receipt by the offeror of notice of the commencement of the act.

### c. CCL Late Acceptance in Practice

Although Chinese case law offers sparse applications of CCL Articles 28 and 29, existing scholarship discusses the meaning and flaws of those articles. First, the bifurcation of late acceptance between the two CCL Articles is less than ideal. Taken alone, Article 28 indicates all late acceptances are counteroffers that the original offeror is free to disregard. But Article 29 indicates that if the late acceptance is due to a delayed transmission, it constitutes an effective acceptance (unless the offeror notifies the offeree otherwise). However, if the offeror expressly fixes a date upon which the acceptance must be received, a late acceptance due to a delayed transmission would be ineffective.

Other Chinese commentary suggests the late delivery of an acceptance sent within the time allowed for acceptance is a “standard” type of late acceptance. Professor Wang Cheng emphasizes that the CCL fails to regulate the scenario in which an acceptance is sent within the period set for receipt of the acceptance (within a reasonable period of time) but is not received within the period implied in law (a reasonable period of time from the receipt of the offer). Therefore, by analogy to Article 29, a timely dispatch followed by untimely delivery (not due to a delay in transmission) is a late acceptance, which means it is treated as a counteroffer. That is, it is to be treated just like an acceptance that is sent late.

In the scenario of a belated dispatch, CCL Article 28 allows the offeror to convert the counteroffer into an effective acceptance by giving timely notice to that effect. The offeror has the option to accept the “late acceptance immediately

156. See SHIYUAN, supra note 132, at 109; LIMING, STUDY ON CONTRACT LAW, supra note 152, at 238.
157. Id. at 121 (which of the two sources in the previous footnote does this refer to?).
158. See, e.g., BGB § 148 (Ger.) (“Fixing a Period for Acceptance”).
159. PRINCIPLES OF NEW CONTRACT LAW AND COMMENTS ON RELEVANT CASES 67 (Cui Jianyuan ed., 1999).
160. Id. See also SHIYUAN, supra note 132, at 100.
or, just decide within a reasonable time whether to accept the new offer [based on market changes].” 161 Other scholars have criticized this view as embracing both the counteroffer and effective acceptance approaches to late acceptance, thus conflating the two theories. Professor Li Yongjun states:

If [we] interpret it literally, it should be that if the offeror notifies the offeree that he considers the late acceptance to be effective, then such late acceptance shall no longer constitute a counteroffer. But in the strict sense, such interpretation contradicts the rules of conclusion of a contract. 162

As noted earlier, Professor Han offers an explanation for Article 28’s “internal contradiction”—a late acceptance is a counteroffer but can also be an effective acceptance at the option of the offeror. 163

Based on the above critique, the Draft Civil Code of the People’s Republic of China (Draft Chinese Civil Code or Draft CCC), edited by Liang Huixing, suggests amending CCL Article 28 to state: “late acceptance of an offer is [always] considered to be a new offer” 164 and cannot be converted to an effective acceptance by the offeror. We argue in the next sub-section that such an approach is not optimal.

Very few Chinese scholars have commented on the meaning of the late acceptance provisions found in CCL Article 29. 165 However, further analysis shows differences between Article 29 and BGB Article 149, as well as, CISG Article 21(2). Both the BGB and CISG emphasize that if the offeror does not want to be bound by a late acceptance due to late dispatch, he need not respond. However, under the BGB and CISG, if the offeror knew or “ought” to have known the late acceptance was caused by a delay in transmission, he is required

161. But even the author himself has to admit that in practice, in order to avoid the opportunistic behavior of the offeror, even when the late acceptance has been considered a new offer by the offeror, the offeror still has to accept it on time, instead of within a reasonable time. LIMING, STUDY OF CONTRACT LAW, supra note 152, at 242; see also THE DRAFT CIVIL CODE OF THE PEOPLE’S REPUBLIC OF CHINA: ENGLISH TRANSLATION art. 1297 (Liang Huixing ed., 2010) [hereinafter DRAFT CCC] (“If the offeree makes an acceptance beyond the time limit for acceptance, it shall constitute a new offer unless the offeror notifies the offeree in time that the acceptance is effective.”); THE DRAFT CIVIL CODE OF THE PEOPLE’S REPUBLIC OF CHINA AND ITS LEGISLATION REASONS: GENERAL PROVISIONS OF OBLIGATIONS & CONTRACTS 223 (Wang Liming ed., 2005) [hereinafter Wang, LEGISLATION REASONS].

162. LI YONGJUN, CONTRACT LAW 121 (2004).

163. SHIYUAN, supra note 132.

164. Liang, PROPOSITIONAL CCC, supra note 133, art. 886(1), at 63-64.

165. See Wang, LEGISLATION REASONS, supra note 161, at 224 (following the provision of CCL Articles 28 and 29 and stating that if an offeree makes within the time limit for acceptance an acceptance that could reach the offeror in time under normal conditions but happens to reach the offeror beyond the time limit due to other reasons, the acceptance shall nevertheless be effective unless the offeror notifies the offeree in time and the acceptance is denied due to its delayed arrival); see also Liang, PROPOSITIONAL CCC, supra note 133, at 63-64 (noting the difference between CCL Article 29 and BGB Article 150 or CISG Article 21(2) and amending CCL Article 29 along these lines: If the offeror could know if its transmission had been normal and the late acceptance would have reached the offeror in due time, the offeror shall inform the offeree of the lateness of the acceptance without undue delay; otherwise such late acceptance shall be deemed as not being late.).
to notify the offeree that the offer has lapsed.\textsuperscript{166} If he fails to provide notice, the late acceptance is effective and binds the contract. If the offeror did not or should not have known of the delay in transmission, he can presume the late acceptance was due to belated dispatch.\textsuperscript{167}

On the other hand, CCL Article 29 does not provide such safeguards. Under CCL Article 29, as noted previously, even if the offeror does not know and could not know the offeree had dispatched the acceptance on time but had simply experienced a delay in transmission, the offeror is still obliged to notify the offeree if he does not want to be bound by the late acceptance. This places an undue burden on the offeror who reasonably believes he had received a late acceptance and is allowed to disregard it as a counteroffer.\textsuperscript{168} Professor Ye Jinqiang argues in cases where the offeror had no indication the acceptance had been delayed in transmission, the requirement of notifying the offeree of the lateness of acceptance under CCL Article 29 is not appropriate both logically and as a value judgment.\textsuperscript{169}

d. Reforming the CCL Late Acceptance Regime

Both the common core and “better rule” comparative law methodologies support the effective acceptance approach, as opposed to the counteroffer approach, to late acceptance. The CCL should be reformed, or the proposed Draft Chinese Civil Code should be drafted,\textsuperscript{170} to fully embrace effective acceptance theory. That said, it must be reiterated that late acceptance cannot be adequately captured by a single rule; instead, it requires a batch of rules. This Article has discussed multiple scenarios and the mix of potential rules that may be used.

An effective reform of the CCL late acceptance rules would need to answer the following questions: (1) how the law should deal with a belatedly sent

\textsuperscript{166} See Huber & Mullis, supra note 47, at 99; European Contract Law, supra note 54, at 34.

\textsuperscript{167} See Huber & Mullis, supra note 47, at 98-99.

\textsuperscript{168} It is possible that CCL Article 29 was transplanted partially from Article 159(1) of the Civil Code of the Republic of China (Taiwan Civil Law). See Draft CCC, supra note 161, art. 159(1). Prior to the amendment to Article 159(1) of the Civil Code, CCL Article 29 was almost identical in establishing that there is no such obligation on the offeror “if the offeror could have known.” However, Article 159 of the Civil Code of the Republic of China was amended on April 2, 1999. The amendment indicated that the original writing of Article 159 made unclear as to whether the offeror is only obligated to notify the offeree of the lateness of acceptance on the condition that the offeror could have known that the lateness was caused by a delay in transmission when the late acceptance was sent under circumstances such that it would have reached the offeror in due time had its transmission been timely. In order to minimize ambiguity and protect the rights and interests of the offeror, this amendment to Article 159(1) has provided the additional obligation of the offeror to notify the offeree “if the offeror could have known.”

\textsuperscript{169} See Ye Jinqiang, supra note 147, at 90.

\textsuperscript{170} Towards a Chinese Civil Code: Comparative and Historical Perspectives (Chen Lei & C.H. von Rhee eds., 2012) (provides history behind the current project of the drafting and enacting a Chinese Civil Code).
acceptance; (2) how the law should deal with a late receipt of a timely dispatched acceptance that has been delayed in transmission; (3) how the law should deal with the issue of late acceptance in a unilateral contract (acceptance by performance); (4) at what time a contract should be binding in different late acceptance scenarios; (5) what the notice obligation of the offeror in accepting or not accepting a late acceptance should be; (6) what the notice obligation of the offeree in a scenario of late acceptance by performance should be; (7) what rights the offeror should have in revoking his notice of effective acceptance; and (8) what rights the offeree should have to withdraw his late acceptance between its receipt and the offeror’s dispatch of notice.

Based upon the review of late acceptance rules, the effective acceptance theory is preferable for three main reasons.\(^\text{171}\) First, compared with counteroffer theory, the effective acceptance theory better protects the offeree against offeror speculation, which is heightened under the former approach due to the extended time that the reasonable period of time standard provides. Since the offeror is entitled to choose to form a contract or not, he can make his decision in light of market changes.\(^\text{172}\) Under the effective acceptance theory, however, the offeror is not entitled to a “reasonable time” to provide notice. He must give notice of an effective acceptance “without delay.” As compared to a “reasonable” time, “without delay” has been interpreted to mean a much shorter time period.\(^\text{173}\) Effective acceptance theory protects the reasonable expectations of the offeree and deters opportunistic behavior of the offeror. Thus, it is the more efficient rule.\(^\text{174}\)

Second, under the effective acceptance approach, the contract is concluded at the time of arrival of the late acceptance, which is when the parties would naturally assume the contract to be binding.\(^\text{175}\)

Third, since China adopted the CISG prior to enacting the CCL, the theory of effective acceptance, presented in CISG Article 21(1), was already the law in China. Professor Feng Datong correctly states:

\begin{quote}
[I]n order to favor the conclusion of [a] contract, [the] CISG has taken a flexible method to provide that under certain conditions the late acceptance can be deemed as an effective acceptance and the contract can be concluded accordingly . . . If Seller is still interested in concluding a contract with Buyer, he may notify Buyer, without undue delay, that he considers Buyer’s acceptance as being effective. If so, the contract is concluded accordingly, and the date of conclusion is the date when the late acceptance arrives to Seller.\(^\text{176}\)
\end{quote}

\(^\text{171}\) Of course, the offeree may characterize its reply to an offer as a counter offer either explicitly or by referring to the offer as having already lapsed. If this is the case, the effective acceptance theory cannot be applied. See CESL, supra note 134, at 200.

\(^\text{172}\) See BIANCA & BONELL, supra note 135, at 192-93.

\(^\text{173}\) PECL, supra note 22, art. 2:207 (“Some legal systems treat a late acceptance as a new offer which the offeror may accept within the time set for acceptance which is often longer than the time [without delay] provided in paragraph 1.”); see also PECL, supra note 22, at 177.

\(^\text{174}\) JINGXIA, supra note 136, at 58-59.

\(^\text{175}\) KÖTZ & FLESSNER, supra note 54, at 33.

\(^\text{176}\) FENG DATONG, INTERNATIONAL TRADE LAW 51 (1995); see also WU JIANBING ET AL.,
In the interest of consistency, the divergent rules in the CISG and the CCL on late acceptance should be reconciled. Given the CCL adopted many rules directly from the CISG, the most reasonable solution would be to amend the CCL’s late acceptance rules to be consistent with those found in the CISG.

In summary, the proposed changes to the current rules would answer the questions posed at the beginning of this Part as follows:

1. The offeror may treat an untimely dispatched acceptance as an effective acceptance by providing notice to the offeree without delay;
2. A timely dispatched acceptance that is delayed in transmission (by no fault of the offeree) should be considered an effective acceptance (unless the offeror duly notifies the offeree of the late delivery and that he is treating the offer as having lapsed);
3. In the case of a late acceptance in a unilateral contract (acceptance by performance), the offeree should be required to provide notice of a delay in performance, and the offeror should be allowed to revoke his offer without delay. In return, the offeree should give notice that the offer has lapsed, or that he intends to accept the late performance as binding the contract;
4. In both the late acceptance scenarios involving an acceptance by promise (bilateral contract), the contract should be binding at the time of the arrival of the late acceptance. In a unilateral contract, the fairest rule would be to bind the contract at the time the offeree begins performance under the presumption that the offeree will complete performance within a reasonable time;
5. In the case where there is a belated dispatch of an acceptance, the offeror has no duty to notify the offeree of the receipt of the late acceptance. The offeror should have a duty to notify the offeree of his intent to accept a late acceptance in cases where the lateness was due to a delay in transmission that the offeror knew or should have been aware of;
6. In the case in which the offeror sends a notice of effective acceptance, he should be allowed to revoke that notice in the event that the revocation reaches the offeree on or before the notice; and
7. The offeree should not be allowed to revoke the late acceptance between its receipt and offeror’s dispatch of notice of effective acceptance. In such a scenario, the contract is binding retroactively to the time of receipt of the late acceptance. This should only be the case until the offeror’s “without delay” period has expired. If the “without delay” period has expired, the late acceptance is not effective. Note, however, that in cases of belated acceptance, the offeree could conceivably insert a provision limiting the time period during which the offeror must provide notice of accepting the late acceptance.

The above set of late acceptance rules best protects the expectations of both parties. It also allows for the consummation of a contract if that is, in fact, what the parties desire. The law should recognize a presumption that the offeree is

INTERNATIONAL BUSINESS LAW 120 (2007).
178. Note, however, that in cases of belated acceptance, the offeree could conceivably insert a provision limiting the time period during which the offeror must provide notice of accepting the late acceptance.
willing to conclude a contract with the original offeror at the time the offeree sends his late acceptance. Late acceptance rules are a way of recognizing the parties’ intent to form a contract despite the late acceptance. The ability of the offeror to conclude the contract by sending notice to the offeree is equivalent to the offeree’s intent in sending the late acceptance.

The question of when the contract comes into being is an important one. There are three options: (1) at the time the offeror dispatches the notice of effective acceptance; (2) at the time the offeree receives notice of offeror’s acceptance; or (3) at the time the offeror receives acceptance (as long as the offeror promptly dispatches notice to the offeree).

The first option has been dismissed previously as nonsensical under the receipt theory. Specifically, it places the risk of delayed or lost transmission on the wrong party. The most efficient insurer is the sending party, not the receiving party. This is the rationale that pervades the civil law and the CISG’s receipt rule, as opposed to the “mailbox rule” of the common law (acceptance effective upon dispatch). The second option is inconsistent with the offeror’s declaration (notice) of effective acceptance. If the acceptance is effective, then under the offer-acceptance model, it binds the contract upon receipt by the offeror.

The third choice is preferred because it protects the offeror from withdrawal of the late acceptance. Treating the late acceptance as merely a counteroffer may result in injustice to the offeror. If the late acceptance is treated as a counteroffer, a subsequent revocation defeats the reasonable expectations of the original offeror intending to send notice of effective acceptance. CISG Article 16 states the right of the offeror to revoke ends upon the dispatch of the acceptance even though the acceptance is not effective upon receipt. By analogy, the offeree’s right to withdraw a late acceptance should be prevented during the “without delay” period, allowing the offeror to convert the late acceptance into an effective acceptance. Furthermore, upon the offeror’s dispatch of its notice of late acceptance, the right of the offeree to revoke (counteroffer theory) should be extinguished.

In the end, the particular rules adopted in relation to the offer-acceptance paradigm are less important than the adoption of a set of rules that are comprehensive and internally consistent. An established English treatise on contract law states:

[T]he phrase ‘offer and acceptance’ . . . is not to be applied as a talisman, revealing, by a species of esoteric art, the presence of a contract . . . The rules which the judges have elaborated from the premise of offer and acceptance are neither the rigid deductions of logic nor the inspiration of natural justice. They are only presumptions, drawn from experience, to be applied in so far as they serve the ultimate object of establishing the phenomena of agreement.

179. See CISG, supra note 4, 18(2) (accepts civil law’s receipt rule over the common law’s dispatch or mailbox rule).

The rule changes proposed in this Article are only the first step in a process. The consistency and comprehensiveness of transplanted rules is more a creation of the interpretation of the rules than the text of the rules. This is especially the case, we argue, when the sources of the rules are in one language and are then translated into the language of the receiving country. It is for the courts through interpretation, or the government through amendment, to remove the gaps and inconsistencies invariably found in translated texts. Incidentally, the problem reappears when the native language version of the country of transplant (i.e., Chinese) is translated “back” into the source rules’ native language.

The next section provides an example of such a problem through a discussion of anticipatory breach and adequate assurance. In the former FECL, a seemingly more demanding threshold of a “sure guarantee” was required to defeat a declaration of anticipatory breach. Such a standard of assurance of performance is well beyond that found in any other law. The English translation of the CCL has rectified such a burdensome and unwieldy standard. That said, different translations of the CCL have used a variety of phrases, such as “appropriate assurance,” “adequate assurance,” and “guarantee.” However, since the CCL fails to define “adequate,” the aforementioned phrases are sufficient because they all expunge the word “sure.”

B. Anticipatory Breach

“Anticipatory breach,” referred to in the common law as “anticipatory repudiation,” is a manifestation—express or implicit—by one party to the other that the first cannot or will not perform at least some of its obligations under the contract at the time set for performance. In such cases, under certain

---

181. Simone Glanert, Speaking Language to Law: The Case of Europe, 28 LEGAL STUD. 161, 165 (2008) (when text is “translated back into the national language, this language becomes the object of an interpretation by the national judge in every specific case.”). This interpretative process captures the text within the cultural and legal traditions of the transplanting country and leads to different meanings from those given the text from the country or legal system from which it was borrowed. The success of transplantation can be measured by whether the different meanings coalesce into a consistent whole.


183. See U.C.C. § 2-609 cmt. 4 (AM. LAW INST. & UNIF. LAW COMM’N 2014); PECL, supra note 22, art. 8:105 cmt. d.


186. See FARNSWORTH ON CONTRACTS, supra note 46, at 558.
circumstances, the party receiving information of a prospective breach can act in
advance of the breach to repudiate the contract. This allows the non-breaching
party to terminate the contract in advance of the breach (i.e., non-delivery of
goods) and commence an action for damages. \footnote{187} The concept of anticipatory
breach was first established in \textit{Hochster v. De la Tour} \footnote{188} and has since been
universally accepted by common law countries, civil law countries, and
international private law instruments, including the CISG and the PICC. \footnote{189} The
rationale behind the doctrine is that the contracting party has the right to expect
not only that the other party will perform when the time comes, but also that it
will do nothing to substantially impair that expectation before the time comes
for performance. \footnote{190}

\textbf{1. Seriousness of Breach}

For an anticipatory breach to have legal effect, the threatened breach must
be serious and based upon credible information. According to the \textit{Restatement}
and the UCC, \footnote{191} the prospective non-performance must be serious enough that
the injured party is able to treat it as a “total breach,” which is defined as a
\textit{substantial impairment} of the contract. \footnote{192} Under the CISG, if, before the
performance date, it becomes “clear that one of the parties will commit a
fundamental breach of contract,” the other party can declare the contract to be
void. \footnote{193} The PICC states that it must be “clear that there will be a \textit{fundamental}
non-performance.” \footnote{194} Finally, the CESL more vaguely states, “the non-
performance would be such as to justify termination.” \footnote{195}

\textbf{2. Express Repudiation and Reasonable Grounds for Implied
Repudiation}

The breach may be anticipated by words or by conduct. Usually, a breach
consists of a statement of the repudiating party that it cannot or will not perform.
The statement must be sufficiently affirmative such that a reasonable person

\begin{footnotesize}
\begin{enumerate}
\item \footnote{187} See id. at 565-68.
\item \footnote{188} Hochster v. De la Tour (1853) 118 Eng. Rep. 922 (Que.).
\item \footnote{189} BGB § 323 (Ger.); CISG, \textit{supra} note 4, art. 72(1); U.C.C. §§ 2-610, 2-611 (AM. LAW
INST. & UNIF. LAW COMM’N 2014); \textit{RESTATEMENT (SECOND) OF CONTRACTS §§ 250-257}
(1981); PICC, \textit{supra} note 21, art. 7(3)(3).
\item \footnote{190} See \textit{FARNSWORTH ON CONTRACTS}, \textit{supra} note 46, at 551-54.
\item \footnote{191} \textit{RESTATEMENT (SECOND) OF CONTRACTS §§ 250-57} (1981); U.C.C. § 2-610 (AM. LAW
INST. & UNIF. LAW COMM’N 2014).
\item \footnote{192} \textit{RESTATEMENT (SECOND) OF CONTRACTS § 243} (1981) provides a measure for
determining total breach: “a breach of non-performance [or anticipated non-performance] gives rise
to a claim for total breach only if it so substantially impairs the value of the contract to the injured
party.” \textit{See also} U.C.C. § 2-610 (AM. LAW INST. & UNIF. LAW COMM’N 2014) (“substantially impair
the value of the contract. . .”).
\item \footnote{193} CISG, \textit{supra} note 4, art. 72(1).
\item \footnote{194} PICC, \textit{supra} note 21, art. 7.3.3.
\item \footnote{195} CESL, \textit{supra} note 134, at 116.
\end{enumerate}
\end{footnotesize}
would understand it to mean the breach will actually occur. A party may repudiate by conduct as well. In that case, an anticipatory breach entails a party’s voluntary affirmative act that renders the party actually or apparently unable to perform. Since the conduct must be an affirmative act, mere delay in performance is not an anticipatory breach.

Since the act must be voluntary, inability to perform due to incompetence or financial difficulties is also not an anticipatory breach. However, if such circumstances give the other party reason to believe the first party will commit a breach, the other party is entitled to exercise a right to self-help by suspending its own performance until the first party performs or provides adequate security relating to the future performance. UCC Section 2-609(1) provides that “[w]hen reasonable grounds for insecurity arise with respect to the performance of either party the other [party] may in writing demand adequate assurance of due performance and until he receives such assurance may if commercially reasonable suspend any performance for which he has not already received the agreed return.” Further, UCC Section 2-609(4) states, “[a]fter receipt of a justified demand failure to provide within a reasonable time not exceeding thirty days such assurance of due performance as is adequate under the circumstances of the particular case is a repudiation of the contract.”

The repudiating party would be in breach if it does not have reasonable grounds for suspending performance or for demanding adequate assurance. If a party does not have reasonable grounds, the court should question whether the party nefariously made the demand in the hope of triggering a breach. This would violate the anticipating party’s duty to act in accordance with the principle of good faith.

3. Anticipatory Breach under the CCL

China “transplanted” the concept of anticipatory breach from its 1985 Foreign Economic Contract Law (FECL), which states:

A party may temporarily suspend its performance of the contract if it has conclusive evidence that the other party is unable to perform the contract.

196. See FARNSWORTH ON CONTRACTS, supra note 46, at 559.
197. See Pappas v. Crist, 25 S.E.2d 850 (N.C. 1943) (where owner leased premises to another lessee, this was an “unequivocal and absolute renunciation of the entire agreement to make the lease to the plaintiff”); FARNSWORTH ON CONTRACTS, supra note 46, at 563-64.
198. FARNSWORTH ON CONTRACTS, supra note 46, at 564.
199. Id. at 572-73.
201. See U.C.C. § 1-304 (AM. LAW INST. & UNIF. LAW COMM’N 2014); CCL, supra note 116, art. 6.
202. The FECL was promulgated on March 21, 1985, became effective as of July 1, 1985, and was simultaneously annulled when the CCL was promulgated in 1999. See generally WANG LIMING, LIABILITIES FOR BREACH OF CONTRACT 146-47 (1996); Nan Zhengxing & Guo Dengke, Comparative Study on Anticipatory Breach, 84 CHINESE J. L. 71-76 (1993). But see Han Shiyuan & Cui Jianyuan, Anticipatory Breach and Chinese Contract Law, 86 CHINESE J. L. 33-38 (1993).
However, it shall immediately inform the other party of such suspension. It shall perform the contract if and when the other party provides a sure guarantee for performance of the contract. If a party suspends performance of the contract without conclusive evidence of the other party’s inability to perform the contract, it shall be liable for breach of contract.203

The FECL places a heavy burden on the non-breaching party; it substitutes the reasonableness standard (“reasonable grounds”) with a “conclusive evidence” standard. The latter sets a much higher standard to advance a claim of anticipatory breach. It also severely deters its use since the anticipating party will be deemed to be in breach if its evidence is determined to be inconclusive. Furthermore, the FECL does not give the non-breaching party a right to terminate the contract.

The shortcomings of the FECL’s anticipatory breach rules were at least implicitly acknowledged when China ratified the CISG in 1986.204 The CISG incorporates anticipatory breach in Articles 71 and 72. These articles are similar to the rules found in the American UCC. The drafters of the CCL recognized the importance of anticipatory breach, but unfortunately decided not to transplant the better, simpler rules found in the CISG. Instead, they decided to amend the rules found in the FECL. Unfortunately, these changes did not amount to a significant improvement.205 First, the changes provide a convoluted set of rules found in different places in the CCL (Articles 68, 69, 94, and 108). Second, Articles 68 and 69 retain the extreme language of “conclusive evidence,” although the CCL, as interpreted, has replaced the notion of providing a “sure guarantee” found in the FECL in favor of something more reasonable, such as the need to provide “adequate assurance” of performance.206

The CCL has dual sets of rules or concepts dealing with non-performance that can easily be conflated. Anticipatory breach is found in Articles 94 and 108, while “defense of insecurity” is found in Articles 68 and 69. The rest of this section and the next one will analyze this dual system relating to a party’s fear of non-performance.

The CCL rules are unclear on the relationship between “defense of insecurity” and the right to anticipatory breach. “Defense of insecurity” is a term used in Chinese law, which is different from the concept of insecurity found in the UCC and Restatement. The Restatement bases the right to declare an anticipatory breach on the appearance of reasonable grounds for insecurity with respect to the performance of the other party. As such, parties to a contract are entitled to “a continuing sense of reliance and security that the promised

203. FECL, supra note 182, art. 17.
204. China signed the CISG on September 30, 1980, which was ratified on December 11, 1986 and became effective on January 1, 1988.
205. Han Guijun & Xiao Guangwen, Comparative Studies of Remedies for Anticipatory Breach of Contract, in HEBEI L. SCI. (2004); Li Wei & Huang Hui, On Defense of Insecurity (Einrede der Unsicherheit) and Anticipatory Breach, in MOD. L. REV. (2002); Li Yongjun, supra note 162, at 594-96.
performance will be forthcoming when it is due." It provides the further reasoning that a sense of security is an implied part of the contract. Furthermore, the nature of the ground for insecurity is a factor in determining the nature, or type, of adequate assurance to be provided to remove the declaration of anticipatory repudiation.

Unlike the UCC and Restatement, defense of insecurity in the CCL has nothing to do with anticipatory repudiation. The Chinese concept of defense of insecurity originated from the concept of Einrede der Unsicherheit ("Defence of uncertainty") in the German Civil Code. Under the defense of insecurity concept in the CCL (implied in CCL Articles 68 and 69), the party who should perform first may suspend his performance or even terminate the contract if certain circumstances are met (e.g., insecurity relating to the other party’s reciprocal performance). There is no reciprocal right of the breaching party to provide adequate assurance. It is important to note Articles 68 and 69 by their express words only allow for the suspension of the contract; they do not provide for a right of termination. Rather, it is Articles 94 and 108 that allow for termination. In sum, both sets of articles are “anticipatory” in nature, but each pair of articles provides different remedies (suspension versus termination). This bifurcation serves no reasonable purpose and has caused much confusion and debate.

4. CCL Articles 94 and 108

The traditional notion of anticipatory breach is found in CCL Articles 94 and 108. CCL Article 94(2) states that the parties may terminate “before the period of performance expires, [if] either party clearly indicates by word or by act that it will not discharge the principal debts.” CCL Article 108 states: “If either party explicitly expresses or indicates by act its intention not to perform its obligations under the contract, the other party may, before the expiration of the period of fulfillment, demand that the party in question bear the liability for breach of contract.” These phrases recognize express anticipatory breach (“clearly indicates by word;” “explicitly expresses”) and implied anticipatory breach (“by act;” “indicates by act”).

It is important to note there is a difference between non-performance and delayed performance. Therefore, if the party only indicates it could not perform

208. Id. cmt. e (Nature and time of assurance).
209. BGB § 321 (Ger.).
210. See Han & Xiao Comparative Studies of Remedies for Anticipatory Breach of Contract, supra note 205, at 38-43; Li & Huang, On Defense of Insecurity (Einrede der Unsicherheit) and Anticipatory Breach, supra note 205, at 54-57; Li YONGJUN, supra note 162, at 594-96.
211. The Restatement gives the example of an “act” as one when there are a series of minor breaches by one of the parties; those acts may be grounds for a declaration of implied anticipatory breach. RESTATEMENT (SECOND) OF CONTRACTS § 251 cmt. c (1981) ("[M]inor breaches may give reasonable grounds for a belief that there will be more serious breaches.").
the contract in due time, the party is not liable for anticipatory breach if it can perform within a reasonable time of the due date. CCL Article 94(3) states the right to terminate commences when the delayed performance moves beyond a reasonable time after the non-performing party has been urged to perform. The language “after being urged” seems to put a duty on the non-breaching party to make a demand for performance.\textsuperscript{212} If the breaching party attempts to perform its delayed performance within a reasonable time, it would seem that under the principle of good faith\textsuperscript{213} the party would be required to keep the non-breaching party aware of the progress of that performance.

5. Defense of Insecurity in CCL Articles 68 and 69

As discussed above, Article 108 provides a broader rule, while Article 68 is narrower in scope. This narrowness is due to two elements. First, it allows only for suspension and not termination of the contract. However, the failure to provide a right to terminate is alleviated by Article 69, which allows the suspending party to terminate the contract after suspension of performance if the other party fails to “reinstate its capacity of performance and does not provide a sure guarantee” of performance.

Second, Article 68 permits a party that must ordinarily perform first to suspend its performance if it can provide conclusive evidence that the other party faces any of the following circumstances: (1) serious deterioration of its business conditions; (2) diversion of its properties and secret withdrawal of capital to evade debts; (3) loss of business credibility; or (4) other situations showing inability or possible inability to meet liabilities. The best that can be said for the dual systems in the CCL—defense of insecurity in Articles 68 and 69 with anticipatory breach in Article 94 and 108—is that unlike FECL Article 17, the CCL provides a right to terminate the contract in cases of either express or implied anticipatory breach. Nevertheless, the CCL also has shortcomings. These shortcomings are discussed in the next section.


The two-pronged approach found in CCL Articles 94 and 108 (anticipatory breach) and Articles 68 and 69 (defense of insecurity) have numerous shortcomings, causing a great deal of confusion. First, the remedial consequences of repudiation are unclear. The CCL only provides that the non-breaching party may terminate the contract and/or hold the repudiating party liable for breach of contract. However, there is no specific provision that allows

\footnotesize{\textsuperscript{212} CCL, supra note 116, art. 94(3) (“[S]ill fails to discharge them within a reasonable period of time after being urged”).}

\footnotesize{\textsuperscript{213} See CCL, supra note 116, art. 5 (“[T]he parties shall abide by the principle of fairness in defining the rights and obligations of each party.”); see also CCL, supra note 116, art. 6 (“[T]he parties must act in accordance with the principle of good faith, no matter in exercising rights or in performing obligations.”).}
the non-breaching party to suspend his performance and require the repudiating party to provide adequate assurance.

Second, taking a literal interpretation, CCL Article 108 unduly enlarges the application of anticipatory breach. Namely, it permits a non-breaching party to terminate the contract regardless of the degree of the breach (material, minor, or de minimis). As noted above, anticipatory breach under CISG Article 71 must be serious in nature: “the other party will not perform a substantial part of his obligations.” However, CCL Article 108 provides that as long as one party indicates his intention not to perform his “obligations” no matter how trivial the breach, the other party is entitled to hold the first party liable for anticipatory breach. An alternative, more reasonable interpretation of the language of CCL Article 108—“will not perform its obligations under a contract”—is the party will not perform at all, or at least not in a substantial way. This interpretation would limit the power to anticipate breach for lesser degrees of non-performance. This view is supported by the CISG, which requires a “fundamental breach” in order to avoid or terminate a contract.214

Third, the relationship between defense of insecurity and anticipatory breach is ambiguous. From the perspective of contextual interpretation and legislative history, it seems these terms, although related, are conceptually different. The defense of insecurity is transplanted from civil law, especially from BGB Article 321,215 while anticipatory breach is transplanted from CISG Articles 71 and 72, as well as common law.216 However, other than the fact that defense of insecurity is only applicable to contracts in which the parties are not expected to perform concurrently, the substance of the two systems are identical in function.217 Whenever the defense of insecurity is applicable, anticipatory breach is also applicable. Despite playing similar functions, their inclusion in the same law is problematic from the perspective of interpretation since they have different origins and are found in different chapters of the CCL. Namely, Articles 68 and 69 are found in the chapter on “Fulfillment of the Contract,” Article 94 is found in the chapter on “Termination of Rights and Obligations under the Contract,” and Article 108 is found in the chapter on “Liability for Breach of Contract.”

As a result, some Chinese courts have wrongfully based their decisions on both Articles 68 and 69 and Articles 94 and 108, while simultaneously equating anticipatory breach with defense of insecurity.218 Some Chinese courts have

214. See CISG, supra note 4, arts. 25, 49(1), 64(1).
215. BGB § 321 is entitled “Defence of uncertainty.”
217. Han & Xiao, supra note 205, at 38-43 (explaining that this is the reason some foreign scholars hold the opinion that CCL Articles 68 and 69 are the “anticipatory breach” provisions.). See also LARRY A. DIMATTEO & LUCIEN J. DHOOGE, INTERNATIONAL BUSINESS LAW: A TRANSACTIONAL APPROACH 229 (2d ed. 2006).
confused the different systems and held CCL Articles 94 and 108 relate to the defense of insecurity, instead of anticipatory breach.\footnote{The \textit{CISG and Modernisation of Chinese Contract Law}, 18 COMP. L.J. THE PAC. 75 (2014) (contributions to the Study of International Trade))}

Fourth, while the above conflation of CCL articles may seem irrelevant because the two concepts serve the same function, the two concepts do in fact conflict.\footnote{See, e.g., \textit{Zhejiang Province Ningbo Hongtu Paper Prods. Indus. & Trade Co. v. Zhejiang Ningbo Jingying Zhiban Color Printing Co.}, (Zhejiang Ningbo Intermediate People’s Ct. 2011) YJSCZD No. 231; \textit{ZYSZZ} No. 30, http://gdlawyer.chinalawinfo.com/newlaw2002/slc/SLC.asp?Db=fnl&Gid=119235901.} For example, Article 68 allows that a party “may” suspend performance in defense of insecurity. But the “may” is converted to a “must” if that party decides to terminate the contract under Article 69. According to CCL Article 69, \textit{only if} the other party has failed to regain his capability of meeting its liabilities and to provide an assurance within a reasonable time, the injured party can terminate the contract. In addition, under Article 69 the non-breaching party must have conclusive evidence to believe the other party will not or cannot perform its obligations and must promptly notify the other party of the suspension. However, if the other party regains the capability to perform and provides an adequate assurance or guarantee, the non-breaching party must continue to perform the contract.\footnote{See \textit{Hunan Debang Med. Co. v. Hunan Liye Gucheng Biotech. Co.}, (Changsha Intermediate People’s Ct. 2009) CZMEZZ No. 0231, http://gdlawyer.chinalawinfo.com/newlaw2002/slc/SLC.asp?Db=fnl&Gid=118746766.} Thus, under the defense of insecurity approach, suspension of performance is a precondition for termination.\footnote{Professor Han Shiyuan holds the view that based on systematic interpretation, the suspension of performance and sufficient assurance should also be preconditions for the termination of a contract due to anticipatory breach. \textit{See} Han Shiyuan, \textit{The CISG and Modernisation of Chinese Contract Law}, 18 COMP. L.J. THE PAC. 75 (2014) (contributions to the Study of International Trade)}

However, a different result holds if the non-breaching party brings a suit against the other party based on the concept of anticipatory breach. Articles 94 and 108 provide that as long as the non-breaching party has evidence to prove that the other party has indicated by words or acts his intention to not perform his obligations, the non-breaching party may immediately terminate the contract and hold the other party liable for the breach.\footnote{See \textit{Xiamen Juying Refrigeration Entm’t Co. v. Xiamen Colorful Era Entm’t Mgmt. Co. Siming District People’s Court}, (2013) SMCZ No. 1963, http://gdlawyer.chinalawinfo.com/newlaw2002/slc/SLC.asp?Db=fnl&Gid=119560631. \textit{See also} Guiding Opinions of the Supreme People’s Court on Several Issues Concerning the Trial of Cases of Disputes over Civil and Commercial Contracts Under the Current Situation (promulgated by Sup. People’s Ct. July 7, 2009, effective July 7, 2009).}

\footnote{See, e.g., \textit{Guiding Opinions of the Supreme People’s Court on Several Issues Concerning the Trial of Cases of Disputes over Civil and Commercial Contracts Under the Current Situation} (promulgated by Sup. People’s Ct. July 7, 2009, effective July 7, 2009).}

\footnote{See \textit{Han & Xiao, supra} note 205, at 38-43; Foreign Civil Law, \textit{supra}, note 16, at 7-13; \textit{Li & Huang, supra} note 205, at 54-57; Ye Jinqiang, \textit{Anticipatory Breach in Chinese Contract Law}, A J. NANJING U. 52-59 (2002).}


7. Reforming the CCL Anticipatory Breach—Defense of Security Regime

In anticipation of the passage of the Chinese Civil Code, Articles 68, 69, 94, and 108 should be amended. The Draft CCC,\textsuperscript{225} in Article 915, retains the substance of the defense of insecurity currently found in CCL Articles 68 and 69. But, in Article 921, the Draft CCC amends the requirements for anticipatory breach in the following manner:

Before the expiration of the contract, if one party has indicated by words that he will not perform his principal obligations, the other party may terminate the contract.

Although Article 915 uses the phrase “principal obligations,” Draft CCC Article 921 fails to define the nature of the breach in determining the appropriateness of anticipating a breach. In particular, Draft CCC Article 921 states: “[i]n case either party indicates expressly by words or by acts that he will not perform the contract, the other party may, before the expiration of the contract, hold the first party liable for breach of contract.”\textsuperscript{226} Therefore, there is still doubt, just as in CCL Article 108, about whether the Draft CCC will continue to unduly enlarge the application of anticipatory breach.\textsuperscript{227} It should also be noted the Draft CCC does not address the right to suspend performance and the requirement to provide notice of suspension.

Before the new Civil Code is adopted, Chinese courts may play an important role in resolving the deficiencies in the CCL. Recently, the court in Xiamen XX Paper Packaging Industry Co. v. Longhai XX Metal Co., Ltd.\textsuperscript{228} held:

The Defendant Longhai XX Metal Co., Ltd. had not performed its obligation of payment in accordance with the sales contract entered into by and between the plaintiff and defendant. The defendant had delayed two installments of payment. When the plaintiff required the defendant to pay the whole contract price, the defendant refused on the ground[s] of lack of money. Since the defendant failed to provide any evidence that he would pay the price when it would be due, the defendant indicates that he would not perform his obligations in accordance with the contract. The defendant’s act has obviously constituted an anticipatory breach under the CCL Article 108.\textsuperscript{229}

The logic behind this civil judgment is that if the defendant could provide sufficient assurance of payment when it is to become due, it is not an

\textsuperscript{225} Liang, PROPOSITIONAL CCC, supra note 133.

\textsuperscript{226} \textit{Id.} at 175, 204 (explaining the Draft Civil Code tries to harmonize the two systems of Defense of Insecurity and Anticipatory Breach, instead of simply deleting one and adopting the other).

\textsuperscript{227} \textit{Id.} (indicating that just like CCL Article 108, Draft CC Article 921 continues to provide that, as long as one party indicates his intention not to perform his “obligations,” no matter how trivial the breach, the other party is entitled to hold the first party liable for the anticipatory breach).


\textsuperscript{229} \textit{Id.}
anticipatory breach. Therefore, if the defendant had provided sufficient evidence that he could pay the price, this would not be an anticipatory breach. This closes a loophole in Article 108 by confirming not all possible future breaches constitute an anticipatory breach. However, when the first party demands adequate assurance of performance, the CCL continues to place a very high burden on the “breaching” party’s ability to obtain a retraction of the anticipatory breach since it must demonstrate a restored capacity to perform and provide an “adequate” guarantee of performance. If the other party fails to provide such assurance within a reasonable time, the first party may hold it liable for anticipatory breach.

Fortunately, courts have stepped in to try to prevent the abuse of the right to declare an anticipatory breach. First, in Shanxi Xinlei Commercial Concrete Co. v. Guangxia Construction Group Co. (2013), the appellate court held, consistently with this Article’s argument, that slight delays of the first two payments did not constitute an anticipatory breach of future payment obligations. Second, in XX Group Co. v. Shanghai XX Concrete Products Co., the court recognized the non-breaching party (Shanghai XX) may choose to hold the other party (XX Group) in anticipatory breach immediately, or it may ignore the breach and wait until the time for performance passes (actual breach), which the CCL does not expressly acknowledge. This right to choose makes some sense when the object of the contract is what the non-breaching party needs. It becomes even more sensible when the object is unique (not replaceable). Therefore, the non-breaching party may try to save the contract by attempting to persuade the breaching party to retract his refusal to perform.

In A Co. v. B Co., the Shanghai First Intermediate People’s Court affirmed the party’s choice to declare or not declare an anticipatory breach. However, it noted the non-breaching party must not misuse its right by ignoring the breach and waiting for the time for performance. This need for the non-breaching party to “accept” the anticipatory breach is grounded in its duty to mitigate damages.

---

230. See LIMING, STUDY ON CONTRACT LAW, supra note 152, at 510-13.


235. Id.

236. CCL, supra note 116, art. 119 (non-breaching “party shall take proper measures to prevent from the enlargement of losses; if the other party fails to take proper measures so that the
breaching party may also violate the principle of good faith. Therefore, the non-breaching party must accept the anticipatory breach and terminate the contract immediately so as to avoid incurring any further damages, if that is the only reasonable course of action.

The above case law shows promise that Chinese courts will, often creatively, interpret and implement provisions of the CCL in a way that produces fair and efficient outcomes. However, nothing short of statutory amendment will be able to resolve the inconsistencies, gaps, and ambiguities presented by Articles 68, 69, 94, and 108. The reform would need to: (1) define “adequate guarantee;” (2) abolish the dual system of defense of insecurity and anticipatory breach and replace it with a single, uniform set of rules; (3) apply a single set of uniform anticipatory breach rules to make clear when suspension rather than termination is appropriate; (4) prohibit the non-breaching party from ignoring an express repudiation of performance or indications of an implied repudiation, as noted in XX Group Co.238 and A Co. v. B Co.;239 and (5) make clear minor breaches in most circumstances should not be grounds for anticipatory breach, especially when this tactic is used opportunistically by the non-breaching party, as noted in Shanxi Xinlei Commercial Concrete.240

C. Lack of a Right to Cure

The seller’s right to cure refers to the breaching party’s right to cure defects in its performance. Usually, in the case of the seller, this right manifests itself either through the repair or replacement of defective goods.241 The relevant laws only offered the non-breaching buyer the choice of either returning the defective goods to the seller in order to recover the full contract price, or keeping the defective goods and recovering the diminution in value of the defective goods, as compared to conforming goods.242 Thus, the seller had no right to cure the defects unless the buyer agreed to receive substituted goods from the seller.

1. Modern Right to Cure

The dawn of industrial production during the middle of the nineteenth century was accompanied by the mass production and supply of goods. As a consequence, repair or replacement of defective goods by the seller, as an alternative remedy to termination and price reduction, surfaced as an issue that losses are enlarged, it may not claim any compensation as to the enlarged losses”.

237. CCL, supra note 116, art. 6 (“The parties must act in accordance with the principle of good faith, no matter in exercising rights or in performing obligations.”).
238. Supra note 192.
239. Supra note 194.
240. Supra note 191.
242. Id.
continued to challenge the civil law and the law of sales throughout subsequent decades and well into the twentieth century. In Germany, the framers of the Civil Code of 1900 failed to depart from tradition and did not insert a right of repair and replacement into statutory remedies.243 As a consequence, commercial practice developed to recognize cure by sellers as a standard remedy for delivery of non-conforming goods.244

In the United States, Karl Llewellyn245 inserted into the 1952 Draft of the UCC the breaching party’s right to cure.246 Since then, the right to cure has been widely accepted in both civil and common law.247 The CISG also provides for a seller’s right to cure. For example, CISG Articles 34 and 37 allow the seller to cure non-conforming documents or non-conforming goods before the sales contract expires, while Article 48 offers the seller the right to cure non-conforming goods or documents for a period of time after the expiration of the delivery date set forth in the sales contract.248 PICC Article 7.1.4 expands the right to cure beyond the sale of goods to other types of contracts.249 It is generally acknowledged that the allowance of a reasonable opportunity to cure is consistent with the notion of good faith and fair dealing250 and with the desire to maintain contractual relations251 where possible and appropriate.252 A reasonable opportunity to cure is also consistent with the spirit of mitigating loss and minimizing economic waste.253

243. See id.
244. See id.; see also SCHLECHTRIEM & SCHWENZER, supra note 103, at 563 n.2; JAMES J. WHITE & ROBERT S. SUMMERS, UNIFORM COMMERCIAL CODE 320 (4th ed. 1995).
245. “Llewellyn held the view that, at least in transactions between merchants, the seller deserved the privilege to cure delivery of non-conforming goods by means of a second tender,” Wagner, supra note 241, at 2; see also K.N. Llewellyn, On Warranty of Quality, and Society: II, 37 COLUM. L. REV. 341, 388-89 (1937).
247. See, e.g., Danish Sale of Goods Act §49, Købeloven, Act no. 102 (April 4, 1906); Sales of Goods Act (SFS 1990:931) (Køjplagen) (Swed.). Even English law recognizes the seller’s right to cure before or after the time for performance if time is not of the essence of the contract. See DCFR, supra note 139, 835-38. For an explanation of the right to cure in Germany, see Andreas Heldrich & Gebhard M. Rehm, Modernisation of the German Law of Obligations: Harmonization of Civil Law and Common Law in the Recent Reform of the German Civil Code, in COMPARATIVE REMEDIES FOR BREACH OF CONTRACT 129 (Nili Cohen & Ewan McKendrick eds., 2005).
248. See SCHLECHTRIEM & SCHWENZER, supra note 103, at 406-08, 440-45, 562-73.
249. See COMMENTARY ON THE PICC, supra note 70, at 747-52.
250. See BIANCA & BONELL, supra note 135, at 291; see also CCL, supra note 116, arts. 5, 6.
251. Contract maintenance is an underlying principle of the CCL, which can be found in the following articles: 8, 10, 19, 22, 28-31, 36, 40, 45, 47, 49-51, 54, 55, 61, 62, 68, 69, 73, 74, 76, 78.
253. Id.; see also CCL, supra note 116, art. 119 (mitigation of damages).
2. **CCL’s Lack of a Right to Cure**

It is the authors’ experience that in Chinese domestic trade practice, it is widely accepted that the breaching party may cure the non-performance before the expiration of the date for performance. In some industries, it is customary practice to provide the breaching party with the ability to repair or replace any non-performance within a fixed period of time, or within a reasonable time after the expiration of the time of performance.\(^{254}\) Despite the commercial practice of acknowledging the right to cure, along with the civil and common laws’ adoption of such a right, the CCL has neglected to do so.

The best remedy would be to amend the CCL to include such a right or to insert such a right in the proposed Chinese Civil Code. The second best remedy would be for the court to imply such a right into the CCL. The Chinese Supreme Court has issued a judicial interpretation recognizing the right to cure in the area of construction contracts.\(^{255}\) For example, the Court implicitly recognized a contractor’s right to cure, upholding the contract-offering party’s claim for a reduction in construction price in Article 11 “if construction fails to comply with the quality requirements as agreed due to the contractor’s fault, and the contractor refuses to repair, rework or modify.”\(^{256}\) Therefore, based on Article 11, if the contractor agrees or offers to repair, rework or rebuild (namely, to cure the non-conformity), the non-breaching party’s claim for price reduction is not supported. Professor Han has suggested a rationale for doing so by linking the right to cure to a price reduction remedy.\(^{257}\) According to CISG Article 50, it is clear that the seller’s right to cure prevails over the buyer’s right to reduce the price.\(^{258}\) Therefore, Han suggests Chinese courts could recognize the right to cure to limit the need for a price reduction in certain situations, with the outcome being the buyer receives conforming goods, and the seller receives the full contract price.\(^{259}\)

---


256. Id. art. 11.


258. See SCHLECHTRIEM & SCHWENZER, supra note 103, at 599.

259. See Logic Structure, supra note 257, at 25.
Another way to incorporate the right to cure into the CCL is by invoking the principle of good faith. If the breaching party can cure the non-performance without undue delay and without causing the non-breaching party unreasonable inconvenience or uncertainty, it would undermine the principle of good faith to not offer the breaching party the right to cure. For example, this would be the case when the breaching party’s non-performance is not fundamental, and the non-breaching party claims damages instead of specific performance (repair or replacement). The lack of a seller’s right to cure produces an inefficient outcome in the form of wasted expenditures and termination of the contractual relationship.

Including the right to cure into the CCL would also address the current imbalance between buyer and seller rights. Currently, a buyer within the CCL remedial scheme can allow the seller to cure by providing a time extension for performance (nachfrist notice), or it can “force” the seller to cure through a demand for specific performance. But, these rights are exercised purely at the discretion of the non-breaching party. As noted above, the non-breaching party may ignore these alternatives by simply declaring an avoidance (termination) of the contract and demanding damages. The CCL should provide a party the right to cure when such a cure can be effectuated promptly and without undue inconvenience to the other party. The strongest case for such a right is when the non-breaching party does not suffer any damages by allowing the breaching party to cure, whether within the contractual time for performance or by an extension of the time for performance. This outcome would align with core principles of contract law, including those of good faith, mitigation of damages, and preservation of the contractual relationship. It would also deter opportunistic behavior, such as when the buyer uses the existence of minor defects that can be easily cured to terminate the contract with the purpose of taking advantage of market changes. Fortunately, the lack of an express right to cure has not caused many cases of injustice or waste because Chinese contracts customarily stipulate such a right. However, the law does not require such a stipulation, leaving open possibilities whereby the parties may inadvertently forget to include such a provision, or a party with superior bargaining power may decline to give the other party such a right.

260. CCL Article 6 is a foundational principle of the CCL and states that: “The parties shall observe the principle of good faith in exercising their rights and fulfilling their obligations.”
261. CCL, supra note 116, art. 111.
262. CCL, supra note 116, art. 94(3).
263. See CCL, supra note 116, arts. 110-11.
3. Reforming the CCL’s Lack of a Right to Cure

The CCL’s failure to provide a right to cure is an unfortunate gap that does not bode well for the future. The Draft CCC provides a “Right to Cure” in Articles 937 and 938:

Article 937:

(1) In case any party to the contract has failed to perform the contract and if the performance period has not become due or his delayed performance has not constituted a fundamental breach, the breaching party may tender a compliant performance at his own expenses.

(2) In case the breaching party intends to cure the contract in accordance with Para. (1) herein, the following conditions shall also be met: (a) without undue delay, it gives notice indicating the proposed manner and timing of the cure; (b) cure is appropriate in the circumstances; (c) the aggrieved party has no legitimate interest in refusing cure; and (d) cure is effected promptly.

Article 938:

(1) Upon effective notice of cure, rights of the aggrieved party that are inconsistent with the breaching party’s performance are suspended until the time for cure has expired.

(2) The aggrieved party may withhold performance pending cure.

(3) Notwithstanding cure, the aggrieved party retains the right to claim damages for delay as well as for any harm caused or not prevented by the cure.265

This provision is drawn from PICC Article 7.1.4. Unfortunately, the Draft CCC Articles 937 and 938 leave a number of questions unanswered. First, they fail to expressly maintain the superiority of the non-breaching party’s right to terminate the contract in cases of fundamental or substantial breach.266 Second, one glaring example of the need to retain this priority is in installment contracts. Because of the obligations to perform numerous installments pursuant to such a contract, one breach can lead to a pattern of breaches, thus making it easier to abuse the right to cure over time. Third, some have argued that, in order to offer consumers better protection, the right to cure should not be available if the non-breaching party is a consumer.267 If this were the case, the consumer, instead of the merchant, would retain more freedom to choose the remedies, including repair, replacement, damages, or termination. Thus, the law should make explicit that the right to cure does not extend to consumer contracts.

265. Liang, PROPOSITIONAL CCC, supra note 133, at 211-12, 215.

266. See generally Michael Bridge, Avoidance for Fundamental Breach of Contract Under the UN Convention on the International Sale of Goods, 59 INT’L & COMP. L.Q. 911 (2010). But for a different opinion, see PICC Article 7.1.4, which provides clearly that the right to cure shall prevail over the right to terminate. In CISG jurisprudence, this issue remains unsettled, but the prevailing view seems to be that the right to terminate (avoid) the contract prevails over the right to cure. But when judging whether the breach is fundamental, the possibility of cure shall be considered. See CISG-AC Opinion no 5, The buyer’s right to avoid the contract in case of non-conforming goods or documents, 7 May 2005, Badenweiler (Ger.).

267. CESL, supra note 134, at 490-92.
CONCLUSION

This Article uses the CCL to explore the pitfalls of the reception and transplantation of foreign laws and influences into a different legal culture. China is an especially interesting case because its CCL reflects a myriad of influences from the civil law (especially the German BGB), common law, and the American UCC, as well as international legal instruments, such as the CISG and the PICC. In this case, the CCL drafters’ study and application of this multitude of foreign and international laws, although admirable, has resulted in a less than clear and comprehensive contract law. This has caused substantial confusion, hindering the Chinese courts’ abilities to properly apply the rules in a uniform manner.

This Article analyzes some of the gaps and inconsistencies in the CCL to illustrate how such borrowing can lead to a less than consistent and comprehensive contract law. The areas of study here include the CCL’s late acceptance rules, its dual system of anticipatory breach and defense of insecurity, and its lack of a right to cure.

The analysis also considers the problem of “double transplantation.” In the case of China, it adopted the CISG as its (domestic) international sales law, which was the first transplant. China then used the CISG as a source in drafting the CCL, the second transplant. Unfortunately, this second transplantation was only partial, since it only transferred some of the CISG rules to the CCL and used other sources for the rest of the CCL, creating a number of problems. First, partial transplantation results in unnecessary inconsistencies between the CCL and the CISG. Second, taking rules out of the context of the body of rules in which they are initially located increases the uncertainty of their meanings when transplanted into a different body of rules. This has certainly been the case in the interpretation and application of the CCL.

Since the CCL is a product of numerous foreign laws, this Article uses a comparative law methodology to try to understand the meaning of the rules in the CCL. It also uses comparative law sources to offer solutions and avenues of legal reform in order to make the CCL a more rational, consistent, and comprehensive contract law. The primary sources analyzed include the German BGB, American UCC, and the common law as represented by the *Restatement (Second) of Contracts*, as well as the CISG and PICC. This Article also draws from other sources, such as the Dutch Civil Code (BW), Louisiana Civil Code, and the Civil Code of Québec, as well as the PECL and the proposed (but, ultimately rejected) CESL. Finally, this Article uses the interpretive guidelines issued by the Chinese Supreme Court and the Draft Chinese Civil Code in its analysis.

In making recommendations to reform the CCL, this Article draws on both comparative law approaches—the “common core” and the “better rules.”268 The wide array of this comparative analysis ferrets out a great deal of commonalities

268. *Supra* notes 39-41 and accompanying text.
between contract rules across different legal systems. This commonality is a powerful rationale for realigning and reforming the CCL. When comparative analysis uncovers differences in rules across legal systems, this Article recommends the use of “better rules” interpretations in the context of the existing CCL. The rules suggested here capture the reasonable expectations of both parties, encourage the efficient conclusion of contracts, and deter opportunistic behavior.
Recovering Lost Tax Revenue Through Taxation of Transnational Households

Ariel Stevenson*

ABSTRACT

This Article addresses the difficult problem of raising revenue in developing countries with significant outmigration. Migrant-source country governments face a unique policy dilemma because emigration reduces domestic human capital and tax revenue, but simultaneously improves outcomes for migrant workers and their families. Thus, governments must balance contrasting needs to maximize government revenue while protecting the welfare of migrant worker households. I argue that migrant-source countries may find a solution to this dilemma by taxing income remitted by migrant workers to family members remaining in their home countries. If constructed properly, a tax on remittance payments could raise revenue without burdening migrant workers or restricting their freedom to migrate.

In this Article, I push back against common anti-remittance-taxation arguments based on both normative and practical considerations, with a focus on improving and updating the taxation of families separated by national borders. After surveying the tax policy instruments available in remittance-receiving developing countries, I offer a menu of policy designs through which policymakers can leverage these important inflows. Proposed policies range from an ideal case of bilateral cooperation between host and home countries to a third-best regime that seeks to harness remittance gains indirectly via consumption and property taxation.

Abstract .........................................................................................................................100
Introduction ......................................................................................................................101
I. Background on Migration and Remittances ...........................................................103
   A. Emigration and Taxation in Source Countries .................................................103
      1. Emigration and Welfare .............................................................................103
      2. Migration Policy Proposals and the Bhagwati Tax .................................105
   B. Remittances and Welfare in Source Countries .............................................109
      1. Background on Remittance Flows ...............................................................109

DOI: http://dx.doi.org/10.15779/Z384G3Z

* Acting Assistant Professor of Tax Law, New York University School of Law. I am grateful to Michael Graetz and Anne Alstott for their thoughtful insights and support on this piece, and to my clients in the Bet Tzedek Legal Services Tax Clinic, for allowing me to help them with their tax problems. Thank you also to the editors of the Berkeley Journal of International Law, especially Nathan Berry and Christopher Yandel.
INTRODUCTION

This Article addresses the difficult problem of raising revenue in developing countries with significant outmigration. Migrant-source country...
governments face a unique policy dilemma because emigration reduces domestic human capital and tax revenue but simultaneously improves outcomes for migrant workers and their families. Thus, governments must balance contrasting needs to maximize government revenue while protecting the welfare of migrant worker households. Migration scholars have proposed various tax policy solutions to this problem, foremost among them being Jagdish Bhagwati’s tax on professional emigrants, which has spawned much scholarly discussion and given rise to a subset of migration-taxation literature.1 However, the Bhagwati tax and many of its intellectual offshoots fail to adequately resolve this policy tension because they prioritize the goal of maximizing government revenue at the expense of protecting migrants’ wellbeing.2 This Article argues that a fairer and more feasible solution to this dilemma is for migrant-source countries to tax income remitted by migrant workers to family members remaining in their home countries. If constructed properly, a tax on remittance payments could raise revenue without burdening migrant workers or restricting their freedom to migrate.

A chorus of voices from the economic development community repeatedly cautions that remittance transfers should not be taxed, as taxation of the flows would result in unfair double taxation, drive the flows underground, or discourage them altogether.3 As a threshold issue, it is important to realize that remitted income is currently being taxed; it is simply a tax imposed by the host country via labor taxation rather than by the recipient developing country.4 By accepting the current structure as neutral non-taxation, the door is closed on shifting the tax revenue from industrialized nations to low-income migrant-source countries.

2. See discussion infra Part I(A)(ii).
4. This Article specifically addresses remittances sent by documented immigrant workers as well as undocumented immigrants who are paying into tax coffers. Although certainly some undocumented immigrants are working “off the books,” the literature suggests that a substantial number of undocumented immigrants pay taxes. See U.S. GOV’T PRINTING OFFICE, ECONOMIC REPORT OF THE PRESIDENT: TOGETHER WITH THE ANNUAL REPORT OF THE COUNCIL OF ECONOMIC ADVISERS 107 (Feb. 2005) (“More than half of undocumented immigrants are believed to be working ‘on the books,’ so they contribute to the tax rolls, but are ineligible for almost all Federal public assistance programs and most major joint Federal-state programs.”); Virginia Harper-Ho, Note, Noncitizen Voting Rights: The History, the Law and Current Prospects for Change, 18 L. & INEQ. 271, 295–96 (2000) (noting that many undocumented immigrants go out of their way to file taxes, even without receiving the refund to which they are entitled); Unauthorized Immigrants Pay Taxes, Too, IMMIGRATION POLICY CTR. (Apr. 18, 2011), http://www.immigrationpolicy.org/just-facts/unauthorized-immigrants-pay-taxes-too (finding that households headed by undocumented immigrants contributed $11.2 billion in taxes in 2010).
This Article pushes back against the anti-remittance-taxation stance based on both normative and practical considerations. While agreeing with the conventional wisdom that the transfers themselves should not be taxed, this Article argues that migrant-source governments should consider taxing recipient households on remitted income sent by relatives working in high-income countries. In order to ensure that the flows are not overtaxed, the ideal policy corollary to this home country tax is non-taxation of remittance payments in the high-income host country. This Article offers several justifications for this shift of tax locus from high- to low-income countries, based on first principles of intra-family transfer taxation, considerations of inter-nation equity, and practical tax goal concerns of home country governments. This Article also proposes several policy structures, taking into account tax administrative realities on the ground in migrant-source developing countries.

Part I provides an overview of current research on migration and remittances and of the role of both phenomena in economic development. Looking to past scholarship on emigration costs, this Article pays particular attention to the Bhagwati tax and the scholarship growing from it, as these provide an intellectual launching pad for the proposals presented later in the Article. Part II addresses theoretical and practical justifications for shifting the locus of remittance taxation from high-income host countries to low-income home countries. Part III describes tax realities in developing countries, offering a survey of available tax instruments as well as specific tax data for twenty remittance-receiving countries. Part IV proposes several different mechanisms through which policymakers might leverage these important inflows, starting with the ideal case of bilateral cooperation between host and home countries, and working down to a third-best regime that seeks to harness remittance gains without the use of income taxation. Finally, Part V addresses administrative concerns including behavioral distortions, political resistance, evasion, and corruption.

I. BACKGROUND ON MIGRATION AND REMITTANCES

A. Emigration and Taxation in Source Countries

1. Emigration and Welfare

Emigration is a double-edged sword, rightfully touted as a great boon to migrant households and industrialized economies that benefit from lower-cost workers, while at the same time draining human capital and tax revenue from migrant-source economies.\(^5\) Migrant-source developing countries seeking to

---

raise revenue find themselves saddled with the difficult task of siphoning water from a leaking bucket as their populace exits in search of better opportunities outside of their borders. Developing creative tax policies is crucial to implementing a lasting solution.

Although the extent of the loss is difficult to calculate, net emigration undoubtedly reduces a country's tax base. Mihir Desai et al. calculate that India loses twelve percent of its income tax base—approximately $700 million—due to high-skilled migration to the United States through the H-1B visa program. Note that this significant figure captures only one kind of migrant to one specific destination country. The authors estimate that only about $300 million makes its way back to India via remittances, leading to a substantial overall loss of government revenue. Additionally, measures of net remittance gains are likely to be overstated, since they fail to account for the local incomes and other household inputs that migrants would have produced domestically had they remained in the home country.

In addition to lost tax revenue, migrant-source countries also lose human capital in the form of their most ambitious and employable citizens who exit in search of greater opportunity. This loss of human capital plagues countries across the developing world and is known as the "brain drain." The migration scholarship documents the negative effect of emigration on human capital development and growth, notably through the work of Jagdish Bhagwati and Koichi Hamada, and more recently by Nadeem Haque and Se-Jik Kim, as well as Kaz Miyagiwa. Haque and Kim go so far as to argue that developing countries should not subsidize higher education, as those citizens who receive advanced education are more likely to migrate to higher-income countries.
However, despite the negative impact that emigration may have on government revenue and human capital stocks, individuals and households do benefit from the increased employment opportunities found by migrant workers abroad. Confusing the matter further, a body of literature argues that emigration can help the domestic economy by putting upward pressure on wages. Thus, while migrant-source countries often lose out from increased emigration, this same migration is a boon to the migrant worker and his household. Such conflicting outcomes create an inherent policy tension for developing country governments that strive to ensure the continued growth and stability of the domestic economy without harming individual citizens.

2. Migration Policy Proposals and the Bhagwati Tax

Before explaining how taxation of remittance flows can balance this inherent policy tension, it is useful to explore other policy solutions that have been proposed to ameliorate the negative effects of emigration. Devesh Kapur and John McHale classify the array of policy options into four broad categories: control, creation, connection, and compensation. Policies in the first two categories aim to directly reduce migration, while those in the third seek to increase remittance payments and spur return migration. This Article is concerned with the final category, that of compensation policies. These policies aim to compensate developing countries and families left back in home countries for the economic losses caused by emigration. Compensation policies can take many different forms, from taxing migration directly via an exit tax, to taxing migrant workers on worldwide income—as the United States does—to sharing tax revenue from rich to poor countries.


18. Id. at 5–6.

19. Id. at 5.

20. Id.
Various economists and migration policy scholars have proposed directly taxing migrant workers as a way to compensate home countries for the losses associated with emigration. Perhaps the most famous tax policy proposal was that suggested by Jagdish Bhagwati in 1972, which has come to be known as the “Bhagwati tax.” This tax would be imposed specifically on high-skilled migrants from low-income countries, with the dual goals of compensating home countries and decreasing human capital flight. Professional emigrants would bear this tax liability for a limited period after emigration, perhaps ten years. Originally, Bhagwati envisioned the tax being collected by host country tax authorities and administered under the auspices of the United Nations. The United Nations would then distribute the tax revenue according to standard development criteria. Many criticized this approach, arguing that U.S. tax collection of migrant income would violate constitutional principles of equal taxation. In response, Bhagwati altered the proposal such that developing country governments would collect the tax directly. Under this structure, developing countries would tax emigrants under the same rationale underlying the U.S. global tax system, whereby citizens are taxed on income earned both at home and abroad. The Bhagwati tax is noteworthy for a number of reasons, not least among them being its recognition of tax policy as a tool for influencing migration outcomes.

Putting aside the preeminent role that the Bhagwati tax proposal plays in moving forward migration policy scholarship, the plan is vulnerable to several criticisms. First and foremost, by increasing the costs of migration the tax imposes a direct restriction on citizens’ international mobility, which reduces personal freedom and economic efficiency. Economic efficiency aside, the


22. Bhagwati & Dallalfar, supra note 21, at 3.
23. Id. at 7.
24. Id. at 6.
25. For example, the United Nations could withhold funds from corrupt or dictatorial regimes.
Id.
27. Wilson, supra note 21, at 2; McHale, supra note 26, at 363–65.
29. Michael A. Clemens, Economics and Emigration: Trillion-Dollar Bills on the Sidewalk?, 25 J. ECON. PERSPECTIVES 83, 83–84 (2011) (detailing the significant economic gains to be had from reducing barriers to emigration); see, e.g., Speranta Dumitru, Skilled Migration: Who Should
right to migrate is itself valuable. Sperantru Dumitru argues that the right to emigrate is a fundamental human right and that restrictions on this right, such as those imposed by the theoretical Bhagwati tax, conflict with important social justice goals.\(^{30}\) The United Nations has codified the right to emigrate in Article 13 of the Universal Declaration of Human Rights, which states that, “[e]veryone has the right to leave any country, including his own.”\(^{31}\) Thus, the Bhagwati tax conflicts with key human rights and social justice goals by reducing individuals’ international mobility.

Additionally, a tax on only high-skilled migration misses potential productivity gains that accrue to unskilled migrants when they move to higher-income countries. Although these unskilled individuals would have perhaps contributed relatively less to source country economies had they remained back home, migration can cause significant productivity gains simply from changing locations.\(^{32}\) A tax based on pre-migration skill level, as the Bhagwati tax is, will fail to capture these productivity gains. While perhaps effective as a means of restricting mobility, this limited policy will therefore be relatively less effective at compensating the home countries for emigration losses, compared with a policy that seeks to harness gains accruing to all migrants.

Desai et al. recently revisited compensatory emigration policies, agreeing with Bhagwati that taxation schemes are an effective way to reduce the negative effects of the brain drain in developing countries.\(^{33}\) The authors propose three possible tax structures that aim to correct various deficiencies of the Bhagwati tax proposal.\(^{34}\) These structures include: 1) taxing emigrants on global income, as the United States does; 2) a cooperative regime in which host countries remit a portion of taxes back to developing source countries; and 3) an exit tax on skilled emigrants.\(^{35}\) Although a step in the right direction, these policies continue to raise administrative, efficiency, and human rights concerns. The second policy is the most promising, but it lacks sufficient detail to make it a workable policy solution for developing countries.

The first policy, taxing global income, involves significant administrative difficulties. It would entail not only maintaining a functional domestic income tax system, but also requires the ability to assess income earned by migrant workers residing in foreign countries. This demands substantial tax collection and enforcement infrastructure, and is likely beyond the capacity of most

---


\(^{32}\) Clemens, *supra* note 29, at 96 (synthesizing current research to conclude that productivity gains associated with migration are based significantly on location effects, rather than innate personal characteristics or self-selection biases).

\(^{33}\) Desai et al., *supra* note 1, at 664.

\(^{34}\) *Id.* at 682–85.

\(^{35}\) *Id.*
developing countries. Further, a global income tax places the migrant taxpayer within the home country’s tax brackets, which assesses his income against a wholly different distribution of wealth compared with the country in which he actually lives. While a migrant taxpayer who earns $15,000 per year in the United States might be subject to a relatively low U.S. tax rate, this might place him in a comparatively high tax bracket in his home country. This raises concerns about deprivation and fairness. Finally, these complications might induce emigrants to renounce their home country citizenship in order to avoid such taxation, reducing the likelihood of return migration or beneficial investment in the home country economy. Thus, although worldwide taxation of citizens is feasible for the United States, administrative challenges and other concerns make it an unrealistic solution for most developing countries.

The third policy option, the exit tax, implicates liberty and efficiency concerns much like the Bhagwati tax. Similar to emigrant taxation, exit taxes reduce overall economic efficiency by reducing voluntary labor exchange. According to Trebilcock and Sudak, shifting from voluntary migration to coercive or planned migration may produce efficiency losses similar to those associated with trade restrictions, thereby reducing overall welfare. Desai et al. propose several modifications to reduce these inefficient distortions, for example, structuring the exit tax as a forgivable education loan, only to be repaid in the event of emigration. Trebilcock and Sudak point out that these proposals do not ameliorate the risk that emigrants will renounce home country citizenship or reduce remittances or investment in home countries.


37. Desai et al., supra note 1, at 682–85.


39. Clemens, supra note 29, at 83–84; Dumitrăș, supra note 29, at 14, 16.

40. Clemens, supra note 29, at 83–84; Dumitrăș, supra note 29, at 14.


42. Desai et al., supra note 1, at 685.

43. Trebilcock & Sudak, supra note 41, at 263.
The second option—in which host countries remit a portion of tax revenue back to developing home countries—remedies many of the complications associated with the first and third policies. Although negotiating a web of bilateral treaties may seem difficult, Desai et al. explain that several trends point to the feasibility of such bilateral cooperation.\textsuperscript{44} Namely, there is increasing need in the industrialized world for developing country labor as well as an increased reluctance on the part of developing countries to relinquish such labor without compensation.\textsuperscript{45} Such pressures might tip the scales in favor of sharing tax revenues from high- to low-income countries. However, this policy option is too imprecise to provide a workable solution, as it fails to adequately address how the tax revenues should be properly divided. Trebilcock and Sudak, working under the assumption that such tax revenue would be divided according to income tax rates in each country, point out that such a policy would likely result in little revenue unless the migrant-source country imposes very high income tax rates.\textsuperscript{46} Thus the underlying basis for the division is integral to the workability and efficacy of such a policy. Ideally such a division would be based on some observable characteristic and firmly justified by basic tax principles, as this will prove more defensible and stable in the long run. Basing the division on remittance payments, as this Article proposes, offers one way to solve this deficiency.

\textbf{B. Remittances and Welfare in Source Countries}

1. Background on Remittance Flows\textsuperscript{47}

Understanding the nature and influence of remittance inflows in developing countries is a prerequisite for understanding their potential role as a tax handle for resource-poor migrant-source countries. There has been a dramatic upsurge in remittance flows over the past few decades. At three times the size of official development aid, they have reached a monumental level.\textsuperscript{48} For most of the past dozen years they have exceeded private debt and portfolio equity inflows, as

\begin{itemize}
  \item \textsuperscript{44}Desai et al., \textit{supra} note 1, at 684.
  \item \textsuperscript{45}Id.
  \item \textsuperscript{46}Trebilcock & Sudak, \textit{supra} note 41, at 263.
  \item \textsuperscript{47}This Article focuses on the specific category of remittance transfers known as worker remittances or personal transfers—typically small, regular financial transfers from migrants living and working abroad sent to support relatives living back home. IMF, \textit{INTERNATIONAL TRANSACTIONS IN REMITTANCES: GUIDE FOR COMPILERS AND USERS} 20–21 (2009), http://www.imf.org/external/np/sta/bop/2008/rcg/pdf/guide.pdf. Other remittance categories not covered here include cross-border employee compensation and remittance of public benefits accrued abroad. Id. at 19–20.
\end{itemize}
well. Flows into developing countries are estimated to have reached $414 billion in 2013, which is a 6.3% increase from 2012. The amounts exceed foreign exchange reserves in at least fourteen developing countries. These figures show that remittances are not merely a stopgap response to a temporary situation, but rather an increasingly significant economic phenomenon that developing countries must learn to harness.

As would be expected, the top remittance recipients measured by aggregate amount are large economies for which remittance inflows comprise merely one source of economic vitality. As reflected in Table 1, in 2013, the World Bank identified India, China, the Philippines, Mexico, and Nigeria as the top recipients in terms of absolute amounts. For smaller economies with a disproportionate share of migrant households, these remittance inflows become a dominating capital source. For example, according to World Bank estimates shown in Table 2, in 2013 remittances comprised 48% of the GDP of Tajikistan, 31% of the Kyrgyz Republic GDP, and 25% each of Nepal and Lesotho’s economies. As these lists demonstrate, remittance-receiving countries vary drastically according to demographic, economic, political, and geographic classifications.

Table 1: Top Ten Recipients by Amount in 2013

<table>
<thead>
<tr>
<th>Country</th>
<th>Amount Received</th>
</tr>
</thead>
</table>


51. Id. (listing countries in which remittance inflows exceed foreign exchange reserves, including Tajikistan, Ecuador, Sudan, and Egypt, among others).

52. Id. at 5.

53. Id.

54. Id.

55 Id.
RECOVERING LOST TAX REVENUE

<table>
<thead>
<tr>
<th></th>
<th>Estimate (in billion)</th>
</tr>
</thead>
<tbody>
<tr>
<td>India</td>
<td>$71</td>
</tr>
<tr>
<td>China</td>
<td>$60</td>
</tr>
<tr>
<td>Philippines</td>
<td>$26</td>
</tr>
<tr>
<td>Mexico</td>
<td>$22</td>
</tr>
<tr>
<td>Nigeria</td>
<td>$21</td>
</tr>
<tr>
<td>Egypt</td>
<td>$20</td>
</tr>
<tr>
<td>Bangladesh</td>
<td>$15</td>
</tr>
<tr>
<td>Pakistan</td>
<td>$15</td>
</tr>
<tr>
<td>Vietnam</td>
<td>$11</td>
</tr>
<tr>
<td>Ukraine</td>
<td>$9</td>
</tr>
</tbody>
</table>

Table 2: Top Ten Recipients as Percentage of GDP

<table>
<thead>
<tr>
<th>Country</th>
<th>Percentage of GDP (based on estimate)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tajikistan</td>
<td>48%</td>
</tr>
<tr>
<td>Kyrgyz Republic</td>
<td>31%</td>
</tr>
<tr>
<td>Nepal</td>
<td>25%</td>
</tr>
<tr>
<td>Lesotho</td>
<td>25%</td>
</tr>
<tr>
<td>Moldova</td>
<td>24%</td>
</tr>
<tr>
<td>Armenia</td>
<td>21%</td>
</tr>
<tr>
<td>Haiti</td>
<td>21%</td>
</tr>
<tr>
<td>Samoa</td>
<td>21%</td>
</tr>
<tr>
<td>Liberia</td>
<td>20%</td>
</tr>
<tr>
<td>Lebanon</td>
<td>17%</td>
</tr>
</tbody>
</table>

2. Remittances and Welfare

Many in the economic development community present remittances as a panacea to the negative effects of migration because remittances inject money into economies that have lost human capital. However, it is questionable whether remittances truly offset the negative effects of labor migration out of source countries. Although they are likely a boon to the households that receive them, economists have been unable to establish a robust link between remittances and aggregate economic growth in low-income countries.

---

56. Id.
58. Trebilcock & Sudak, supra note 41, at 259 (“Whether remittance payments on their own are a net benefit is a question less complex than whether remittance payments might sufficiently spur development to offset the human capital concerns raised above. That question is largely an open one.”).
Understanding the limitations of remittances in spurring economic growth demonstrates why compensatory tax policies remain necessary in migrant-source countries, despite the large and often increasing size of remittance inflows.

At the household level, data supports the intuition that remittance payments reduce deprivation by providing resources for basic household needs. The corollary to this is that remittances are not typically invested or saved in a way that supports long-term growth. According to research collected by the International Fund for Agricultural Development, eighty percent to ninety percent of remitted income is used for basic consumption, healthcare, and education.60 The remaining ten percent to twenty percent is saved or invested in either formal or informal financial instruments.61 Research conducted among Latin American immigrants living in Connecticut reveals similar findings, with over half of respondents listing food as an important use of remittances.62 Over a third of respondents in the Connecticut study also listed home maintenance as an important expense, suggesting that the remitted funds need not be used for only the most basic needs.63 Less than ten percent of respondents reported that remitted funds were used to finance investments in home countries.64 Thus, across different populations, research tends to agree that remitted funds are used to boost consumption rather than increase long-term savings and investments.

Lack of savings aside, a good deal of evidence supports the assumption that remittances benefit the individuals and communities that receive them in myriad ways, including by alleviating poverty,65 contributing to human capital development,66 and increasing investment in microenterprise,67 among other

---

60. INTERNATIONAL FUND FOR AGRICULTURAL DEVELOPMENT, SENDING MONEY HOME: WORLDWIDE REMITTANCE FLOWS TO DEVELOPING AND TRANSITION COUNTRIES 7 (2007), http://www.ifad.org/remittances/maps/brochure.pdf [hereinafter IFAD, SENDING MONEY HOME].
61. Id.
63. Id.
64. Id. at 32.
65. See Pablo Acosta et al., Do Remittances Lower Poverty Levels in Latin America?, in REMITTANCES AND DEVELOPMENT: LESSONS FROM LATIN AMERICA 87, 128 (Pablo Fajnzylber & J. Humberto López eds., 2008) (the authors find that remittances are positively correlated with poverty reduction as well as economic growth, with data showing a 0.4% decrease in the fraction of the population living in poverty for every one percent increase in the share of remittances to GDP); see also Pia M. Orrenius et al., Do Remittances Boost Economic Development? Evidence from Mexican States, 16 L. & BUS. REV. AMERICAS 803 (2010) (finding that remittances lessen certain measures of income inequality).
things. A World Bank study examining data from 115 developing countries found that international remittances reduced both the level and depth\textsuperscript{68} of poverty.\textsuperscript{69} The study also found that remittance-receiving households spent less on immediate consumption goods for each additional dollar received and more on investments such as education, housing, and entrepreneurial activities.\textsuperscript{70} Research from Guatemala shows that remittances have played a large role in reducing the depth of poverty, meaning that they are particularly helpful to the poorest of the poor.\textsuperscript{71} A study in El Salvador found that remittances reduce the probability of children leaving school, even when compared with other sources of income.\textsuperscript{72} Additionally, although investment might not comprise their primary use, remittances have been found to support urban microenterprises in Mexico\textsuperscript{73} and to ease credit constraints for new businesses in the Philippines.\textsuperscript{74}

From poverty alleviation, to educational attainment, to business investments, remittance transfers ease constraints and provide beneficial support for countless migrant households in the developing world.

Despite the notable positives listed here, researchers have consistently been unable to robustly link international remittances with long-term economic

\textsuperscript{67.} See Christopher Woodruff & Rene Zenteno, Remittances and Microenterprises in Mexico 3–4 (Aug. 14, 2001) (unpublished manuscript) (on file with author), http://dx.doi.org/10.2139/ssrn.282019 (finding that remittances play an important role in funding microenterprises, estimating that nearly a third of funds invested in Mexican microenterprises come from remittances).

\textsuperscript{68.} The level of poverty refers to the percentage of the population living below the poverty line, while the depth of poverty refers to the amount by which the average income of the poor falls below the poverty line.


\textsuperscript{70.} Id. at 26 (measuring marginal budget shares on expenditures in remittance receiving and non-remittance receiving households).

\textsuperscript{71.} Richard H. Adams, Jr., \textit{Remittances and Poverty in Guatemala} 12–13 (WBG, Policy Research Working Paper No. 3418, 2004); but see Alejandro de la Fuente, \textit{Remittances and Vulnerability to Poverty in Rural Mexico}, 38 \textit{World Dev.} 828, 838 (finding that remittances tend to benefit households that are already better off financially compared with non-recipients).

\textsuperscript{72.} Alejandra Cox Edwards & Manueltita Ureta, \textit{International Migration, Remittances, and Schooling: Evidence from El Salvador}, 72 \textit{J. Dev. Econ.} 429, 450 (2003); see also Gordon H. Hanson & Christopher Woodruff, Emigration and Educational Attainment in Mexico 16 (Apr. 2003) (on file with author) (finding that Mexican children in transnational households completed significantly more schooling, with the largest impact being on girls with mothers with low levels of education).


growth in recipient countries.\footnote{The economic literature points out several growth-hampering negative consequences of remittances, including reduced labor market participation by recipient family members, exchange rate appreciation and Dutch Disease, and volatility due to the uncertainty of remittance flows.}

Further, as explained above, remittances are more likely to spur consumption rather than productive investment, limiting an important potential avenue of long-term growth.\footnote{Although it may seem that remittance-receiving households should save more than non-receiving households, Pablo Acosta et al. find that remittance inflows are actually correlated with lower savings rates among remittance-receiving households in certain Latin American countries and for high-income remittance recipients in general. The authors also find that remittances may harm growth by reducing the quality of government institutions; Pablo A. Acosta, Emmanuel K.K. Larrey & Federico S. Mandelman, Remittances and the Dutch Disease (Federal Reserve Bank of Atlanta, Working Paper No. 2007-8, 2007) (finding that remittance flows, whether altruistically motivated or not, lead to a reduction in the labor supply and exchange rate appreciation that hampers growth); Pablo Acosta, Pablo Fajnzylber & J. Humberto López, Remittances and Household Behavior: Evidence for Latin America, in REMITTANCES AND DEVELOPMENT: LESSONS FROM LATIN AMERICA 133, 161 (Pablo Fajnzylber & J. Humberto López eds., 2008) (using data from all ten countries for which data is available, the authors find that receiving remittances reduces the number of hours worked by per week by recipient family members); Dutch Disease, named for a notable instance of the phenomenon in the Netherlands in the 1960s, occurs where increased foreign currency inflows lead to exchange rate appreciation that in turn reduces exports, hinders import competition, and slows growth. See, e.g., Christine Ebrahimzadeh, Dutch Disease: Too Much Wealth Managed Unwisely, 40 FIN. & DEV. 1 (2003), http://www.imf.org/external/pubs/ft/fandd/2003/03/ebra.htm; Joong Shik Kang, Alessandro Prati & Alessandro Rebucci, Aid, Exports, and Growth: A Time-Series Perspective on the Dutch Disease Hypothesis 3 (IMF Research Dep’t, Working Paper No. WP/13/73, 2013) (exploring links between aid and Dutch Disease). For a discussion of the link between remittances and Dutch Disease, see J. Humberto López, Luis Molina & Maurizio Bussolo, Remittances, the Real Exchange Rate, and the Dutch Disease Phenomenon, in REMITTANCES AND DEVELOPMENT: LESSONS FROM LATIN AMERICA 217, 232–34 (Pablo Fajnzylber & J. Humberto López eds., 2008).}
recipient family members work fewer hours per week, confirming other evidence that remittance payments reduce labor market participation in home countries.81

These negative effects may be especially pronounced in smaller countries. Several studies highlight that smaller economies with high remittance-to-GDP ratios face special challenges in harnessing remittance inflows for economic growth. J. Humberto López et al. explore fears that remittance inflows may lead to exchange rate appreciation and Dutch Disease, especially where inflows are too large relative to the domestic economy.82 Their research shows that indeed remittance inflows are associated with exchange rate appreciation in Latin America.83 They further find that these results are broadly applicable across regions, suggesting that any remittance-receiving countries may be at risk of Dutch Disease, particularly where remittance inflows are large relative to GDP.84 In addition to exchange rate appreciation, smaller economies are more sensitive to the inherent volatility of remittance flows. Katsushi Imai et al. find that remittance volatility can have a negative effect on economic performance where recipient countries are unable to protect themselves against sudden swings in flows.85

It would be folly to argue that remittances are wholly detrimental at the household, community, or even country level. Rather, the literature demonstrates that remittance inflows can have varying effects on different kinds of households and in different domestic settings—many positive and some negative.86 Further, researchers are repeatedly unable to establish a robust link between remittance inflows and sustained economic growth, and some even find negative consequences of the flows in the aggregate.87 Thus, remittances alone are not the solution to persistent poverty in migrant-source countries. Remittance-receiving countries must do more to ensure that remittance inflows contribute to improved long-term growth and aggregate welfare.

3. The Argument Against Remittance Taxation

Despite the massive and increasing size of remittance flows, and their uncertain effect on growth, experts tend to agree that developing countries should not seek to ameliorate emigration costs through remittance taxation.88

81. Id. at 158–66, 161.
82. López et al., supra note 77.
83. López et al., supra note 77, at 232–34.
84. Id.
86. Abdih et al., supra note 75, at 664, and accompanying text; Acosta et al., Remittances and the Dutch Disease, supra note 75, at 1, and accompanying text; see generally Barajas et al., supra note 59 and accompanying text; Chami et al., supra note 75, at 1, and accompanying text.
87. Id.
88. See, e.g., Dilip Ratha, Leveraging Remittances for Development, MIGRATION, TRADE, &
According to the most common arguments, taxing remittances would: 1) result in unfair double taxation; 2) reduce the incentive to send remittances, thereby reducing total flows; 3) drive the transfers into informal channels, which reduces the use and attendant benefits of formal financial institutions; and 4) entail significant administrative challenges. Although not without merit, these concerns are insufficient to justify a blanket anti-remittance taxation stance. Importantly, these arguments fail to recognize that remitted income is currently being taxed—it is simply a tax imposed by host countries via labor taxation rather than the recipient developing country. By accepting the current structure as neutral non-taxation, the door is closed on shifting the tax revenue from industrialized nations to low-income migrant-source countries. Further, much of the concerns, such as fears of double taxation and high administrative costs, can be addressed through targeted tax policy design.

C. Migration Taxation in Practice

Not blind to the costs of emigration, developing country governments have attempted in various ways to compensate domestic economies for shrinking human capital stocks and associated dwindling tax revenues. Although most countries avoid taxing remittances, those countries that do tax remittances largely do so indirectly by requiring recipients to convert the payments to overvalued local currency at uncompetitive official exchange rates. Ethiopia, Pakistan, Venezuela, and Cuba are examples of countries that have employed such policies. Hidden taxes such as this preclude effective public oversight, thereby engendering corruption. Cuba previously imposed this method of taxation by requiring all remittances sent from the United States to Cuba be paid to recipients in Cuban Convertible Pesos, then levying a ten percent conversion tax. The United States originally stymied this tax policy by allowing remittances to Cuba to be paid directly in the Cuban currency, enabling remitters to

---

Dev. 173, 180 (James F. Hollifield, Pia M. Orrenius & Thomas Osang eds., 2006), http://www.dallasfed.org/assets/documents/research/pubs/migration/migration.pdf#page=171 (arguing against remittance taxation in part because remittances are private transfers that “should not be expected to fund public projects”); Sanket Mohapatra, Blanca Moreno-Dodson & Dilip Ratha, Migration, Taxation, and Inequality, Econ. Premise, May 2012, at 2 (noting that countries typically resist the temptation to tax remittances, which would drive them into informal channels).

89. Mohapatra, supra note 3.

90. See discussion infra Part IV.

91. Mohapatra, supra note 3.

92. Id.

93. See Susan Eckstein, Remittances and Their Unintended Consequences in Cuba, 38 World Dev. 1047, 1051–52 (describing the Cuban government’s remittance appropriation practices, which undermined the state’s moral authority and contributed to corruption and state profiteering).

to avoid the conversion tax\textsuperscript{95}—although these efforts will be moot once Cuba realizes its intention to eliminate the Cuban Convertible Peso altogether.\textsuperscript{96} While the Philippines does impose a documentary stamp tax on remittance transfers, in November 2010 the Philippine Bureau of Internal Revenue passed a regulation exempting all overseas foreign workers from the tax.\textsuperscript{97} The regulation covers foreign workers who are registered with the Philippine Overseas Employment Administration and who are able to show valid proof of foreign employment.\textsuperscript{98} Because only Filipino emigrants in the formal economy will benefit from the exemption, informal workers may decide to remit funds back to the Philippines via informal channels or to reduce remittance transfers overall. Thus, while remittance taxation does occur, so far no government has developed an effective and fair remittance tax policy that compensates the domestic economy without distorting transfer behavior or engendering corruption.

Outside of remittance taxation, some migrant-source countries impose direct and indirect burdens on emigrants in an effort to reduce emigration costs. Because these policies seek to reduce individual mobility, they often raise liberty concerns. In the most restrictive cases, certain countries impede emigration through the use of exit visas. For example, the former Soviet Union employed exit visas to prevent emigration, and today Cuba and Nepal still require citizens to obtain a permit to leave.\textsuperscript{99} Other countries, such as Saudi Arabia and Qatar, require foreign workers to obtain exit permits.\textsuperscript{100} These policies raise significant human rights concerns.

\begin{itemize}
 \item[95.] Id.
 \item[96.] Marc Frank, Cuba Likely to End Dual Currency System, FIN. TIMES (June 15, 2015), http://www.ft.com/cms/s/0/b34fd7b8-fb12-11e4-9aed-00144feab7de.html#axzz3pGfJb4Q6.
 \item[97.] Tax Treatment of Income Earnings and Money Remittances of an Overseas Contract Worker (OCW) or Overseas Filipino Worker (OFW), REVENUE REG. 1-2011 § 3(c) (Phil.) (“The remittances of all [overseas Filipino workers], upon showing of [the same proof of entitlement by the overseas Filipino worker’s] beneficiary or recipient, shall be exempt from the payment of documentary stamp tax . . . .”).
 \item[100.] Bradford, supra note 9, at 39–40 n.39; Heather E. Murray, Note, Hope for Reform Springs Eternal: How the Sponsorship System, Domestic Laws and Traditional Customs Fail to Protect Migrant Domestic Workers in GCC Countries, 45 CORNELL INT’L L.J. 461, 471 (2012); Katherine Scully, Note, Blocking Exit, Stopping Voice: How Exclusion from Labor Law Protection Puts Domestic Workers at Risk in Saudi Arabia and Around the World, 41 COLUM. HUM. RTS. L. REV. 825, 852–53 (2010) (explaining that Saudi Arabia’s kafala immigration law requires a domestic worker who seeks to leave prior to the expiration of her contract to obtain an exit visa, which is
Moving away from direct migration restrictions, certain countries have adopted forms of emigrant taxation much like that envisioned by Bhagwati and Desai et al. Eritrea is a primary example, imposing a two percent “voluntary tax” on emigrants’ annual income. The tax applies to all Eritreans living abroad, regardless of income. In exchange for the tax, the government provides emigrants full citizenship and robust political rights. Studies suggest that the vast majority of citizens pay the tax, either out of patriotism or social pressure, as paying the tax is a “public” act. The Eritrean government has also undertaken significant steps to attract remittances from Eritreans living abroad. Government policy enables tax-free remittance transfers in any major global currency and with favorable exchange rates.

In a more targeted attempt to harness gains in migrant workers’ income, South Korea has previously imposed taxes on certain emigrant citizens working abroad under construction contracts that the Korean government negotiated. Under this program, the Korean government assisted domestic companies that used Korean migrant workers to secure projects in the Middle East. The government then withheld income taxes from workers’ income and required that their Korea-based employers deposit a percentage of the salaries into Korean bank accounts. Kim Barry suggests that this type of program unfairly burdens participating migrant workers and conflicts with a progressive taxation structure.

controlled by the visa sponsor, as well as a court order releasing the worker from her contract); HUMAN RIGHTS WATCH, “AS IF I AM NOT HUMAN”: ABUSES AGAINST ASIAN DOMESTIC WORKERS IN SAUDI ARABIA 26–33 (2008), http://www.hrw.org/sites/default/files/reports/saudiarabia0708_1.pdf (describing how kafala sponsorship system gives the visa sponsor near total control over the migrant workers’ entry into and exit from the country).

102. Id.
103. Id.
104. Id. The tax is often called a “healing tax,” and is viewed by some as an “affirmation of citizenship.” Id. at 39. “In the words of one Eritrean in Germany, nonpayment ‘would be declaring that I am not an Eritrean.’” Id.
106. As explained infra, Part IV, this Article does not advocate for taxation of remittance transfers directly. Rather, the Article argues that remitted income should be relieved of taxation in the host country and then subjected to a progressive income tax in the recipient nation, thereby shifting tax revenue from high- to low-income nations. This Article agrees that remittances should free from transfer taxation, as Eritrean policy provides.
107. Id.
109. Id.
110. Id.
since migrant workers in more lucrative industries are not subject to the same obligations.\footnote{Id. at 37–38.}

Encouraging emigrants to send remittances is the least burdensome emigration compensation policy. Many remittance-receiving countries have established programs designed to harness remittances for development purposes by creating specialized development funds or targeted financial instruments known as remittance or diaspora bonds.\footnote{B. Lindsay Lowell \& Rodolfo O. de la Garza, \textit{Inter-American Dialogue, The Developmental Role of Remittances in U.S. Latino Communities and in Latin American Countries} 11–12 (2000), http://www.thedialogue.org/PublicationFiles/Final%20report.pdf.} Remittance-sending nations have started capitalizing on remittances as well, often in partnership with recipient countries. For example, the U.S. and Mexican governments have entered into a joint financial program, called \textit{Directo a México}, which aims to assist banks in the United States in remitting customers’ funds to Mexico by promoting the use of an Automated Clearing House channel.\footnote{See Fed. Reserve Banks Servs., \textit{Directo a México: Frequently Asked Questions} (2005), http://www.frbservices.org/files/help/pdf/DirectoMexicoFAQ.pdf.} Both governments stand to profit from these money transfers. Mexico also matches migrants’ contributions made via “hometown associations,” which are organizations that link migrants with home communities and utilize funds for local community development.\footnote{Francisco Javier Aparicio \& Covadonga Meseguer, \textit{Collective Remittances and the State: The 3x1 Program in Mexican Municipalities}, \textit{40 World Dev.} 206 (2012).} Since 2002 this matching-funds program has financed more than 6,000 development projects, taking advantage of an average annual investment of $15 million from the Mexican federal government.\footnote{Id.}

Altogether, the policies discussed above demonstrate the recognized need to compensate developing countries for economic losses due to net emigration. Taxing remittance inflows could provide a partial answer to the problem of lost capital in migrant-source countries. A well-constructed tax on remitted income would build on the longstanding scholarly discussion of migration-compensation policies, as epitomized by the Bhagwati tax and similar proposals. Remittance taxation improves on these earlier proposed policies by preserving migrant workers’ freedom to emigrate and reducing various administrative limitations.

Before describing potential policy regimes in greater detail, the next section explores normative and practical justifications for taxation of remittance inflows by recipient developing countries.

\section*{II. Normative and Practical Justifications for Remittance Taxation}

Through taxation of remittance inflows, migrant-source country governments would gain a robust and visible income stream from which to draw public revenue. It is important to realize that even if a recipient country does not
tax remitted income, the income remains subject to taxation by the host country. Both theoretical and practical considerations support shifting the point of taxation from the host country to the low-income recipient country.

Experts correctly argue against directly taxing remittance transfers, which would reduce transfers overall and drive them underground. Further, a tax on transfers is likely regressive because it is calculated independent of the remitter’s ability to pay. However, by moving one step down the transfer stream and taxing remittance recipients, governments would be able to raise revenue in a progressive way that compensates for emigration without restricting labor mobility or drastically reducing remittance inflows. Under an ideal policy structure, the transferred income would be free of taxation in the host country to enable subsequent taxation in the home country. This contravenes the standard practice in which labor income is taxed where the work is done. Such a policy would also breach current accepted doctrine regarding non-deductibility and non-taxation of gifts and intra-family transfers. Thus, it is first necessary to explain why remittance taxation justifies such a divergence from tax policy custom.

A. Taxation of Intra-Family Transfers

Crafting an ideal tax policy for international remittances entails defining the proper tax treatment of transnational families. Looking to first principles of family taxation, it actually makes good sense to tax recipient family members on remittance inflows received, rather than taxing remittance senders on relinquished labor income. This is because transnational families are not subject to the same concerns that underlie customary non-deductibility and non-taxation of intra-family transfers—namely, income pooling and bracket shifting.

With regard to income pooling, the first question in determining how best to tax transnational families is whether they should be taxed jointly or separately. Under joint taxation of spouses or family units, intra-family
transfers are neither deductible to the transferor nor taxable to the transferee since relevant members are taxed as one unit. Although governments adopt family taxation policies for a variety of reasons, the standard justification for joint taxation of households is that families act as single economic units that pool their income.\(^{122}\) However, income pooling is practically impossible for transnational families because national borders separate their household finances. Consequently, it is more appropriate to tax transnational family members as separate entities in each country of residence.

Furthermore, much of the policy underlying family taxation arises out of a concern over bracket shifting, where a high-tax spouse shifts income-producing property to a low-tax spouse in order to reduce tax liability.\(^{123}\) This same concern partly explains the parallel treatment extended to gifts, which are nondeductible to the giver and tax-free to the recipient.\(^{124}\) Bracket shifting is hardly a concern in the context of transnational remittances. Because many remittance senders are low-income taxpayers in host countries,\(^ {125}\) they have little incentive or ability to engage in tax evasion through bracket shifting. Bracket shifting among transnational families is also unlikely because family members often reside in home countries, making transferring capital across borders much more difficult. This suggests that non-deductibility and non-taxation of intra-family transfers may be less justified in the case of international remittance payments.

Instead of treating transnational households like standard families, it may be more appropriate to treat them like divorced families. Under this framework, remittances are akin to alimony payments, which are taxed as income to the recipient and deductible by the sender.\(^{126}\) The tax treatment of alimony demonstrates the principle that where intra-family payments result in a fully executed transfer—i.e., where there is no risk of bracket shifting or income pooling—then taxing the income to the recipient is more appropriate than continuing to tax it to the earner.\(^{127}\) This is certainly the case for remittance...
transfers, as national boundaries separate the recipients from the sender who is unlikely to benefit from the funds.

One possible counterargument to taxing remittance recipients is that income should be taxed to the person who earns and controls the income, rather than the person who consumes or benefits from the income.\textsuperscript{128} This concept, sometimes referred to as the “control principle,” is a pillar of tax policy design.\textsuperscript{129} However, taxation of remittances to the recipient rather than the sender still satisfies the control principle for a number of reasons. First, unlike single-nation families in which the earner can maintain practical control of his earned funds, the recipients in transnational families obtain full control of the remittance payments due to physical separation between the senders and recipients. This means that an earner relinquishes actual control of the income once he transmits the money abroad. At best he can request how the money be used, but he likely has little recourse or knowledge if the family members back home choose an alternative expense.

The control principle further supports taxation of recipient family members because remittance-receiving households should properly be considered the earners of transferred income rather than the beneficiaries of charitable transfers. This perspective arises out of an evolving understanding in remittance research, which suggests the payments should be viewed as rightful investment income rather than mere gifts. A great deal of empirical research shows that migrant households are more likely to view remittances as an agreed-upon return on an investment made by the entire household, rather than a result of altruism on the part of the sender.\textsuperscript{130} Although there is perhaps no legal requirement that family

\textsuperscript{128} Zelenak, \textit{supra} note 121, at 343 (“In an income tax, control should govern, not consumption.”). Although the control principle establishes that an ideal tax is levied on the controller of income, tax systems instead tax income earners because it is simple to determine who earns income, but much more difficult to determine who controls income. For a discussion of the control principle and taxation of earners as controllers, see Lucas v. Earl, 281 U.S. 111, 114–15 (1930); Zelenak, \textit{supra} note 121, at 343 (“The evidence on marital pooling suggests that, although spouses may share the consumption of resources rather evenly, the control over income remains with the earner.”).

\textsuperscript{129} Lucas, 281 U.S. at 114–15; Zelenak, \textit{supra} note 121, at 343.

\textsuperscript{130} The development community has recently started to understand this and to reframe remittances as a household investment rather than a windfall gain. See, e.g., Michael Clemens & Timothy Ogden, \textit{Migration as a Strategy for Household Finance: A Research Agenda on Remittances, Payments, and Development} (Ctr. for Global Dev., Working Paper No. 354, 2014) (arguing that migration is seen as an investment by the migrant household, and remittances as a return on investment). This view is bolstered by migration research that emphasizes that migration decisions are based on joint decision-making by the entire household, which is a line of discourse termed the “new economics of migration.” For an in-depth explanation of the new economics of migration, see Oded Stark & David E. Bloom, \textit{The New Economics of Labor Migration}, 75 AM. ECON. REV. 173 (1985). See also Douglas S. Massey et al., \textit{Theories of International Migration: A Review and Appraisal}, 19 POPULATION & DEV. REV. 431, 436 (1993) (summarizing recent theories of migration); Oded Stark, Book Review, 14 J. DEV. ECON. 251 (1984) (reviewing \textit{Migration Decision Making} (Gordon F. DeJong & Robert W. Gardner eds., 1981)). For example, research
members remit money back home, as there is for alimony payments, in many cases the senders have a binding moral obligation to make remittance payments. Conceptualizing the payments as a return on investment thereby further suggests that remittances are properly taxed to the recipients as a form of rightful income, as the control principle argues. Thus, in the case of a transnational household, satisfying the control principle may actually entail taxing recipients rather than earners.

Upon considering first principles of family taxation, we see that the concerns that justify the customary treatment of intra-family transfers do not apply to families separated by national borders. From the perspective of the host country, taxing remitted income to recipients rather than senders would achieve a more appropriately tailored transnational family tax policy, and one that is better supported by first principles of family taxation.

B. Inter-Nation Equity

Pulling back from the household to the international level, concepts of residence-based taxation and inter-nation equity provide support for developing countries’ claims to remittance tax revenue. As a threshold matter, it is widely accepted that both the source country and the residence country have a claim to foreign income earned by a resident taxpayer. The “residence principle” states that taxpayers who earn money through labor or investment abroad are considered to owe tax allegiance to their residence countries in return for the rights and privileges that they receive as residents. In this context, “residence”

conducted among cyclical migrants in Albania found that migrants’ decisions to return home were based on risk-pooling among all household members, rather than individual decisions by the migrants themselves. Talip Kilic et al., Investing Back Home: Return Migration and Business Ownership in Albania, 17 ECON. TRANSITION 587, 591 (2009). A case study of migrant households in Santa Ana, Mexico also found that migration resulted from household-level decision making among nuclear or extended families. Dennis Conway & Jeffrey H. Cohen, Consequences of Migration and Remittances for Mexican Transnational Communities, 74 ECON. GEOGRAPHY 26, 37 (1998). The Santa Ana case study revealed that migration decisions are typically based on improving family or household livelihoods. Working together, the family decides which member to send abroad in order to realize sufficient gains to counteract stagnant home economies and market failures. Thus, empirical evidence suggests that remittances are not actually considered gratuitous gifts, but rather entail a joint household investment in the migrant, who then repays this investment through sending remittances home.

131. See Charles Tilly, Trust Networks in Transnational Migration, 22 SOC. FORUM 3 (2007) (exploring various trust relationships that arise in migration, particularly the obligation of sending remittances back home).

132. See, e.g., Nancy H. Kaufman, Fairness and the Taxation of International Income, 29 LAW & POL’Y INT’L BUS. 145, 148 (1998) (“Source, residence, and citizenship taxation provide the framework within which each country legislates its domestic rules for the taxation of international income.”).

133. See, e.g., Peggy B. Musgrave, Sovereignty, Entitlement, and Cooperation in International Taxation, 26 BROOK. J. INT’L L. 1335, 1336 (2001) (explaining that under the “residence principle”, “[r]esidents are held to owe tax allegiance in return for the rights and privileges which they receive as residents”); Kaufman, supra note 132, at 148.

134. See, e.g., Kaufman, supra note 132, at 148; Peggy B. Musgrave, Sovereignty, Entitlement,
refers to the residence country of the recipient household, rather than the host country where the migrant worker lives and works. Given the long-accepted right of both source and residence nations to tax income, the question then becomes how the tax revenue should be divided by both nations. Although host countries currently dominate taxation of remitted income, considerations of inter-nation equity and international distributional concerns justify shifting some portion of remittance tax revenue to home countries.

The concept of inter-nation equity, developed by Peggy and Richard Musgrave, provides a framework to challenge the current source country primacy in the taxation of remitted income. The theory posits that considerations of equity between national units, as opposed to purely inter-individual equity concerns, might provide guidance as to how to distribute tax revenue between home and host countries. Unlike inter-individual equity, wherein countries seek to promote equity among individual taxpayers, inter-nation equity refers specifically to equity at the national level. The concept attempts to address at what point a government is entitled to collect tax revenue, based on the allocation of gains and losses between nations. Imagine the case of an investor from Country B earning income on a project operating within Country A’s borders. Under the inter-nation equity framework, when Country A taxes that foreign investor on the income earned within its borders, Country B’s potential gain decreases by the amount of the tax. The Musgraves’ work, on which the hypothetical is based, specifically addresses corporate taxation, but the general principles are broadly applicable to individual income taxation and taxation of remitted income.
The concept of inter-nation equity is more than just a positivist description of inter-nation distributions of gain loss. It suggests that distributational concerns ought to govern the division of tax entitlements between high-income and low-income nations.\textsuperscript{142} Although largely absent from the scholarly discussion of inter-nation tax equity, which focuses on corporate investment, tax policies generate gains and losses in the context of labor income as well. When a migrant worker from Country B earns income in Country A with the intention of remitting a certain portion back home, a tax on this income in Country A reduces the potential gain to Country B. Recognizing that taxation of remittances in the host country represents a net loss to the home country, high-income host countries should consider ceding certain limited tax authority to low-income recipient countries in the furtherance of international equity concerns.

The actual size of the loss created by the host country’s tax will depend on the elasticity of remittances to taxation, which is largely an empirical question. To see why, imagine a scenario in which the host country, Country A, imposes no income tax. A migrant worker, X, from Country B earns $100 per week while working in Country A and remits $20 of his earnings to family back home each week. Now suppose that Country A implements a twenty percent income tax, thus reducing X’s after-tax income to $80. There are three possible outcomes. First, knowing his family needs the full $20, X does not reduce the remitted amount. Here there has been no loss to Country B, but X bears the cost in that he now must live on only $60 per week. Second, knowing he needs the full $80 to survive in the host country, X cuts out his remittance entirely. In this case, Country A has effectuated a $20—or one hundred percent—loss on Country B.\textsuperscript{143} Third, X splits the difference between himself and his family, reducing remittances by some amount—perhaps twenty percent, to $16. In this case, Country A has effectuated a $4—or twenty percent—loss on Country B. If either the second or third scenarios dominate, which they likely do in many cases, then taxation by the host country results in a redistribution of gain from Country B to Country A. In the case of low-income migrant-source countries, practical application and offering no guidance in the actual distribution of tax entitlements among nations. See, e.g., Avi-Yonah, supra note 138, at 1648–49 (describing Musgrave’s formulation of inter-nation equity as vague and offering no practical guidance in allocating tax entitlements among countries); Michael J. Graetz, Taxing International Income: Inadequate Principles, Outdated Concepts, and Unsatisfactory Policies, 54 TAX L. REV. 261, 284–94 (2001) (explaining that application of Musgrave’s principles would have destroyed nearly all incentives for U.S. taxpayers to invest abroad).

\textsuperscript{142} See, e.g., Avi-Yonah, supra note 138, at 1650 (“More specifically, when a choice is presented between two otherwise comparable alternative rules, one of which has progressive and the other regressive implications for the division of the international tax base between poorer and richer countries, the progressive rule should be explicitly preferred to the regressive one.”).

\textsuperscript{143} Note that this example equates the gain to the recipient household with the gain to Country B. The actual loss to Country B’s treasury would depend on complicated factors such as tax effort and tax efficacy, which become yet more complex in a developing country setting. This simple hypothesis assumes that Country B seeks to maximize gain to households, ignoring country-level gain for the moment.
the result is redistribution from a low-income country to a high-income country, worsening preexisting global inequity.

This discussion presumes that international equity is indeed a concern for high-income host countries, and further that countries are willing to address this concern via taxation. This need not be the case. It remains an open question whether the tax system is an appropriate mechanism through which to pursue redistribution of global income. However, high-income nations should at least seek to not exacerbate an already inequitable global distribution. This is particularly true where the home and host countries both have some claim to tax the income at issue, as is the case with remitted labor income. In such an instance, the primary explanations for the regressive redistribution from poor to rich countries are the weak bargaining power and low tax capacities of low-income migrant-source nations. Neither of these reasons ought to impose a permanent barrier to a more equitable distribution of tax revenue between home and host countries.

C. Government Tax Goals Principles

Finally, looking to the home country perspective, taxation of remittance payments to the recipient household would support important tax goals of migrant-source country governments. Specifically, a properly crafted remittance tax policy would mobilize needed government revenue, and do so in a relatively progressive way.

1. Increasing Revenue Mobilization

Increasing tax revenue is an important objective of resource-poor countries seeking to improve development outcomes via the stable provision of security and other public goods. Although certainly there are reasons to be cautious when considering taxation of remittance flows, these challenges need not

---

144. See, e.g., Graetz, supra note 141, at 300-01 (discussing whether tax law’s role in redistributing income from rich to poor should stop at national borders).

145. See, e.g., Fiscal Affairs Dep’t, IMF, Revenue Mobilization in Developing Countries 7 (2011), http://www.imf.org/external/np/pp/eng/2011/030811.pdf (explaining that the spending needs of developing countries are substantial and constant); Thomas Baumsgaard & Michael Keen, Tax Revenue and (or?) Trade Liberalization, 94 J. PUB. ECON. 563, 565 (2010) (“[I]ncreased domestic revenue mobilization clearly is, and has long been, a central element of the development strategies of many low income countries. . . . [A]ny substantial loss of total tax revenue following trade reform . . . is prima facie cause for significant concern.”); Burgess & Stern, supra note 36, at 769 (“[T]here is no viable, long-term, and substantial alternative to taxation as a means of financing government expenditures.”); Deborah Itriago, Owning Developing: Taxation to Fight Poverty, OXFAM RESEARCH REPORT 9 (Sept. 2011) (describing the relatively low tax collections of developing countries compared with high-income countries, arguing that increased tax collections would help governments fight poverty).
overshadow the ongoing need for sustainable government resources in developing countries.

A government needs revenue to provide the public services necessary to ensure stability and long-term growth. Individual households cannot provide public goods. Thus, the government must impose taxes in order to raise the requisite funds. Underlying the opposition to remittance taxation is an implicit belief that the individual household is the optimal unit to leverage remittance payments to improve overall welfare. However, as explained above, empirical data fail to show a robust positive relationship between remittance receipts and long-term growth.146 This is partly due to the fact that remitted funds are often used for nondurable consumption rather than being invested in pro-growth activities.147 Although the government may not make better use of the funds in all instances, the fact remains that individual households have been unable to leverage remittances for sustained growth. Taxing remittances would help ensure that the transfers are partially funneled toward providing the public goods necessary to engender long-term growth.

Opponents to raising revenue via remittance taxation may counter that remittance recipients should not bear a greater burden than their neighbors in supporting overall domestic welfare. Although all taxation rests on the assumption that the government has a right to transfer resources from the private to the public sector in return for the provision of public goods, the distribution of this tax burden must be considered fair.148 This idea is expressed in part by the “benefit principle,” which states that a taxpayer’s tax liability should correspond to the benefits that the government provides.149 Taken a step further, this principle leads to the normative conclusion that those who receive greater benefits should pay higher taxes.150 Thus, in order for the benefits principle to justify remittance taxation, recipient households must receive greater public benefits than non-recipient households. The evidence suggests that they likely do—although admittedly the gap in benefits is difficult to measure and may be

146. See Abdih et al., supra note 75, at 664, and accompanying text; Acosta et al., supra note 75, and accompanying text; see, e.g., Barajas et al., supra note 59, and accompanying text; Chami et al., supra note 75, at accompanying text.

147. Barajas et al., supra note 59, at 6 (explaining that remittances may not lead to growth because they likely finance consumption rather than investment).

148. Perceptions of the fairness of the tax system are very important to constructing a functional tax system and ensuring broad compliance. See, e.g., Deborah A. Bräutigam, Introduction: Taxation and State-Building in Developing Countries, in TAXATION AND STATE-BUILDING IN DEVELOPING COUNTRIES 7 (Deborah A. Bräutigam, Odd-Helge Fjeldstad & Mick Moore eds., 2008) (“[C]ompliance will be affected by perceptions of the government’s legitimacy and the fairness of the tax system . . . .”).


150. See SLEMROD & BAKIJA, supra note 149, at 63.
small. As one example, many developing country governments undertake specific efforts to promote remittances and ensure safe and cost-effective transfers. Additionally, remittance transfers result from migration of family members who often benefitted from state-subsidized education. These public services are typically considered to be government investment in the economy, but without a way to harness remittance inflows, much of the consequent gains are lost or, at best, spent on basic consumption rather than long-term growth. Taxing the transfers upon their return is perhaps the most straightforward way for governments to ensure a positive return on their public investment.

The revenue argument is particularly compelling for countries with very high emigration relative to the total population, as taxing remittances may be one of the few ways to compensate the government for the tax revenue lost due to emigration. In a country where remittances make up a large percentage of domestic GDP, there may be few other viable domestic sources of tax revenue. In Tajikistan, for example, remittances comprise forty-eight percent of GDP. Countries such as Tajikistan may have few other options outside of taxing remittances in order to raise enough revenue to provide basic public services for their citizens.

2. Progressivity

The goal of raising revenue must be balanced with the equally important goal of ensuring progressive taxation. Taxing remitted income offers a relatively progressive tax policy compared with both current non-taxation and with alternative emigrant tax proposals such as the Bhagwati tax or the global citizenship taxation proposed by Desai et al.

Taxing remittance-receiving households may improve progressivity in part by targeting a group of taxpayers that is often relatively better off than their non-recipient neighbors. Of course, the progressivity of a remittance tax policy will depend on the demographics of each particular country, so taxing recipient

151. See, e.g., Zeno Ronald R. Abenoja, Presentation at the 9th National Convention on Statistics, Promoting Greater Use of Formal Remittance Systems by Overseas Filipinos 12 (Oct. 4–5, 2004) (listing various efforts by the Philippine government to promote remittances, including international agreements to reduce remittance costs, support for overseas workers, and easing restrictions on banks to enable them to increase money-changing services, among other things); see discussion supra Part I(C) (describing Mexico’s efforts to promote and protect remittance transfers).


154. See supra Part I(A)(ii).
households may not improve progressivity in all contexts. That said, remittance recipient families often inhabit middle- or high-income groups in the home country, even if a worker occupies a low-income group in the host country.155 Further, research from Egypt demonstrates that remittances can worsen inequality in situations where higher-income households contribute more to emigration.156 Ensuring a progressive tax system is particularly important where remittances tend to worsen inequality or benefit higher-income households.

Taxing remittance recipients also offers a more progressive method of compensating migrant-source countries for emigration compared to emigrant tax regimes that Bhagwati and Desai et al. propose. Specifically, remittance taxation is superior to emigrant taxation because it can account for a taxpayer’s ability to pay the tax. The “ability-to-pay principle” states that tax burdens should be measured in accordance with a taxpayer’s economic well-being.157 Determining a taxpayer’s ability to pay is inherently relative, introducing additional complexities in an international tax context.158 When contemplating a migration-related tax policy that affects citizens living abroad, policymakers must decide whether to compare a taxpayer’s ability to pay with the host country’s population or the home country’s population. A taxpayer may be relatively well off compared to one group while being relatively poor compared to the other. For this reason, imposing a home-country tax on an emigrant living abroad may worsen distributional outcomes compared to the host-country populace. Taxing recipient households on remittance payments, on the other hand, ensures an appropriate comparison of tax burden only among residents in the same domestic economy. Therefore, where progressivity is a significant concern, a tax on remittance recipients is preferable to the emigrant tax policy proposed throughout migration taxation scholarship.

As the above analysis demonstrates, taxing remittance recipients would advance the important home-country tax goal of raising revenue, and do so in a more progressive way than is currently proposed in the literature.

155. See, e.g., de la Fuente, supra note 71, at 838 (finding that remittances tend to benefit households that are better off than non-recipients).


157. See, e.g., Slemrod & Bakić, supra note 149, at 165–67 (explaining that distributional equity is achieved in part through the use of ability-to-pay measurements); Kirsch, supra note 152, at 479 (describing the ability-to-pay principle as one “traditionally applied to questions of distribution”).

158. See Kirsch, supra note 152, at 480–84.
III.
TAX REALITIES IN DEVELOPING COUNTRIES

A. Tax Policy Building Blocks

Policymakers seeking to compensate domestic economies for emigration will need to consider on-the-ground tax policy realities. Governments can reach recipient household finances through three main tax instruments: the individual income tax, the consumption tax, and the property tax. Remittance-receiving countries with a functioning progressive income tax regime should be able to tax remittance inflows primarily via taxing income to the recipients. Those without an income tax will have to look to other options. For example, these nations can rely on indirect taxation or forge tax-sharing treaties with host countries. To better understand the policy proposals described below, this section provides a brief background on tax structures in developing countries and highlights specific tax characteristics among a group of remittance-receiving nations.

1. The Personal Income Tax (PIT)

The income tax should be the first choice for policymakers seeking to harness remittance gains because it is one of the few moderately progressive taxes among developing countries. Unfortunately, however, the personal income tax in developing countries faces significant constraints. Although wealthy countries raise substantial revenue via income tax, on average collecting the equivalent of seven percent of GDP, developing nations raise only the equivalent of two percent of GDP through the tax. This is due in part to local administrative and political difficulties in implementation and enforcement. In practice, developing country governments rarely collect more than withholding taxes on employees working in the formal sector, specifically those working for the government or for large corporations. Tax agencies are often unable to reach labor income earned in the large informal sector.

In spite of continually low collection rates, Alberto Barreix and Jerónimo Roca describe the income tax as one of the primary fiscal pillars in Latin America because it promotes progressivity. Barreix and Roca argue that

162. REVENUE MOBILIZATION IN DEVELOPING COUNTRIES, supra note 145, at 31; Bird & Zolt, supra note 36, at 1629.
developing countries should pay greater attention to the personal income tax because of its potential ability to reduce income inequality and, by extension, promote social cohesion. Ky-young Chu et al. find further support for the tax’s progressivity, demonstrating that virtually all studies of the income tax’s incidence from 1975–1998 document its progressivity and redistributive effect. However, much of the literature also notes income tax’s declining progressivity over time as well as its diminishing importance in the developing world. Thus, while an ideal remittance policy would utilize the personal income tax to capture gains, a realistic policy portfolio must look to other options, as well.

2. Consumption Taxation—The VAT

The widespread adoption of the value-added tax (VAT) over the past two decades has nearly amounted to a tax policy revolution in developing countries. The VAT has been promoted and adopted in part for its large revenue-raising potential because of the efficacy of the tax on its own, as well as its tendency to improve overall tax administration and compliance. Indeed, generally the VAT is found to increase revenue. Further, despite popular criticisms of the tax as regressive, it is generally found to have little effect on income distribution, with the exception of certain sector-specific taxes. For example, a tax on kerosene will be more regressive than a tax on luxury goods.

However, the literature notes that certain countries have had less success raising revenue through the VAT. Poor performance is associated with political instability and administrative limitations, particularly in Sub-Saharan African countries. Joshua Aizenman and Yothin Jinjarak caution that enthusiasm for

---

165. Id. at 138.
167. See also Bahl & Bird, supra note 163, at 289 (noting a sharp reduction in corporate and personal income tax rates around the world, reflecting increased international tax competition); Bird, supra note 160, at 2; Chu, supra note 166, at 42–46.
the VAT should be tempered by the awareness that the VAT is not costless to administer, due to large expenditures to collect and process information, prosecute agents found underpaying the tax, and otherwise ensure broad compliance with the complex tax structure.\textsuperscript{173} Importantly, Joseph Stiglitz criticizes overzealous promotion of the VAT because of its inability to reach the informal sector, leading to significant distortions in consumption and employment.\textsuperscript{174} In all, while the VAT is undoubtedly a useful tax for developing countries, it may not be the ideal tax to capture remittance gains because of its weaknesses in capturing the informal sector and other limitations discussed in the following section.\textsuperscript{175}

3. The Property Tax

Developing countries infrequently use property taxation, on average raising revenue equal to only 0.6\% of their GD\Ps.\textsuperscript{176} Common explanations for the dearth of property taxation include the high costs of accurate valuation,\textsuperscript{177} political unpopularity,\textsuperscript{178} and enforcement difficulties.\textsuperscript{179} However, despite the

\begin{itemize}
\item \textsuperscript{173} Aizenman & Jinjarak, supra note 172, at 393; see also Patrick Fossat & Michel Bua, Tax Administration Reform in the Francophone Countries of Sub-Saharan Africa 29 (IMF Fiscal Affairs Dep't, Working Paper No. WP/13/395, 2013) (detailing problems with VAT administration in Sub-Saharan Africa, including burdensome paperwork requirements, rent-seeking through lobbying for exemptions, lack of confidence in the tax system, etc.).
\item \textsuperscript{174} Joseph E. Stiglitz, Development-Oriented Tax Policy, in TAXATION IN DEVELOPING COUNTRIES 11, 11–12 (Roger Gordon ed., 2010) (inefficiency occurs as taxpayers shift into the relatively less productive informal sector, reducing growth and increasing unemployment as labor shifts to the informal sector as well).
\item \textsuperscript{175} See infra Part III(B)(i).
\item \textsuperscript{176} See Roy Bahl & Jorge Martinez-Vazquez, The Property Tax in Developing Countries: Current Patterns and Prospects 1 (Lincoln Inst. of Land Policy, Working Paper No. WP07RB1, 2007).
\item \textsuperscript{178} Id.
\item \textsuperscript{179} Enforcement difficulties occur because local officials lack the capacity and desire to enforce the tax, particularly because property owners are often community leaders. One exception here has been South Africa, where local authorities have used the threat of cutting off electricity for failure to pay property taxes. See Bahl et al., supra note 177, at 42. At the other end of the spectrum, in the United States local tax authorities have gone so far as to impose tax liens on property and eventually foreclose on homes for unpaid property taxes, even where tax debts are as low as $134. See Michael Sallah, Debra Cenziper & Steven Rich, Left with Nothing, WASH. POST, Sept. 8, 2013, http://www.washingtonpost.com/sf/investigative/2013/09/08/left-with-nothing/ (reporting on a controversial local tax policy of foreclosing on homes with unpaid tax debts). Clearly enforcement of property taxes must find a middle ground between no enforcement and foreclosure for minor
costs and challenges of imposing the tax, property taxation boasts many benefits that make it a good choice for developing countries, in particular at the local government level. For one, property taxes theoretically reflect the actual government services provided as long as these services are capitalized into property values.\footnote{180} Further, because the taxes will finance local services, there is a high level of correspondence between those who pay the tax and those who receive the benefits, which should increase the accountability of government officials.\footnote{181} Property also provides a highly stable revenue source for resource-starved localities, one that other levels of government mostly fail to tax.\footnote{182} This characteristic renders the property tax especially important for fiscal authority decentralization and development of local government capacity and autonomy.\footnote{183} At best, however, the current weakness of property taxation regimes relegates it to being merely one component of a migration tax policy portfolio. In the absence of significant expansion of the property tax, it is unlikely to adequately compensate home countries for outmigration.\footnote{184}

\section*{B. Tax Capacity of Remittance-Receiving Countries}

\subsection*{1. Survey of Tax Policies in Selected Migrant-Source Countries}

Migrant-source countries might rely on a mix of tax structures to harness remittance flows as compensation for outmigration. Table 3 provides a snapshot of the tax policy landscape among a group of remittance-receiving countries, including data for the top ten aggregate remittance recipients as well as data for the top ten recipients by percentage of GDP.\footnote{185} The table reveals the wide range in tax policies and taxing capacities in remittance-receiving developing countries. The absence of data for property taxation reveals both the relatively low utilization of the tax as well as the lack of institutional knowledge about property tax implementation in developing countries.

The data in Table 3 is drawn from the U.S. Agency for International Development’s (USAID) Collecting Taxes project, which assembles national tax debts.

\footnote{180. Bahl & Martinez-Vazquez, supra note 176, at 35; see also Enid Slack, Presentation at the Fourth IMF-Japan High-Level Tax Conference, Property Tax Reform in Developing Countries 4 (Apr. 3, 2013).}

\footnote{181. Bahl & Martinez-Vazquez, supra note 176, at 4–5.}

\footnote{182. Id. at 2–3, 8.}

\footnote{183. See, e.g., Richard M. Bird & Enid Slack, Property Tax and Rural Local Finance, in MAKING THE PROPERTY TAX WORK: EXPERIENCES IN DEVELOPING AND TRANSITIONAL COUNTRIES 103, 103–04 (Roy Bahl, Jorge Martinez-Vazquez & Joan Youngman eds., 2008) (arguing that the property tax is vitally important for rural development because it is necessary to allow localities to provide public services and to ensure viable and effective local government).}

\footnote{184. For the property tax to capture emigration gains, some portion of remittance-receiving households would need to own property. The likelihood of this is addressed infra Part IV(c)(i).}

\footnote{185. For remittance amounts and percentages for these countries, see supra Tables 1 and 2.}
structure and performance data for all countries annually. Although most of the categories are relatively straightforward, several of the performance indicators require further explanation. The PIT Productivity column presents figures from the USAID’s “PITPROD” indicator. This indicator seeks to capture the revenue production performance of the PIT in each country and is calculated by dividing the total PIT revenue as a percentage of GDP by the average PIT rate. The indicator will fall between zero and one. The VAT Gross Compliance column presents figures from the USAID’s “VATGCR” indicator, which is intended to measure how well the VAT produces revenue for the government. The figure is calculated by dividing total VAT revenue as a percentage of GDP by the product of total private consumption and the VAT rate. In this way, it measures actual VAT collections divided by potential VAT collections, expressed as a percentage.

The broad takeaways from the table touch upon tax policy design in several ways. For one, relatively low levels of tax effectiveness—as in, for example, India, Liberia, Haiti, and Nigeria—might counsel policymakers in those countries toward tax reforms that avoid direct self-reporting and other input-intensive methods. In contrast, those countries with a relatively productive personal income tax—such as Vietnam, Lesotho, and Mexico—will likely be more successful at capturing remittance gains through a standard household income tax compared with other low-income countries. Countries that lack a robust income tax—such as Bangladesh and Liberia—need not avoid income taxation altogether but should perhaps instead implement tax policies that do not rely on self-reporting, such as source withholding or presumptive taxation.

---


188. Id.

189. Id.

190. Id.

191. Presumptive taxation entails utilizing administrative assessments where possible in order to avoid reliance on self-assessments for “hard to tax” entities that can easily evade tax authorities. See Richard M. Bird & Sally Wallace, Is It Really So Hard to Tax the Hard-to-Tax? The Context and Role of Presumptive Taxes (Andrew Young Sch. of Policy Studies, Int’l Tax Program Paper No. 0307, 2003); Sona Gandhi, Presumptive Direct Taxes, WBG (2011), http://go.worldbank.org/OC0j6VGS1 (last visited Feb. 12, 2015). Methods of presumptive taxation can include: standard lump-sum taxes based on visible indicators such as occupation or business activity, estimated income assessments based on various indicators such as number of employees or land value, assessed agricultural income based on potential land output, comparing net wealth at the start and end of the year, taxation of visible signs of wealth, and minimum taxes. Id. In the context of remittance taxation the obvious proxy for income would be the amount of remittances received.
Importantly, the table also shows that some countries—such as Vietnam, Lebanon, and Pakistan—have particularly high thresholds before income tax applies. Without domestic tax reforms in these countries, many remittance recipients would be exempted from income taxation because they fall below these thresholds. Should these nations maintain their high income tax thresholds, they would need to capture remittance gains instead through indirect tax methods such as consumption taxation and property taxation.
Table 3: Remittance-Receiving Countries’ Tax Policies

<table>
<thead>
<tr>
<th>Country</th>
<th>Income Tax Threshold(^{193})</th>
<th>PIT Revenue to GDP</th>
<th>PIT Productivity</th>
<th>Average VAT rate</th>
<th>VAT Revenue to GDP</th>
<th>VAT Gross Compliance</th>
<th>Tax Revenue to GDP(^{194})</th>
</tr>
</thead>
<tbody>
<tr>
<td>India</td>
<td>$3,766</td>
<td>1.89%</td>
<td>0.07</td>
<td>12.5%</td>
<td>0.02%</td>
<td>0.2%</td>
<td>10.72%</td>
</tr>
<tr>
<td>China</td>
<td>$541</td>
<td>4.80%</td>
<td>0.11</td>
<td>17.0%</td>
<td>9.6%</td>
<td>164.2%</td>
<td>18.96%</td>
</tr>
<tr>
<td>Philippines</td>
<td>$1,084</td>
<td>1.99%</td>
<td>0.06</td>
<td>12.0%</td>
<td>1.9%</td>
<td>21.2%</td>
<td>12.35%</td>
</tr>
<tr>
<td>Mexico</td>
<td>0</td>
<td>5.40%</td>
<td>0.18</td>
<td>16.0%</td>
<td>3.8%</td>
<td>36.9%</td>
<td>15.90%</td>
</tr>
<tr>
<td>Nigeria</td>
<td>1,005</td>
<td>-</td>
<td>-</td>
<td>5.0%</td>
<td>0.2%</td>
<td>5.7%</td>
<td>11.55%</td>
</tr>
<tr>
<td>Egypt</td>
<td>1,351</td>
<td>1.39%</td>
<td>0.07</td>
<td>10.0%</td>
<td>3.0%</td>
<td>39.7%</td>
<td>14.01%</td>
</tr>
<tr>
<td>Bangladesh</td>
<td>$2,813</td>
<td>1.10%</td>
<td>0.05</td>
<td>15.0%</td>
<td>3.5%</td>
<td>29.8%</td>
<td>9.30%</td>
</tr>
<tr>
<td>Pakistan</td>
<td>$3,820</td>
<td>3.34%</td>
<td>0.17</td>
<td>16.0%</td>
<td>3.6%</td>
<td>27.0%</td>
<td>9.31%</td>
</tr>
<tr>
<td>Vietnam</td>
<td>$4,845</td>
<td>8.80%</td>
<td>0.25</td>
<td>10.0%</td>
<td>6.1%</td>
<td>97.2%</td>
<td>24.26%</td>
</tr>
<tr>
<td>Ukraine</td>
<td>0</td>
<td>0.47%</td>
<td>0.03</td>
<td>20.0%</td>
<td>10.0%</td>
<td>75.9%</td>
<td>18.49%</td>
</tr>
<tr>
<td>Tajikistan</td>
<td>$104</td>
<td>2.50%</td>
<td>0.23</td>
<td>20.0%</td>
<td>10.5%</td>
<td>57.1%</td>
<td>18.22%</td>
</tr>
<tr>
<td>Kyrgyz Republic</td>
<td>0</td>
<td>2.35%</td>
<td>0.24</td>
<td>12.0%</td>
<td>7.8%</td>
<td>77.1%</td>
<td>16.13%</td>
</tr>
<tr>
<td>Nepal</td>
<td>0</td>
<td>0.87%</td>
<td>0.03</td>
<td>13.0%</td>
<td>4.6%</td>
<td>43.2%</td>
<td>12.60%</td>
</tr>
<tr>
<td>Lesotho</td>
<td>$2,071</td>
<td>9.90%</td>
<td>0.35</td>
<td>14.0%</td>
<td>7.8%</td>
<td>56.4%</td>
<td>21.50%</td>
</tr>
<tr>
<td>Moldova</td>
<td>$516</td>
<td>2.20%</td>
<td>0.12</td>
<td>20.0%</td>
<td>12.7%</td>
<td>65.9%</td>
<td>18.28%</td>
</tr>
<tr>
<td>Armenia</td>
<td>0</td>
<td>2.15%</td>
<td>0.11</td>
<td>20.0%</td>
<td>8.7%</td>
<td>53.4%</td>
<td>17.32%</td>
</tr>
<tr>
<td>Haiti</td>
<td>$335</td>
<td>2.50%</td>
<td>0.08</td>
<td>10.0%</td>
<td>-</td>
<td>-</td>
<td>12.80%</td>
</tr>
<tr>
<td>Samoa</td>
<td>$3,706</td>
<td>2.60%</td>
<td>0.14</td>
<td>15.0%</td>
<td>6.1%</td>
<td>83.7%</td>
<td>22.84%</td>
</tr>
<tr>
<td>Liberia</td>
<td>$123</td>
<td>2.73%</td>
<td>0.08</td>
<td>7.0%</td>
<td>0.80%</td>
<td>9.1%</td>
<td>17.36%</td>
</tr>
<tr>
<td>Lebanon</td>
<td>$4,973</td>
<td>0.60%</td>
<td>0.03</td>
<td>10.0%</td>
<td>5.2%</td>
<td>65.1%</td>
<td>15.70%</td>
</tr>
</tbody>
</table>


\(^{193}\) Thresholds are the average rate for a standard single taxpayer with no dependents, converted to U.S. dollars using the exchange rate on April 11, 2016.

\(^{194}\) This measures total tax revenues broadly defined, including both domestic taxes and customs duties, as a percent of GDP. See USAID 2012–2013, supra note 187.
2. Why the VAT is Not Enough

Before describing specific policy structure proposals, it is worthwhile to address why it is necessary to harness remittance gains through direct taxation of remittance-receiving households. Intuition suggests that if a country has a functional VAT it should already effectively capture the inflows when recipients spend the remitted funds on consumption. As Table 3 demonstrates, nearly all developing countries utilize a VAT, albeit with varying levels of efficacy. Yet consumption taxation alone has not managed to transform remittance inflows into long-term economic growth and stability. There are several possible reasons for this, which suggest that targeted remittance taxation would be more effective at compensating developing countries for emigration losses compared to relying on consumption taxation alone.

Various factors reduce the revenue collection efficacy of the VAT, including political instability, burdensome information and processing costs, and broad administrative limitations due to low levels of institutional development.195 Although the widespread adoption of the VAT has increased revenue collection on average throughout the developing world, certain countries have been less successful due, in part, to these political and administrative constraints. In particular, Sub-Saharan African countries have had difficulty utilizing the VAT to increase revenue collection because of burdensome paperwork requirements, political pressure to expand exemptions, and a lack of confidence in the tax system.196 Table 3 demonstrates this national variation in VAT efficacy. Countries with low compliance rates—such as Nigeria, Liberia, and India—or those in which the VAT raises little revenue relative to GDP—such as the Philippines, Egypt, and Bangladesh—should not expect the VAT to adequately compensate their domestic economies for emigration losses.

The liberal use of VAT exemptions and zero rates in developing countries further weakens the tax’s ability to harness gains from remittance inflows. Table 4 lists categories of exempted consumption goods in remittance-receiving countries. In particular, many developing country governments face political pressure to exempt basic consumption goods from taxation, such as food, medical services, and cooking fuel.197 These exemptions will prevent remittance-receiving governments from harnessing much of the gain from remittance inflows because much of those remittances are spent on exempted basic consumption goods.198

195. See, e.g., Aizenman & Jinjarak, supra note 172, at 393 (finding a correlation between political instability and reduced VAT effectiveness); Edmiston & Fox, supra note 172, at 249–59 (arguing that VAT underperformance arises due to political and administrative realities and constraints); Fossat & Bua, supra note 173, at 29 (describing VAT paperwork requirements, rent-seeking, and lack of confidence in the tax system, etc.).
196. Fossat & Bua, supra note 173, at 29.
198. IFAD, SENDING MONEY HOME, supra note 60, at 7.
Table 4: Examples of Goods Exempted from Consumption Taxation in Select Remittance-Receiving Countries

<table>
<thead>
<tr>
<th>Country</th>
<th>Selected Consumption Goods Exempted from Taxation or Subject to a Zero-Rate Tax</th>
</tr>
</thead>
<tbody>
<tr>
<td>India</td>
<td>Items subject to a zero-rate tax include agricultural tools and products, such as fresh fruits, vegetables, and other food products; books, periodicals, and journals; and electricity. 199</td>
</tr>
<tr>
<td>China</td>
<td>Agricultural products, contraceptive drugs and devices, antique books, and other items declared by the State Council are exempted. 200</td>
</tr>
<tr>
<td>Philippines</td>
<td>Goods sold by small businesses or small producers, 201 educational services, 202 and sales by registered cooperatives are exempted. 203</td>
</tr>
<tr>
<td>Mexico</td>
<td>Items subject to a zero-rate tax include most non-industrialized animals and vegetables; patent medicines; most products intended for food, including meat, milk, and eggs; ice and water; tractors and other farm equipment; jewelry; and books and magazines. 204</td>
</tr>
<tr>
<td>Nigeria</td>
<td>Essential goods are exempted including medical and pharmaceutical products, basic food, books and educational materials, baby products, fertilizer and agricultural materials, veterinary medicine, and farming transportation equipment. 205</td>
</tr>
<tr>
<td>Vietnam</td>
<td>Certain agricultural products and salt are exempted entirely, and a reduced five-percent rate applies to essential goods and services</td>
</tr>
</tbody>
</table>


201. E&Y VAT GUIDE, supra note 199, at 752 (noting that businesses with gross receipts below a certain amount are not required to register as a VAT taxpayer); see also Value-Added Tax, REPUBLIC PHILIPPINES BUREAU INTERNAL REVENUE, http://www.bir.gov.ph/index.php/tax-information/value-added-tax.html (last visited Mar. 20, 2016).

202. KPMG, ASIA PACIFIC GUIDE, supra note 199, at 42.

203. E&Y VAT GUIDE, supra note 199, at 756.


such as fresh foodstuffs, fertilizer, clean water, medical and educational equipment, and books.\textsuperscript{206}

| Ukraine | Infant foods, education and healthcare services, books and periodicals, housing, religious services, and funeral services are exempted.\textsuperscript{207} |

In addition to administrative and political constraints, the presence of a large informal sector in many developing countries will further reduce governments’ abilities to harness remittance inflows via consumption taxation.\textsuperscript{208} Many low-income countries have large informal cash economies, particularly for items such as food and agricultural products,\textsuperscript{209} on which a large proportion of remitted income is spent.\textsuperscript{210} When purchased products are created or distributed through informal markets, they will escape consumption taxation either partially or fully. Certain remittance-receiving countries may be especially sensitive to this concern. For example, IMF research from 2008 on informal economies in Central Europe estimates that the informal sector accounts for 26.3\% of GDP in Kyrgyz Republic, 32.8\% of GDP in Tajikistan, and 35\% of GDP in Armenia.\textsuperscript{211} Thus, for countries with a large informal sector, consumption taxation will prove a particularly ineffective method for capturing remittance gains.

Finally, taxing remittance receipts through the income tax is worthwhile in and of itself because the income tax is the most progressive tax instrument available to policymakers seeking to improve economic outcomes.\textsuperscript{212} Despite administrative difficulties and low collection rates, the taxation and development literature continues to describe the personal income tax as one of the few progressive taxes in the developing world.\textsuperscript{213} The point here is not that consumption taxation is ineffective or necessarily regressive, but rather that a remittance-receiving government will likely capture remittance gains more effectively and more progressively if consumption taxation is combined with household income taxation in some capacity.\textsuperscript{214}

\textsuperscript{206} KPMG, ASIA PACIFIC GUIDE, supra note 199, at 56.
\textsuperscript{207} E\&Y VAT GUIDE, supra note 199, at 992–93.
\textsuperscript{208} See, e.g., Stiglitz, supra note 174, at 11–12 (arguing that the VAT produces inefficient outcomes in developing countries due to the strength of the informal sector, which escapes VAT taxation); M. Shahe Emran & Joseph E. Stiglitz, On Selective Indirect Tax Reform in Developing Countries, 89 J. PUB. ECON. 599, 621 (2005).
\textsuperscript{210} IFAD, SENDING MONEY HOME, supra note 60, at 7.
\textsuperscript{212} See supra Part III(A)(i).
\textsuperscript{213} See, e.g., Bird & Zolt, supra note 36, at 1682–83.
\textsuperscript{214} Utilizing multiple tax mechanisms to compensate for lost government revenue is not unprecedented and often found to be a more effective revenue recovery strategy. For example, after the international development community advised developing countries to reduce or eliminate trade taxes, research has shown that developing country governments are better able to recover lost...
IV. PROPOSED POLICY STRUCTURES

The following policies aim to compensate developing countries for emigration losses via taxing income remitted to transnational households. This Article lays out three different policy structures to account for differing tax capacities among developing countries. All three policies are designed to ameliorate potential double taxation and include modifications that allow those policies to work in countries with weak personal income tax systems. In broad brushstrokes, an ideal remittance taxation policy would utilize bilateral treaties between host and home countries to allow deduction of remittance payments by the sender and enable subsequent taxation of recipient households in home countries. Where bilateral treaties are not feasible, under a second-best policy home countries would tax recipient households directly, either through self-reporting or withholding by transfer companies, and utilize tax credits to counteract double taxation. Finally, the third policy package suggests ways in which developing countries can capture remittances through indirect taxation of consumption and property, either in lieu of or in addition to income taxation.

A. Policy 1: Deduction in the Host Country via Bilateral Tax Treaties

This policy involves negotiating bilateral tax treaties between wealthy host nations and low-income home countries to provide for deduction of remittances in the former and subsequent taxation in the latter. Coordinating such a policy would entail significant cooperation by high-income host countries because they would be required to cede tax revenue on remitted funds. Although perhaps politically challenging, such a concession is justified based on first principles of family transfer taxation as well as inter-nation equity considerations, as described above.215

Inter-nation tax generosity is not unprecedented. Even as far back as the 1960s, U.S. President John F. Kennedy proposed ceding source-based corporate tax revenue to developing nations in the spirit of equity.216 In seeking tax generosity towards developing economies, President Kennedy argued that, "The free world has a strong obligation to assist in the development of these economies, and private investment has an important contribution to make."217 In the context of labor migration, this kind of tax policy can be presented as a revenue where they combined consumption taxation with improved income tax collection. FISCAL AFFAIRS DEP’T, IMF, DEALING WITH THE REVENUE CONSEQUENCES OF TRADE REFORM 22 (2005).

215. See discussion supra Parts II(A)–(B).
217. Id.
trade-off whereby the benefiting host nation\textsuperscript{218} cedes tax revenue in return for the home country allowing unfettered emigration of labor. As Desai et al. argue, this mutual dependency between high-income labor consumers and low-income labor providers makes such bilateral cooperation increasingly likely in today’s world.\textsuperscript{219} A country such as India would be a prime candidate for this policy, given the combination of its relatively weak tax capacity\textsuperscript{220} and relatively strong bargaining power due to its size and the desirability of high-skilled Indian labor to host economies.\textsuperscript{221}

In addition to political feasibility, there are various important logistical considerations in determining how the tax revenue will be tabulated by the sending country and collected in the recipient country. Calculation in the host country could occur either at the individual level or in the aggregate for each recipient country. If calculated at the individual level, individual remittance senders would deduct remittance payments from their gross income when they file taxes, perhaps providing receipts or other proof to verify the reported transfers. In the United States, this proof would be facilitated by the Electronic Fund Transfers Act, which requires that all remittance services in the United States provide receipts to customers.\textsuperscript{222} This remittance payment would subsequently be taxed to the recipient household by the home country government. If done at the aggregate level, host country governments would send tax revenue associated with remittances to each home country to approximate the revenue collected from remitted funds. The following subsections explore both of these options in greater detail.

---

\textsuperscript{218} Although perpetually controversial, abundant empirical research generally finds that migration, even unskilled migration, benefits host countries’ economies in certain ways. See Patricia Cortes, \textit{The Effect of Low-Skilled Immigration on U.S. Prices: Evidence from CPI Data}, 116 J. Pol. Econ. 381, 401 (2008) (finding that a 10\% increase in unskilled labor in the United States results in a 2\% decrease in prices of immigrant-intensive services); George J. Borjas & Lawrence F. Katz, \textit{The Evolution of the Mexican-Born Workforce in the United States} 42–43 (Nat’l Bureau of Econ. Research, Working Paper No. 11281, 2005) (finding that the influx of immigrants from Mexico between 1990–2000 improved earnings of college graduates and lowered prices of goods and services that rely on low-skill labor). Further research suggests that an increase in unskilled migrant labor increases the labor market participation of high-skilled women. See Patricia Cortés & José Tessada, \textit{Low-Skilled Immigration and the Labor Supply of Highly Skilled Women}, 3 AM. ECON. J.: APPLIED ECON. 88 (2011) (finding that low-skilled immigration increases average hours of market work and the probability of working long hours for women at the top quartile of the wage distribution).

\textsuperscript{219} Desai et al., \textit{supra} note 1, at 684.

\textsuperscript{220} See \textit{supra} Table 3.

\textsuperscript{221} See Ayelet Shachar, \textit{Highly Skilled Immigration: The New Frontier of International Labor Migration}, 105 AM. SOC’Y INT’L LAW: PROC. ANN. MEETING 415, 416 (2011) (exploring increased international competition for high-skilled immigration, in particular from China and India, “the major sending countries for highly skilled migrants in the fields of science and technology”).

\textsuperscript{222} Electronic Fund Transfers (Regulation E), 12 C.F.R. § 1005.31(b)(2)(i) (2011).
1. Aggregate Transfer of Tax Revenue

If calculated in the aggregate, the migrant worker would remit money to his family back home, and that remittance would end the individual households’ transaction and their role in the transfer policy. The host country government would then transfer associated tax revenue to the recipient country based on information received from transfer companies about the total quantity of remittances sent to that country and likely a prior agreed-upon assumed average tax rate. This aggregate calculation method may be preferable to individualized deductions and subsequent taxation for a number of reasons. For one, the administrative costs of an aggregate policy are much lower than an individualized approach, as neither the host nor the home country needs to invest resources in educating the public and monitoring a new national tax program. The policy still entails certain administrative costs, especially for the host country, which would need to calculate the share of tax revenue to transfer to each home country. Such a policy also essentially eliminates the risk of taxpayer error and tax gaming, both of which are significant problems plaguing the individualized approach. Finally, the success of the policy would be independent of the income tax capacity in developing home countries, which, as explained above, is often limited.

However, an aggregate high-to-low-income country revenue transfer is not without its drawbacks. Importantly, such a policy would be largely divorced from direct contact with citizens, which reduces public oversight.\(^{223}\) Indeed, in practice this type of transfer feels more like development aid than tax revenue. Additionally, home countries would be largely unable to utilize the tax revenue transfer to promote progressive redistribution through the tax system since individual households would essentially be removed from the tax revenue transaction. One possible progressivity-enhancing remedy would be for recipient countries to allow very low-income remittance recipients—those that fall below home country income-tax thresholds—to file for a refund of withheld tax revenue. These households could do so by providing proof of remittances received and household income. A rebate policy of this kind might even encourage increased use of formal remittance services, as the use of formal services would be a necessary prerequisite for the tax recognition of the flows.

In all, an aggregate inter-nation revenue transfer would be less costly and present fewer risks than an individual deduction and subsequent taxation. However, despite the inherent complexities—or perhaps because of them—an

\(^{223}\) There is some evidence that a government’s fiscal dependence on its citizens is positively correlated with governance quality. See Mick Moore, *Revenues, State Formation, and the Quality of Governance in Developing Countries*, 25 INT’L POL. SCI. REV. 297 (2004) (investigating whether the quality of governance in developing countries would improve if States were more dependent for their financial resources on domestic taxpayers, cautioning against firm conclusions).
individualized approach offers numerous advantages that may counsel a remittance-receiving nation to favor such a policy.

2. Individual Deductions and Subsequent Taxation in Home Countries

Collection at the individual level entails higher administrative costs compared with an aggregate transfer, as well as increased risk of error, tax gaming, and behavioral distortions. Notwithstanding these drawbacks, however, such a policy also allows for greater progressive tailoring and creates the possibility of improving home country income tax capacity, among other advantages. If calculated individually, the migrant remittance sender would reduce his gross income on his tax return by the amount remitted annually. The recipient country would then separately tax the recipient household on the income received, presuming that the household receives income above the minimum income tax threshold. Bilateral tax treaties could be used to ensure that only migrant workers from certain nations—i.e., those that agree to tax recipient households on received funds—would be allowed to deduct the remitted income. In this way, host countries can cede the right to tax remitted income to the recipient nation while still ensuring that the income exclusion is not used to avoid taxation altogether.

Because this policy is a tax subsidy in the host nation, it is worthwhile to consider its possible effects on remittance activity. In all, the total behavioral effect of the deduction and subsequent taxation will depend on the interaction of various factors, including the difference between home- and host-country tax rates, the elasticity of remittances to taxation, and the role of remittances in the lives of senders and recipients. A tax deduction in the host country may incentivize increased remittance activity to some extent. However, subsequent taxation by the home country may ameliorate this incentive. Thus, overall behavioral changes should be minimal as long as the tax rates of the home- and host-countries do not diverge greatly. Consider, for example, a situation where the remittance sender provides his home-country relatives with a fixed amount of regular income for basic consumption. In this case, the sender will increase the overall payment to account for the subsequent tax, resulting in a similar post-tax, post-remittance income compared to the case if there were no deduction. The same result occurs where the sender provides himself a certain fixed income in the host country. In such a case, the tax deduction allows the sender to increase the remittance payment and remain left with the same net income. This increased remittance payment would again be reduced by a subsequent tax in the

224. Another interesting, albeit, initially more complex option is for migrant workers’ employers to contract with workers to remit a certain amount of income pre-tax, much like employer-provided health insurance premiums or healthcare spending accounts. Of course, an employer-centered policy would omit self-employed migrant workers and those working in the informal sector.

225. For a discussion of remittance elasticity and taxation, see supra Part II(B) and infra Part V(A).
home country. The end result, as with the aggregate transfer described above, is a net transfer of tax revenue from the host country government to the home country government. Individual participants are left with roughly the same income after taxes. Keep in mind, however, that where host- and home-country tax rates diverge greatly, there is higher risk of behavioral distortions as remittance senders may either shift income to lower-tax locales or avoid increased taxation by reducing remittance payments.226 Thus, where countries want to avoid changes in remittance-sending behavior, the bilateral tax treaties that create such a system should seek to ensure that average tax rates are roughly the same between host- and home-countries.

The individualized policy’s success requires relatively high income tax capacity in the home country, such as in Vietnam or Mexico.227 Even in nations with a robust income tax system, reliance on self-reporting creates significant opportunities for evasion by taxpayers—a persistent problem confronting income taxation in developing countries. However, hope is not lost merely because income taxation in developing countries is complicated. Indeed, remittances might actually prove easier to tax than other forms of income in developing countries. This is because the hard-to-tax nature of income in the developing world arises in part out of the pervasive informality of low-income economies.228 Unlike other forms of income, remittance transfers are often formal, offering a rare and ready third-party to enforce a remittance tax policy. Financial transfer companies could either report the transfers to the national tax agency or provide direct withholding of tax upon receipt in the home country. Where withholding is used, households could apply for an income tax refund to ensure the progressivity of the policy. Such third-party reporting might even offer tax administrators a unique opportunity to bring income into the tax net, capturing households that otherwise escape the reach of taxing authorities.229

226. Note, however, that increasing a tax on remittances need not reduce overall payments and some evidence suggests that it may even increase payments. See discussion infra Part V(A).
227. See supra Table 3.
228. See Burgess & Stern, supra note 36, at 776–77 (explaining why the personal income tax accounts for so little tax revenue in developing countries); Keen, Taxation and Development, supra note 36, at 10 (noting that effective progressive personal income taxation, which is often bolstered by withholding and third-party reporting, has eluded developing country officials).
229. Although beyond the scope of a remittance taxation policy proposal, remittance transfer companies could play a broader role in expanding income tax capacity by facilitating presumptive taxation of remittance recipients. For a brief explanation of presumptive taxation, see supra note 191. Many countries utilize presumptive taxes as a stopgap measure that allows them to capture mid-level taxpayers (e.g., small businesses), while at the same time sheltering taxpayers from tax complexities and allowing tax agencies to focus on the bigger fish. Bird, supra note 160, at 3–4. However, according to Bird these systems are usually poorly designed and poorly integrated into the regular tax system, making them a “dead end.” Id. at 4. Presumptive taxation is not exactly on point here because most presumptive taxation systems are based on assessed income and thus distinct from comprehensive approaches that attempt to tax actual income. See Gandhi, supra note 191.
Although self-reporting would be administratively more difficult than the inter-nation aggregate transfer described above, an individualized policy has certain advantages. First, such a policy would allow the low-income home country to impose specific domestic tax rates, brackets, and thresholds. Allowing this kind of tailoring would enable the use of a progressive rate structure targeted to achieve domestic income redistribution goals. Of course, if increased progressivity entails lowering the tax rate on low-income households, this heightens the risk of behavioral distortions and tax gaming as described above. These concerns would need to be navigated by specific countries in the course of negotiating bilateral treaties. A second advantage of self-reporting is that it would strengthen the connection between the taxpayers and the State—a commonly cited strength of the personal income tax in general. This increased connection may lead to better fiscal oversight and thus improved governance outcomes and more targeted public spending. Finally, relying on self-reporting and domestic collection of revenue would help develop tax capacity in the low-income home country. This is partly because, as explained above, requiring reporting of remittances received would bring more households into the income tax net. Thus, although self-reporting poses administrative challenges, the benefits of such an approach might outweigh the drawbacks in the long run and counsel low-income migrant countries to strengthen their income tax capacity in order to implement such a policy.

As a final consideration, high-income host countries may wish to improve the efficacy of the policy by limiting the remittance deduction to countries with a certain minimum level of income tax administration and enforcement capacity or to those with demonstrably low levels of corruption. Without such restrictions, allowing carte blanche deduction of all transfers to low-income countries would likely siphon off tax revenue without contributing to growth. However, host countries should avoid imposing too many restrictions beyond those that seek to ensure the general growth-enhancing utility of the policy. Imposing restrictions that require specific fiscal policies or governance structures, for example, would undermine national sovereignty of migrant-source countries and likely weaken support for an otherwise worthwhile program.

B. Policy 2: Tax Inflows to Recipients, Offsetting for Double Taxation

Where negotiating a pro-growth bilateral tax treaty is not politically feasible, developing home countries can instead tax remittance inflows to recipient households by utilizing a tax subsidy to prevent double taxation and offset taxes already paid. This policy option is politically more feasible than Policy 1, but it entails greater administrative complexities and results in less tax revenue for home country governments. These heightened complexities occur both because of basic collection difficulties in countries with weak personal

230. See, e.g., Barreix & Roca, supra note 164, at 138; Bird, supra note 160, at 1; but see Moore, supra note 223, at 300–04.
income tax regimes and also because of accounting complications necessary to offset taxes already paid in the host country. Vietnam would be a possible candidate for this second-best policy because of its relatively strong tax capacity. 231

Before addressing the specific policy structure, it is worth digressing for a moment to address double taxation in the context of remittance payments. In general, taxing the same income twice is considered a major problem to be avoided in tax system design.232 It is particularly salient in international taxation, where the country of income source and the country of taxpayer residence differ, yet both have tax claims to the same income. 233 Avoiding double taxation is one of the primary arguments against taxing remittance transfers in home countries;234 nevertheless, such fears should not wholly preclude remittance taxation for two reasons. First, where a migrant remitter pays no income taxes in the work country, this income would be exempted from taxation entirely unless the home country taxes it. 235 This will occur where a migrant worker occupies a low-income group in the host country and thus falls below the standard deduction and personal exemption amount. 236 It may also occur because the worker’s pay is not subject to employer withholding and the worker chooses not to satisfy his income tax obligations—perhaps because he does not feel a connection to the host country or is simply ignorant of tax filing laws. In the United States, for example, nonpayment of taxes often occurs because a worker’s employer has misclassified him as an independent contractor whose pay does not require tax withholding. 237 Where the host country fails to tax workers’ income, double taxation is not a concern.

231. See supra Table 3.
233. See supra note 133–135 and accompanying text.
234. See Mohapatra, supra note 3.
235. Note that this Article refers specifically to income taxation and not to other payroll taxes, such as Social Security or Medicare taxes.
236. See KRISHNASWAMI ET AL., supra note 62, at 12–13 (in a survey of Latin American immigrants to the United States, in which ninety percent of respondents sent regular remittances, that thirty-seven percent earn less than $1,200 monthly).
237. Employers often misclassify workers to avoid tax obligations, leading to complications for workers who are left with higher tax liabilities and lack of access to certain social services. See, e.g., TREASURY INSPECTOR GEN. FOR TAX ADMIN., EMPLOYERS DO NOT ALWAYS FOLLOW INTERNAL REVENUE SERVICE WORKER DETERMINATION RULINGS 2 (2013), https://www.treasury.gov/tigta/audireports/2013reports/201330058fr.pdf (“The IRS estimates that employers misclassify millions of workers as independent contractors instead of employees . . . .”); WORKERS DEF. PROJECT, BUILD A BETTER TEXAS ii (2013), http://www.workersdefense.org/Build%20a%20Better%20Texas_FINAL.pdf (finding that more than forty percent of construction workers in Texas are misclassified as independent contractors).
The second response to the double-taxation concern is that a well-crafted tax policy can largely ameliorate burdensome double taxation. In the absence of a bilateral treaty described above, a home country can mitigate double taxation in several ways. One option is to deduct the foreign source funds from income. In theory, this is how countries currently address this concern, ceding tax authority over remitted income to the high-income source country. Another option, detailed further below, is to provide a credit for foreign taxes paid. The United States does this through a foreign tax credit for income taxes paid abroad.238 A third option is to tax remittances at a lower rate.239 Although this would not eliminate all double taxation, if the rate is sufficiently low, it would mitigate concerns of overly burdensome taxation while avoiding the complexities inherent in the tax credit structure. The second and third options merit further discussion.

Under the second option, the recipient country could avoid double taxation by providing a tax credit equal to the amount of tax already paid on the remitted funds.240 One element of complexity here is that the taxes are paid by the sender and then credited to the recipient. A useful policy model for this kind of credit is that utilized in I.R.C. Section 902. This section allows a credit for foreign income taxes attributable to a dividend that a foreign subsidiary pays to its U.S. parent corporation.241 The foreign tax is deemed “paid” by the qualified domestic corporation claiming the credit, despite the fact that it was directly paid by the foreign subsidiary.242 Box 1 provides details on how this would work in the remittance context. The noteworthy accounting trick to adopt from Section 902 is that the recipient taxpayer must “gross up” the repatriated remittance by the amount of foreign taxes paid, in order to avoid a double deduction. Where a worker has not paid taxes in the destination country, the remitted funds would simply be taxed at the full home-country rate.

238. I.R.C. §§ 901, 902 (West 2010); Kirsch, supra note 152, at 504–05 (explaining how the United States’ foreign tax credit addresses double taxation concerns for U.S. citizens residing abroad).

239. The reasoning behind a reduced remittance tax rate would be akin to that used to support the reduced capital gains tax rate in the United States. See, e.g., Chris Edwards, Six Reasons to Keep the Capital Gains Tax Rate Low, CATO INST. (Dec. 2012), http://www.cato.org/publications/commentary/six-reasons-keep-capital-gains-tax-rates-low (arguing that the capital gains tax rate should remain low in order to mitigate burdens associated with double taxation of corporate profit at the corporate and shareholder dividend level).

240. The United States provides such a foreign tax credit to individuals and corporations. See I.R.C. §§ 901, 902 (West 2010). The goal of such a tax credit is not to eliminate U.S. taxation of the income entirely, but to allow the United States to impose its domestic tax rate without overburdening taxpayers that have already paid some amount of tax abroad. As a rough example, where the United States levies a twenty percent tax, and a U.S. taxpayer paid a fifteen percent tax abroad, the foreign tax credit ensures that the U.S. taxpayer will only be subject to an additional five percent tax on that income. Id.


242. Id.
The third option to mitigate double taxation is to apply a reduced rate to remittance inflows. In this case, the recipient country could impose a reduced tax rate designed to capture the gap between the home country tax rate and the likely tax rate imposed in a common migrant host country. In the example in Box 1, the country would apply a five percent tax rate, as that is the difference between the average host country tax rate of twenty percent and the home country rate of twenty-five percent. Applying a reduced tax rate would raise comparable revenue to the tax credit method with much less accounting complexity. However, the reduction in complexity carries a risk of burdensome taxation if the host country already imposes a high rate of income tax.

Developing countries would need to decide whether to impose remittance taxation via self-reporting or withholding at the source, in this case through the transfer company. Source withholding would be more administratively feasible if the country chooses to utilize the reduced tax rate rather than a tax credit. However, the reduced rate method carries certain significant drawbacks. Namely, it risks becoming a de facto flat tax on remittance transfers, which would eliminate the progressivity that makes income taxation of remittance inflows more attractive than other methods of taxation. Remittance recipients who fall below the income tax threshold would theoretically be able to apply for a rebate of withheld taxes, which would mitigate this danger somewhat. However, absent some additional benefit or financial incentive, such as the tax deduction senders receive under Policy 1, the additional burden of applying for a tax rebate may drive remittance flows to informal channels to avoid the additional tax.

C. Policy 3: Indirect Remittance Taxation

Where bilateral treaties, source withholding, and income taxation are simply impossible, countries still need not ignore remittance flows as potential sources of revenue. In these cases, policymakers in remittance-receiving nations would do well to consider how best to capture remittance inflows indirectly. While consumption taxation may be insufficient to capture the maximum
possible gains from remittance inflows, migrant-source countries seeking to harness inflows indirectly can at least improve their revenue outcomes by plugging holes in consumption taxation and strengthening (or instituting) property taxation. Although this indirect taxation would not raise as much revenue as the above-described income tax policies, it would nonetheless be an improvement on the current tax policies that fail to specifically target remittance inflows. These indirect remittance tax reforms could also be implemented alongside the above-described policies, further maximizing the public revenue gains from remittance inflows.

1. Property Taxation

Recipient country policymakers should consider capturing gains through remittance-conscious property taxation. Remittance recipients often invest transferred funds in property through purchasing real estate, building new structures, or improving existing structures. In fact, the frequency with which remitted funds are used on real estate is often cited as a reason why remittances fail to contribute to long-term economic growth—compared with, for example, if the transfers were more often invested in business enterprises. Despite this practice, as explained above, the unpopularity of the property tax means it is infrequently used in developing countries. Migrant-source countries or localities that do not utilize property taxation should implement such a tax in order to capture the remittance gains that are currently slipping through tax cracks. Those that already do utilize property taxation should ensure that the tax captures not only purchases but improvements as well, since migrant households often save remittances through real estate improvements.

This kind of policy could be implemented at either the central or the local level. Central governments seeking to harness remittance transfers may wish to define the policy centrally, then work with localities to ensure a properly targeted remittance taxation policy. For example, central governments can ascertain regions and villages with a large proportion of migrant households and offer technical assistance and other incentives to encourage those local governments to institute property taxation. Likewise, a local government whose

243. See discussion supra Part III(B)(i).
244. See, e.g., KRISHNASWAMI ET AL., supra note 62, at 13 (finding that thirty-six percent of remitters in New Haven reported that remitted funds were used for home maintenance); Debra Roberts, Presentation at the 8th Annual Conference of the Sir Arthur Lewis Institute of Social and Economic Studies, The Developmental Impact of Remittances on Caribbean Economies: The Case of Guyana 13 (Mar. 26–28, 2007) (finding that sixteen percent of surveyed Guyanans saved remitted funds in the form of real estate).
245. See, e.g., Ralph Chami & Connel Fullenkamp, Beyond the Household, FIN. & DEV., Sept. 2013, at 48, 50, http://www.imf.org/external/pubs/ft/fandd/2013/09/Chami.htm (explaining that the common use of remittances to purchase assets such as real estate does not increase capital stocks, and thus is counterproductive to growth enhancing capital accumulation).
246. See supra Part III(A)(iii).
region hosts a sizable population of remittance recipients could unilaterally undertake special efforts to ensure an effective property tax system.248

Wielding the property tax in this way has several possible upshots. One advantage is that it would not reduce incentives to remit funds or drive remittances underground. Additionally, because the property tax is often administered at the local level,249 tax revenue would more likely be used for local programs rather than national programs. This can be a positive or a negative, depending on the capacity and quality of local institutions compared with national institutions. Because the property tax would only capture revenue from remittance gains invested in property, expansion of the property tax should be coupled with improved consumption taxation as well. As explained in the following section, these two policies together would cast a wider net, ensuring more effective capture of remittance gains.

2. Consumption Taxation

As mentioned above, consumption taxation is already a mainstay of tax policy throughout the developing world.250 Thus, for developing countries aiming to utilize consumption taxation to capture gains from remittance transfers, the trick will be ensuring that the tax specifically captures remittance-related gains. Given the proliferation of consumption taxation in developing countries, certainly some amount of remittance capture is already happening. Indeed, Christian Ebeke finds evidence that remittances significantly increase both the level and stability of government tax revenue in remittance-receiving countries that have adopted the VAT.251

The primary way that policymakers can ensure that their country’s VAT captures remittance gains is to implement the tax on a broader base, encompassing the types of items on which remittances are commonly spent. For

---

248. The issue of what an effective property tax entails is quite another question that requires more detail and attention than this Article can provide here. For an excellent discussion of property taxation in developing countries from varying perspectives, see MAKING THE PROPERTY TAX WORK: EXPERIENCES IN DEVELOPING AND TRANSITIONAL COUNTRIES (Roy Bahl, Jorge Martinez-Vazquez & Joan Youngman eds., 2008) See also John Norregaard, Taxing Immovable Property: Revenue Potential and Implementation Challenges (IMF Fiscal Affairs Dep’t, Working Paper No. WP/13/129, 2013).

249. See Bird & Slack, supra note 183, at 103–04.

250. See Keen & Lockwood, supra note 168, at 138 (describing the VAT as the centerpiece of many developing country tax reform efforts); see also Vito Tanzi & Howell H. Zee, Tax Policy for Emerging Markets: Developing Countries 21 (IMF Fiscal Affairs Dep’t, Working Paper, No. WP/00/35, 2000) (although the “overwhelming majority of developing countries” have adopted a VAT, the VAT systems are largely incomplete for exempting major sectors of the economy like wholesale, retail, and services).

example, because a high proportion of remitted income is spent on food items and other basic consumption goods.\textsuperscript{252} Food consumption should not be exempted from taxation across the board.\textsuperscript{253} Vietnam provides a useful model for a VAT that encompasses food, while still offering a reduced five percent rate on basic consumption goods.\textsuperscript{254} Housing expenditures should likewise not be exempted from taxation, for the reasons explained in the previous section.\textsuperscript{255}

One difficulty inherent in capturing remittance gains via an indirect consumption tax is that the link between remittance gains and tax mobilization will be unclear. Where remittances are not directly taxed, and where a government benefits indirectly from windfall gains, the risk that the gains will generate government corruption increases.\textsuperscript{256} To mitigate this concern, policymakers could consider linking tax collections directly with remittances by offering sales tax credits or refunds to non-migrant households. However, this may raise legitimate equity concerns and perhaps engender resentment among remittance-receiving taxpayers. It would also reduce incentives to remitting and entail significant enforcement challenges as remitters turn to informal channels.

In all, indirect remittance taxation provides a less effective and less transparent method for harnessing remittance gains for public benefit. Both income taxation policies described above would raise more revenue, more efficiently, and with greater transparency—albeit at likely higher administrative cost. Despite the positives of a remittance income tax regime, developing country policymakers may be unable to negotiate bilateral treaties or institute a complex tax credit policy. This third policy option is more of a gap-filling measure than an actual cohesive policy, intended to highlight that policymakers may still account for remittances when considering tax policy reforms, despite practical limitations.

V. ADMINISTRATIVE CONSIDERATIONS

Tax reform is a tricky business, particularly when a government is aiming to tax a much-celebrated financial phenomenon such as remittances. This section discusses additional administrative difficulties associated with the above-described policies.

\textsuperscript{252} See IFAD, SENDING MONEY HOME, supra note 60, at 7.
\textsuperscript{253} See supra Table 4; LIAM P. EBRILL ET AL., THE MODERN VAT 84 (2001) (listing key nonstandard VAT exemptions in selected countries including flour, milk, agricultural products, rice, onions, etc.).
\textsuperscript{254} See supra Table 4.
\textsuperscript{255} See, e.g., Roberts, supra note 244, and accompanying text; KRISHNASWAMI ET AL., supra note 62, at 13, and accompanying text.
\textsuperscript{256} See Abidh et al., supra note 75, at 657 (explaining that, where government officials benefit from remittance transfers through mechanisms other than direct taxation of flows, such as through indirect consumption taxation, corruption is more likely).
A. Distorting Remittance Behavior

As mentioned above, opponents of remittance taxation cite fears that increasing transfer costs will reduce flows overall.257 If true, such fears would counsel policymakers to utilize less visible and less burdensome tax methods, such as the aggregate national transfer option under Policy 1 or indirect taxation of remittances through property and consumption taxes. Withholding taxes at transfer sites poses particular risks of distorting remittance behavior. However, the risk of reduced flow may be overstated. There is some evidence that remittance transfers are relatively cost inelastic, as workers send them for cost-independent reasons.258 Dean Yang even finds evidence that a decrease in remittance costs in some cases reduces the total amount remitted, as senders find it easier to reach the specific target intended for families back home.259 The inverse may be true of an increase in cost engendered by taxation—the cost increase might actually drive senders to remit more overall. Thus, a tax on remitted funds might not actually decrease total transfers.

Another concern is that senders will instead choose to remit funds through informal channels in an effort to avoid withholding or third-party verification. Driving remittance flows underground is another primary argument cited by opponents of remittance taxation.260 Although estimates suggest that informal remittances are significant—anywhere between thirty-five percent and 250% of formal remittance flows261—the exact relationship between remittance cost and the incidence of informal flows is difficult to measure. Further, cost is just one factor for senders deciding whether to use formal or informal channels.262 Other considerations include security, speed, accessibility for the sender and recipient, and convenience in terms of familiarity and language.263 Thus, where policymakers fear driving remittances underground, they should work to increase the attractiveness of formal channels by enhancing speed, security, and accessibility of formal providers. They should also work to reduce the private costs of sending remittances, for example, by reducing unnecessary regulatory red tape and encouraging competition among remittance transfer providers.264

257. See supra Part I(B)(iii).
259. Yang, supra note 74, at 16 (finding that Philippine migrant workers remitted less money, measured in the foreign currency, in response to a favorable shift in the home country’s exchange rate).
260. Mohapatra, supra note 3; supra Part I(B)(iii).
261. See Freund & Spatafora, supra note 258, at 1.
262. Id. at 4.
263. Id.
264. See Lenora Suki, Competition and Remittances in Latin America: Lower Prices and More
B. Political Will

Low-income countries collect relatively little tax revenue through income and property taxation in part because of a lack of political will to tax income and property.\textsuperscript{265} Yet there is evidence that willingness to tax depends in part on perceptions of fairness.\textsuperscript{266} Thus, a tax that is levied to promote fairness and progressivity may encounter less political resistance. Specifically, the public and policymakers may support a remittance tax in a country with significant human capital flight and where remittance recipients tend to belong to high-income groups. In these contexts, such a tax may be considered a response to an important domestic need.

Conversely, remittance tax policies could face resistance where they disproportionately burden elites in developing countries, as these citizens will have more power to block the reforms. However, taxing remittance payments via income taxes will affect a small sub-group of the population, which may limit the risk of political resistance among domestic elites. In all, while political feasibility is certainly a practical consideration, it need not be an insurmountable barrier where migration compensation is an important domestic policy goal.

C. Evasion

Evasion is always a serious concern when considering tax administration. Self-reported taxes, such as income taxes, suffer from greater risk of evasion compared to property and consumption taxes, which are more difficult to avoid. Third-party reporting by remittance transfer companies, as described above, should help reduce evasion by alerting tax officials to income received by migrant households.\textsuperscript{267} Expanding the income tax to include remittance payments may even reduce income tax evasion more broadly by bringing additional households to the attention of domestic tax agencies. This expansion of the tax base could result in revenue increases beyond merely the transfers themselves.

Evasion may still occur through senders using informal transfer services. Policymakers can reduce this risk by ensuring the competitiveness and attractiveness of formal transfer methods, as described above.\textsuperscript{268} Additionally,

\textit{Efficient Markets} 5 (Org. for Econ. Cooperation and Dev., Working Paper No. 2/2007, 2007) ("Increasing competition in remittances markets has been identified as a means of lowering transaction costs and improving the efficiency of the market.").

\textsuperscript{265} See Bahl et al., \textit{supra} note 177, at 3–4; Bird & Zolt, \textit{supra} note 36, at 1670–71 ("[L]ow levels of taxation are often interpreted as the result of the unwillingness of the richest to pay taxes to provide public services for the masses because the elite can generally provide their own public services privately . . . .").

\textsuperscript{266} See Richard M. Bird et al., \textit{Societal Institutions and Tax Effort in Developing Countries} 29 (Joseph L. Rotman Sch. of Mgmt., Univ. of Toronto, Int’l Tax Program Paper No. 04011, 2004); Bird & Zolt, \textit{supra} note 36, at 1670–71.

\textsuperscript{267} See discussion \textit{supra} Part IV(A)(ii).

\textsuperscript{268} See discussion \textit{supra} Part V(A).
studies show that tax evasion rates are significantly correlated with notions of tax fairness and government responsiveness. For this reason, policymakers should ensure that any remittance tax policy is perceived as fair and transparent, and that mobilized revenue is used to provide needed public goods and services. Methods for improving the tax policy’s fairness and transparency are discussed in the next section.

D. Corruption

Once the tax is implemented and collected, an optimal policy should ensure that the funds are utilized for beneficial, pro-growth public goods. There are several possible ways to defend tax revenues from corruption and misuse, largely through increased monitoring of agencies and government staff. For example, one method to ensure the beneficial use of remittance tax revenues is to earmark them for specific uses. Environmental taxation provides an example of issue-specific earmarking, wherein local governments administer an environmental tax and are then required to spend the garnered revenue on environmental protection. Remittance taxation could operate similarly. For example, governments could spend remittance tax revenue on scholarships that incentivize working domestically or to improve the fiscal environment to encourage investment by emigrants.

There are various other oversight mechanisms used by developing country governments to combat misuse of public funds and to ensure accountable, holistic fiscal policies. In Turkey, for example, the Minister of Finance is required to seek input on tax and expenditure legislation from local authorities, universities, trade unions, public professional organizations, and civil society organizations. Peru and Argentina also require public oversight of budgeting agencies and have enacted limits on their legislatures’ ability to introduce new expenditures. Participatory budgeting is another option for holding government accountable for expenditures. The process entails citizens meeting in public assemblies to discuss the government budget. This process has been tried in several cities including Buenos Aires, Montevideo, Caracas,

271. Id. at 32.
272. See Desai et al., supra note 1, at 685.
273. Leyla Ates, Domestic Political Legitimacy of Tax Reform in Developing Countries: A Case Study of Turkey, 30 WIS. Int’l L.J. 706, 752 (2012).
274. Lledo et al., supra note 270, at 34.
275. Id.
276. Id.
San Denis, and Cordoba.\textsuperscript{277} Finally, Bird et al. suggest that decentralization may help increase accountability,\textsuperscript{278} which may counsel policymakers towards use of the property tax or local consumption taxation where misfeasance is a significant concern.

The main takeaway from this brief survey of administrative feasibility is that implementation difficulties need not overshadow the potential benefits of remittance taxation.

**Conclusion**

When crafting migration-related economic policies, developing country policymakers face an inherent policy tension in balancing individual welfare with broader domestic economic goals. Past proposals to compensate developing countries for migration losses often focus on maximizing the latter at the extent of the former, largely through discouraging migration or burdening emigrants living abroad. However, an ideal migration compensation policy should seek to regain lost tax revenue without distorting migration decision-making. Taxing remittance payments to recipient households is one way to solve this policy dilemma by compensating home economies for lost revenue without overburdening migrant workers abroad.

Migration and development advocates have remained staunchly opposed to taxation of remittances by recipient countries. This view ignores the fact, however, that remitted income is already taxed in high-income host countries via labor taxation of migrant workers. In contrast to the current arrangement, both normative and practical considerations support shifting the locus of taxation from high-income host countries to low-income recipient countries. This kind of tax arrangement would better comport with first principles of family taxation and inter-nation equity concerns and better serve domestic tax goals of home country governments. This Article has provided an ideal structure for such a remittance taxation policy, involving bilateral treaties between home and host countries, as well as alternative tax policies in the event that bilateral cooperation is infeasible.

The purpose here is not necessarily to advocate for one policy over another, but rather to highlight that migrant-source countries may not be well-served by the blanket stance against remittance taxation that has been adopted by the economic development community. Migrant-source countries face special constraints in constructing their tax systems, as they are forced to cede a significant portion of their tax base to higher-income host countries. Under the current regime host countries benefit both from cheaper labor and additional tax revenue, while the home country government suffers a net loss. For resource-poor governments, this loss can perpetuate their inability to provide necessary public goods and maintain reliable domestic institutions.

\textsuperscript{277} Id.

\textsuperscript{278} Bird et al., supra note 266, at 15.
Low-income home country governments have both theoretical and practical arguments why they should not be forced to surrender remittance tax revenue to host countries. Although this battle is one that ultimately must be fought in political and diplomatic arenas, scholarship on the topic undoubtedly plays a role in shaping public perception. Scholars in the field who dogmatically oppose remittance taxation do a disservice to migrant-source countries by taking off the table a valid and feasible source of revenue. Adding remittance taxation to the developing country tax arsenal would better equip these nations to capture migration gains, and in so doing, to improve development outcomes and aggregate wellbeing. Improving wellbeing, after all, is the ultimate goal.
A Comity of Errors: The Rise, Fall, and Return of International Comity in Transnational Discovery

Diego Zambrano*

ABSTRACT

No feature of U.S. law has rankled foreign nations more than the supposed “legal imperialism” of discovery requests for information located abroad to be used in U.S. litigation or investigations. China, France, Germany, and Switzerland have threatened the stability of bilateral relations with the United States due to overbroad transnational discovery requests. For three decades, when faced with concerns of international comity in the discovery context, U.S. courts ruled overwhelmingly in favor of discovery through the Federal Rules, rendering international comity a dead concept.

Recent case law, however, shows that this paradigm is coming to an end. In a trilogy of cases decided, respectively, by the United States Supreme Court (Daimler), the Second Circuit (Gucci), and the New York State Court of Appeals (Motorola), each court rejected attempts by plaintiffs to subject foreign entities to jurisdiction in the United States or otherwise impose on them overbroad duties, including those in conflict with foreign laws. Prominently relying on “international comity,” each decision limited the reach of U.S. courts and emphasized the need for harmony in the international legal system. These three cases are groundbreaking and should lead to changes in U.S. transnational discovery.

The Article analyzes this recent revival of international comity. First, it explores the history of international comity and its interaction with broad U.S. discovery rules. Second, it briefly reviews the Supreme Court case Aérospatiale, which dealt a blow to international comity. Third, this Article analyzes how

DOI: http://dx.doi.org/10.15779/Z384K2P

* Associate at Cleary Gottlieb Steen & Hamilton in New York. J.D. 2013, Harvard Law School; B.A. 2010, University of Virginia. I would like to thank Carmine Boccuzzi for his mentorship and support for this article. I am also grateful to Sophie Kaiser, Corey Banks, Hilary Wong, Rusty Feldman, Adam Olin, Katherine Denby, Eugene Karlik, James Blakemore, Gregory Wolfe, Joseph Brothers, and the members of BJIL for helpful comments and conversations. The views expressed herein are my own and do not necessarily reflect the views of Cleary Gottlieb.
Daimler, Gucci, and Motorola relied on comity to reach their holdings and argues that international comity has been revived in the context of discovery. Finally, this Article takes a normative approach and argues that U.S. courts should engage in a qualitative limitation of the kinds of U.S. interests that are significant in the transnational discovery context.

TABLE OF CONTENTS

Abstract .................................................................................................................. 157
Introduction ........................................................................................................... 159
I. The History of International Comity & Discovery .................................. 161
   A. What is International Comity? ............................................................... 161
   C. Conflicts Between U.S. Discovery and Foreign Laws .................. 167
   D. The Hague Convention .......................................................................... 171
II. Comity in Retreat: Aérospatiale and the Defeat of the Hague Convention .............................................................. 173
III. The Return of International Comity: Daimler, Gucci, and Motorola
   Establish a new Paradigm ........................................................................... 179
      A. Daimler: The Supreme Court Revives International Comity .......... 181
      B. Gucci: The Second Circuit Takes International Comity a Step Further ........................................................................................................ 184
      C. Motorola: The New International Comity Paradigm is
         Established ............................................................................................ 188
IV. Economics and Diplomacy: The Main Factors Behind Daimler, Gucci, and Motorola ................................................................. 193
      1. Reciprocity, Retaliation, and Foreign Relations ....................... 194
      2. The International Economy ................................................................. 195
V. How Daimler Should Reshape Comity: Imposing Limits on U.S.
   Interests and Requiring Input from Foreign Sovereigns .................... 198
      A. The Problems with the Categorical Approach that Overvalued
         U.S. Interests ........................................................................................... 200
      B. Daimler Demands a More Individualized Analysis of U.S.
         Interests .................................................................................................. 207
      C. The Need for Input from Foreign Countries .................................. 211
Conclusion ............................................................................................................. 214
INTRODUCTION

It has been widely noted that no feature of U.S. law has rankled foreign nations more than discovery requests for information located abroad to be used in U.S. litigation or investigations. For example, Chinese regulators recently threatened the stability of bilateral relations with the United States due to overbroad transnational discovery requests against the Bank of China stemming from litigation in New York. France and Switzerland have enacted statutes criminalizing the production of documents to U.S. authorities, and Germany has called it “an intrusion into its sovereignty.” Not only do U.S. discovery procedures affect international relations, they also deter foreign companies from doing business in the United States due to fears of overbroad jurisdiction assertions. Precisely for this reason, the New York Court of Appeals recently intimated that New York’s place as the commercial and financial center of the world is endangered by uninhibited personal jurisdiction and the judiciary’s overbroad exercise of extraterritorial power over foreign matters and parties.


2. See Letter from Huai Peng Mu, Director-General of the Legal Affairs Department of the People’s Bank of China, and Yi Huang, the Director-General of the Supervisory Rules and Regulations Department of the China Banking Regulatory Commission, to Catherine O’Hagan Wolfe, Clerk of Court, United States Court of Appeals for the Second Circuit (Dec. 19, 2013) (writing with respect to the pending appeals in Gucci America, Inc. v. Bank of China, Nos. 11-3934(L)).


4. See Brief for Federal Republic of Germany as Amicus Curiae Supporting Appellants at 1, In re Auto. Refinishing Paint Antitrust Litig., 358 F.3d 288 (3rd Cir. 2004) (No. 02-4272), 2003 WL 2413699. In many instances, foreign nations are interested in resolving these disputes domestically, rather than allowing U.S. courts to sanction their corporations or order turnover of local citizens’ accounts.


Without a doubt, the repercussions of developments in this area of law are increasingly important in the modern global economy.\(^7\)

In the United States, discovery is a routine procedural issue that courts, armed with broad jurisdiction and subpoena powers, are well equipped to supervise.\(^8\) However, when a lawsuit involves foreign parties and documents located in foreign nations, discovery can generate complex and difficult conflicts between U.S. procedures and foreign laws.\(^9\) This kind of transnational discovery has seen much activity recently because foreign corporations with affiliates in the United States are faced with an increasing barrage of lawsuits, subpoenas, and turnover actions from litigants seeking judgment in U.S. courts. Confronted with these conflicts between U.S. discovery rules and foreign laws, courts seek to promote international harmony by giving deference to the sovereign interests of the affected nations, a principle called “international comity.”

This Article identifies and explains a recent trend in U.S. case law towards renewed respect for international comity and foreign laws in the particular context of transnational discovery. In an era of austere U.S. foreign and domestic policy, courts are following the executive’s lead in refurbishing their international comity bona fides when faced with overbroad discovery requests. This judicial development is of particular importance for foreign relations and the global economy because it will alter the operations of thousands of multinationals, the international trade system, and data protection laws.

For three decades, when faced with concerns of international comity and discovery requests, U.S. courts applied a balancing test to weigh the interests of foreign countries against U.S. interests, and ruled almost unanimously in favor of U.S. interests and the judiciary’s power to reach foreign documents or assets.\(^10\) Due to the unjustified emphasis on U.S. interests in patent laws, antitrust laws, criminal laws, and other broad categories, foreign defendants could hardly use international comity as a shield. Instead, comity became a frivolous argument raised by foreign litigants as a last, and ultimately unsuccessful, resort.

---


8. Société Nationale Industrielle Aérospatiale v. U.S. Dist. Court for S. Dist. of Iowa, 482 U.S. 522, 551 (1987) (Blackmun, J., concurring in part and dissenting in part) (“The discovery process usually concerns discrete interests that a court is well equipped to accommodate—the interests of the parties before the court coupled with the interest of the judicial system in resolving the conflict on the basis of the best available information. When a lawsuit requires discovery of materials located in a foreign nation, however, foreign legal systems and foreign interests are implicated as well.”).

9. See id.

10. See infra notes 253–62.
But, as this Article will show, this thirty-year paradigm seems to be coming to an end. In a trilogy of recent cases decided, respectively, by the United States Supreme Court (Daimler AG v. Bauman), the Second Circuit Court of Appeals (Gucci America, Inc. v. Weixing Li), and the New York State Court of Appeals (Motorola Credit Corp. v. Standard Chartered Bank), each court prominently relied on “international comity” in refusing to subject foreign entities to jurisdiction in the United States or otherwise impose on them overbroad duties, thereby limiting the reach of U.S. courts. In the context of Motorola and Gucci, the courts protected non-party foreign banks from discovery or turnover of documents and funds located abroad. These three cases are groundbreaking. They may significantly affect the development of transnational discovery and strengthen alternative avenues to such discovery.

This Article analyzes this recent revival of international comity. First, it explores the recent history of international comity and its interaction with broad U.S. discovery rules. Second, it reviews the Supreme Court case Société Nationale Industrielle Aérospatiale v. United States District Court for the Southern District of Iowa, which dealt a blow to international comity. Third, this Article analyzes how Daimler, Gucci, and Motorola relied on comity to reach their holdings and argues that international comity has been revived in the context of discovery. Finally, this Article takes a normative approach and argues that U.S. courts should engage in a qualitative limitation on the kinds of U.S. interests that are significant in the transnational discovery context.

I.
THE HISTORY OF INTERNATIONAL COMITY & DISCOVERY

A. What is International Comity?
At its simplest, international comity is the concept of judicial respect for the sovereignty of foreign nations. Courts have long recognized that international comity “is neither a matter of absolute obligation . . . nor of mere courtesy and good will.” Instead, comity involves “the recognition which one nation allows within its territory to the legislative, executive or judicial acts of another nation, having due regard both to international duty and convenience, and to the rights of its own citizens, or of other persons who are under the protection of its laws.” Thus, even where a court is within its powers to hear a

12. Id.; see also SEC v. Banner Fund. Int'l, 211 F.3d 602, 612 (D.C. Cir. 2000) (noting that comity in the context of discovery means “the degree of deference that a domestic forum must pay to the act of a foreign government not otherwise binding on the forum.”); Hessel E. Yntema, The Comity Doctrine, 65 Mich. L. Rev. 9 (1966) (indicating that the concept of comity was developed in the late 17th century).
13. Hilton, 159 U.S. at 164. See also Van Den Biggelaar v. Wagner, 978 F. Supp. 848, 857 (N.D. Ind. 1997) (“In the United States, the comity concept was imported by Joseph Story but later modified into a discretionary principle with an ambiguous status between law and policy.”).
case or force a foreign corporation to comply with an order, international comity compels courts to consider the interests of foreign nations in the dispute.

Comity could be considered the judicial way of conducting diplomacy. 14 Although not a political branch, the judiciary is often involved in issues of great international consequence.15 Whether addressing treaties, foreign wars, historical claims, or other important global issues, U.S. courts at times act in the name of the country, and as such must consider the repercussions of their decisions on foreign relations.16 Justice Breyer recently affirmed the growing need for “coordination with other jurisdictions . . . for the smooth functioning of our economy and our various institutions.”17 Because of this need for coordination, Justice Blackmun once noted that “[c]omity is not just a vague political concern favoring international cooperation when it is in our interest to do so. Rather it is a principle under which judicial decisions reflect the systemic value of reciprocal tolerance and goodwill.”18

Commentators typically marshal four major arguments in support of the continued existence of international comity in the civil context: (1) the danger of double liability that a person or corporation may face at home and abroad when there are conflicting laws; (2) the promotion of international commerce; (3) the high burden and cost of requiring a foreign party to appear in front of U.S. courts; and (4) the interest of U.S. courts in having their rulings recognized abroad.19 These arguments can only become more pertinent in the face of globalization, where modern corporations have branches and affiliates in dozens of countries.20 Ultimately, as described by the Second Circuit Court of Appeals,
“international comity is clearly concerned with maintaining amicable working relationships between nations, a shorthand for good neighbourliness, common courtesy and mutual respect between those who labour in adjoining judicial vineyards.”

Before undertaking a historical analysis of international comity, two points of clarification are in order. First, this Article deals with a particular type of international comity analysis that arises in the civil law context of transnational discovery. Although international comity plays an important role in criminal, bankruptcy, tax, antitrust, and other areas of law,22 Daimler and its progeny have only addressed international comity in the context of civil lawsuits. Second, this Article will rely, partially but not entirely, on cases in the Second Circuit and Southern District of New York. As the financial capital of the world, New York is the nerve center for multinational corporations and banks with branches in the United States.24 Because of their status as garnishees, banks are popular targets for transnational discovery and turnover requests.25 Therefore, decisions in the Second Circuit, and even the New York State Court of Appeals, have an outsized influence on transnational discovery and international comity.


Permissive discovery rules have characterized U.S. federal courts since 1938, when the Federal Rules of Civil Procedure were adopted. At the urging of Roscoe Pound, American procedural reforms beginning in 1906 culminated with the adoption of rules that allowed “increased relaxation and expansion of procedure.” Federal Rule of Civil Procedure 34 provided parties with the power to inspect documents and things “material to any matter involved in the action.” The rule allowed parties to “examine” any person who might have assets belonging to the defendant or, in post-judgment actions, the judgment debtor. In 1948, an amendment to the Federal Rules expanded the scope of discovery to the more permissive language of Rule 26, allowing the court to order the production of documents “relating to any of the matters within the scope of the examination permitted by Rule 26(b) and which are in his possession, custody, or control.” At that time, the only vehicle available for courts to request help from a foreign government or party was the Letter Rogatory—a formal request for discovery assistance.

By the 1960s, courts had determined they had the power to order the production of documents located abroad. This conclusion came as the logical consequence of an expanding personal jurisdiction and discovery jurisprudence. It is axiomatic that without personal jurisdiction a court cannot order a party to produce documents because it has no power over that party. However, once a court finds it has personal jurisdiction, there are few limits on what it can order a party to produce. Further, courts concluded that possession, custody, or control over the documents or assets being sought is necessary, because without it a party has no practical ability to obtain the documents and thus cannot be required to do so. These two ingredients became what can be called the

27. Roscoe Pound was one of the most influential legal figures of the twentieth century. As Dean of Harvard Law School, he was a prolific scholar and noted legal realist. Pound was a towering legal figure at a crucial time for American law.
29. History of Rule, 8B FED. PRAC. & PROC. CIV. § 2201 (3d ed.).
30. Discovery in Aid of Execution, 12 FED. PRAC. & PROC. CIV. § 3014 (3d ed.).
31. 8B FED. PRAC. & PROC. CIV. § 2201, supra note 29.
32. 5 F.R.D. 433, 463 (1946).
33. 22 C.F.R. § 92.54.
34. United States v. First Nat’l City Bank, 396 F.2d 897, 900–01 (2d Cir. 1968).
35. Id.; Gucci Am., Inc. v. Weixing Li, 768 F.3d 122, 141 (2d Cir. 2014) (“A district court, however, must have personal jurisdiction over a nonparty in order to compel it to comply with a valid discovery request under Federal Rule of Civil Procedure 45.”); In re Sealed Case, 141 F.3d 337, 341 (D.C. Cir. 1998) (finding it “elementary” that “courts lacking jurisdiction over litigants cannot adjudicate their rights”); RESTATEMENT (THIRD) OF FOREIGN RELATIONS LAW § 442 (1987) (“A court . . . may order a person subject to its jurisdiction to produce documents”).
36. In re Warrant to Search a Certain E-Mail Account Controlled & Maintained by Microsoft Corp., 15 F. Supp. 3d 466, 472 (S.D.N.Y 2014) (“It has long been the law that a subpoena

Published by Berkeley Law Scholarship Repository, 2016
“personal jurisdiction plus control” test that underlies the discovery of documents located abroad: U.S. courts need only find personal jurisdiction over the party and possession, custody, or control of the requested documents by the target of the subpoena.\(^{37}\) Because of these simple requirements, and the permissive nature of the Federal Rules, it is typical for judgment creditors to demand transnational asset discovery from parties and non-parties alike—and courts usually oblige.\(^{38}\) This reach extends to non-party banks that may have information about a debtor’s assets.\(^{39}\)

To complement the “personal jurisdiction plus control” test, courts recognized early on the importance of balancing foreign interests when foreign laws or parties were involved. It was this recognition that created judicial concerns with what courts began to call “international comity.” These concerns came into play, however, only when there was a “true conflict” between domestic and foreign law.\(^{40}\) The Supreme Court has recognized this initial inquiry:

The threshold question in a comity analysis is whether there is in fact a true conflict between domestic and foreign law. When there is a conflict, a court should seek a reasonable accommodation that reconciles the central concerns of both sets of laws. In doing so, it should perform a tripartite analysis that considers the foreign interests, the interests of the United States, and the mutual interests of all nations in a smoothly functioning international legal regime.\(^{41}\)

---


38. See, e.g., EM Ltd. v. Republic of Arg., 695 F.3d 201, 207–08 (2d Cir. 2012), aff’d sub nom. Republic of Arg. v. NML Capital, Ltd., 134 S. Ct. 2250, 2254 (2014). But see Ings v. Ferguson, 282 F.2d 149, 152 (2d Cir. 1960) (refusing to compel production of documents located in Canada because, among other reasons, “[u]pon fundamental principles of international comity, our courts dedicated to the enforcement of our laws should not take such action as may cause a violation of the laws of a friendly neighbor or, at the least, an unnecessary circumvention of its procedures. Whether removal of records from Canada is prohibited is a question of Canadian law and is best resolved by Canadian courts.”).

39. “It is not uncommon to seek asset discovery from third parties, including banks, that possess information pertaining to the judgment debtor’s assets.” EM Ltd., 695 F.3d at 207. “[I]n a run-of-the-mill execution proceeding, we have no doubt that the district court would have been within its discretion to order the discovery from third-party banks about the judgment debtor’s assets located outside the United States.” Id. at 208; see also Eitzen Bulk A/S v. Bank of India, 827 F. Supp. 2d 234, 238–39 (S.D.N.Y. 2011).


41. Id. This Article will only deal with cases where there is a “true conflict.” Cases falling before this threshold question present comity concerns but are generally less relevant because they do not involve an analysis of conflicting laws.
Recognizing the problems presented by these instances of “true conflict,” the 1965 Restatement Second of Foreign Relations Law highlighted that conflict with foreign laws did not deprive a U.S. court of jurisdiction but nonetheless required balancing the interests of the relevant sovereigns. The balancing test announced by the Restatement Second of Foreign Relations urged courts to weigh five factors: (1) the “vital national interests of each of the states,” (2) the “hardship” imposed on the person, (3) the “extent” to which “required conduct is to take place” in the foreign country, (4) the nationality of the person, and (5) the extent to which enforcement can “be expected to achieve compliance.”

Although courts began to weigh these interests in the 1960s, they continued to routinely exercise their jurisdiction over documents located abroad, finding that U.S. interests generally prevailed over foreign interests.

Exemplifying the three-part test that had developed by the late 1960s—(1) personal jurisdiction, (2) control, and (3) balancing of foreign interests—the Second Circuit in 1968 compelled a bank to produce documents located in its branch in Frankfurt, Germany, noting:

The basic legal question confronting us is not a total stranger to this Court. With the growing interdependence of world trade and the increased mobility of persons and companies, the need arises not infrequently, whether related to civil or criminal proceedings, for the production of evidence located in foreign jurisdictions. It is no longer open to doubt that a federal court has the power to require the production of documents located in foreign countries if the court has in personam jurisdiction of the person in possession or control of the material. The court noted that difficulties arose where “the country in which the documents are located has its own rules and policies dealing with the production and disclosure of business information—a circumstance not uncommon.” Recognizing that it involved an “extremely sensitive and delicate area of foreign affairs,” the court noted that a rule that ignores foreign laws except when a party shows it will suffer criminal liability would “show scant respect for international comity.” Despite this, the court analyzed the different interests and upheld the district court’s decision holding the bank in contempt for failure to produce the documents.

42. See generally Restatement (Second) of Foreign Relations Law of the United States § 39(1) (1965) (“A state having jurisdiction to prescribe or to enforce a rule of law is not precluded from exercising its jurisdiction solely because such exercise requires a person to engage in conduct subjecting him to liability under the law of another state having jurisdiction with respect to that conduct.”).

43. Id. § 40.

44. See, e.g., Société Internationale Pour Participations Industrielles et Commerciales, S.A. v. Rogers, 357 U.S. 197 (1958); Application of Chase Manhattan Bank, 297 F.2d 611, 613 (2d Cir. 1962); Ings v. Ferguson, 282 F.2d 149, 152 (2d Cir. 1960); First Nat’l City Bank of N.Y. v. IRS, 271 F.2d 616 (2d Cir. 1959).

45. United States v. First Nat’l City Bank, 396 F.2d 897, 900–01 (2d Cir. 1968).

46. Id. at 901.

47. Id.

48. Id. at 902.

49. Id. at 902–05.
Court decisions ordering the production of documents held abroad contributed to the broadening of the Federal Rules, which in turn have cemented the development of transnational discovery, even over non-parties, in four Rules: 26, 34, 45, and 69. Rule 26 allows broad discovery of “any nonprivileged matter that is relevant.”50 This language applies in the context of Rule 34, which provides that, in general, “a party may serve on any other party a request within the scope of Rule 26(b).”51 This can include a request to produce any designated document or electronically-stored information.52 Rule 45 specifically allows litigants to issue document subpoenas to non-parties, limiting this power only to the extent that it imposes an “undue burden,” it fails to provide a reasonable time to comply, or it requests privileged materials.53 Finally, Rule 69 allows litigants in the post-judgment context to “obtain discovery from any person . . . as provided in these rules or by the procedure of the state where the court is located.”54

Notably absent from these rules is any limitation on the geographical scope of information discovery requests, unlike the limitations imposed on deposition subpoenas.55 Therefore, these rules provide a potent weapon for U.S. litigants seeking transnational discovery. In the post-judgment context, for example, the Second Circuit has interpreted these rules to mean that a judgment creditor is “entitled to discover the identity and location of any of the judgment debtor’s assets, wherever located.”56 Given such broad language, it is easy to see how conflicts between domestic and foreign law became commonplace.

C. Conflicts Between U.S. Discovery and Foreign Laws

The contrast between U.S. and foreign discovery practices is stark. As explained above, American courts have long been comfortable exercising their broad discovery and jurisdictional powers over parties wherever located. Discovery in civil law countries is drastically different from U.S. methods. Because the inquisitorial system predominates in civil law countries, it is judges, not the parties themselves, who have the exclusive power to gather facts.57 After compiling evidence, civil law judges produce an official summary, or dossier,

51. FED. R. CIV. P. 34(a) (emphasis added).
52. FED. R. CIV. P. 34(a)(1)(A).
53. FED. R. CIV. P. 45.
54. FED. R. CIV. P. 69(a)(2).
55. See FED. R. CIV. P. 45.
56. First City, Tex.-Hous., N.A. v. Rafidain Bank, 281 F.3d 48, 54 (2d Cir. 2002) (quoting Nat’l Serv. Indus., Inc. v. Vafla Corp., 694 F.2d 246, 250 (11th Cir. 1982)) (emphasis added); see also FED. R. CIV. P. 69(a)(2).
that is used at trial. With regards to document production, some foreign countries provide severe restrictions. In France, for example, there is no U.S. concept of “blanket” requests for documents. Instead, parties can make specific requests to the judge, who can then order the production of identified documents. In Germany, parties are not obligated to conduct a search for information that is not readily available. Scholars suggest that Europeans’ respect for privacy rights explains their overarching anxiety with broad discovery. As an extension of this general narrowing of discovery, in civil law countries there is no concept of pretrial discovery. In the United Kingdom, procedural rules limit discovery of non-parties. In other countries, there is no post-judgment discovery either, making it difficult for judgment creditors to find


59. NML Capital, Ltd. v. Republic of Arg., No. 03 Civ. 8845(TPG), 2013 WL 491522 (S.D.N.Y. Feb. 8, 2013) (reviewing the data secrecy laws of Brazil, Spain, Bolivia, Chile, Panama, Paraguay, Argentina, and Uruguay); Diana Lloyd Muse, Discovery in France and the Hague Convention: The Search for a French Connection, 64 N.Y.U. L. REV. 1073, 1075 n.8 (1989) (discussing French limitations and noting that other civil law countries are similar); Langbein, supra note 57 (discussing German law).

60. Muse, supra note 59, at 1080–81 (“In France, the current Code of Civil Procedure (the Code) vests all fact-finding authority in the judge. For example, each party, through its attorney (avocat), must make any request for written evidence to the judge, who then has the discretion to order an opposing party to produce the evidence. Even though the Code authorizes the avocat to ask a judge to order document production, judges do not always grant such requests. Thus, although the current Code appears to give the judge broad powers to require the production of evidence, commentators agree that in practice, the fact-finding process in civil cases has, to a large extent, retained its traditionally limited scope.”)

61. Langbein, supra note 57, at 827 (“The defendant’s answer follows the same pattern. It should be emphasized, however, that neither plaintiff’s nor defendant’s lawyer will have conducted any significant search for witnesses or for other evidence unknown to his client. Digging for facts is primarily the work of the judge.”).


63. Muse, supra note 59, at 1075 (“Moreover, France, like most civil law countries, does not have any form of pretrial discovery as it exists in the United States.”). It is also somewhat limited in the United Kingdom. First Am. Corp. v. Price Waterhouse LLP, 154 F.3d 16, 22 (2d Cir. 1998) (noting an English court’s rejection of a discovery request because the particular discovery in that case was not “provided for under the Hague Convention or British law”).

64. South Carolina Ins. v. Assurantie Maatschappij “de Zeven Provincien” N.V., [1986] 3 W.L.R. 398 (HL) (statement of Lord Brandon) (noting that because of certain limitations, “there is no way in which a party to an action in the High Court in England can compel pre-trial discovery as against a person who is not a party to such action”).
assets. Because of these strict limits on discovery, it is no surprise that Europeans worried not only about expansive U.S. judicial power, but also expansive discovery in general.

Given the unrestricted exercise of U.S. discovery allowed by the “personal jurisdiction plus control” test, conflicts with foreign law were a common occurrence in the middle and later part of the twentieth century. Courts routinely ordered production of documents held by bank branches in Panama and Canada, banking records in Switzerland and Germany, foreign shipping lines’ documents “wherever located,” and oil company documents in “foreign countries.” Some of the most offensive practices in the eyes of foreign sovereigns included the taking of depositions by American lawyers in foreign countries without the consent of local authorities. Not surprisingly, foreign laws imposed strict restraints to prevent these abuses. If an American lawyer sought to take evidence in France, for example, where discovery is a judicial task, the French considered it an “unlawful usurpation of the public judicial function and an illegal intrusion on the nation’s judicial sovereignty.” This stance holds sway in many other civil law countries, including Japan, where regulations place strenuous requirements on the taking of a deposition therein, authorizing it “only if (1) the witness or party is willing to be deposed, (2) the deposition takes place on U.S. consular premises, (3) a consular officer presides over that deposition . . . and each participant traveling from the United States to Japan to participate in the deposition obtains a ‘deposition visa.’”

65. For example, in France there are only self-help attachment procedures. CODE DE PROCEDURE CIVILE [C.P.C.] art. L111-1 (Fr.).
70. United States v. First Nat’l City Bank, 396 F.2d 897, 901–03 (2d Cir. 1968).
73. See Graco, Inc. v. Kremlin, Inc., 101 F.R.D. 503, 520 (N.D. Ill. 1984); see also Muse, supra note 59, at 1073 (discussing U.S. “legal tourists” who went to France in search of evidence).
deposition can be a criminal violation given that “Brazilian law subjects foreign attorneys who conduct depositions of Brazilian nationals in Brazil to potential arrest, detention, expulsion or deportation.” In Switzerland, there have been criminal penalties for such “intrusive” discovery since 1937.

It is worth highlighting that antitrust investigations by the U.S. government fueled much of the backlash from European countries. In response to antitrust investigations in the shipping industry, France passed a blocking statute in 1968 prohibiting the production of information to foreign judicial authorities “related to maritime transport.” In 1980, with the French Assembly complaining about U.S. “fishing expeditions” and “legal tourism,” France expanded the blocking statute to prohibit the production to foreign legal authorities of any “economic, commercial, industrial, financial or technical” information “which is capable of harming [the] . . . interests of France.” Even the United Kingdom took part in this backlash, enacting its own limitations on the gathering of evidence therein for use in other countries. This backlash was evidence of a broad failure of international comity. Years later, this reaction against U.S. discovery practices prompted the following comment from the reporter to the Restatement of Foreign Relations Law:

No aspect of the extension of the American legal system beyond the territorial frontier of the United States has given rise to so much friction as the requests for documents in investigation and litigation in the United States. As of 1986, some 15 states had adopted legislation expressly designed to counter United States efforts to secure production of documents situated outside the United States.


78. Laker Airways Ltd. v. Pan Am. World Airways, 607 F. Supp. 324, 327 (S.D.N.Y. 1985) (“The failure to use the Hague Convention is more than a mere technicality. The extraterritorial jurisdiction asserted over foreign interests by the American antitrust laws has long been a sore point with many foreign governments, including that of the United Kingdom.”).


80. Toms, supra note 1, at 596 n.41, 611.

81. Protection of Trading Interests Act 1980, c. 11 (U.K.); see also Laker Airways Ltd., 607 F. Supp. at 327 (“The English Protection of Trading Interests Act of 1980 . . . authorizes and empowers the Secretary of State for Trade and Industry to interpose the official power of the British Government so as to prevent persons conducting business in the United Kingdom from complying with foreign judicial or regulatory provisions designated by the Secretary of State as intrusive upon the sovereignty of that nation.”).

There is no doubting the strength of hostility against U.S. courts in international circles. Many countries view transnational discovery as judicial usurpation, a wasteful exercise, and a direct threat to their sovereignty.\(^{83}\) By the late 1960s there was a clear conflict between the civil law world and the U.S. discovery system. A treaty designed to bridge the gap between U.S. discovery and civil law countries seemed necessary.\(^{84}\) Both parties had interests at stake: the United States in creating a system that would facilitate the production of evidence, and foreign countries in moderating liberal U.S. discovery practices.\(^{85}\)

**D. The Hague Convention**

The Hague Convention on the Taking of Evidence Abroad in Civil or Commercial Matters was signed on October 26, 1968, by all the present delegations (the “Hague Convention” or the “Convention”).\(^{86}\) The primary goal of the Convention was to “bridge differences between the common law and civil law approaches to the taking of evidence abroad.”\(^{87}\) As the Letter of Submittal to the President of the United States noted, the signatories were willing “to proceed promptly for work on the evidence convention” because of “the difficulties encountered by courts and lawyers in obtaining evidence abroad from countries with markedly different legal systems.”\(^{88}\) The United States led the negotiations, represented by Philip W. Amram, who was appointed rapporteur of the commission and co-chairman of the drafting committee.\(^{89}\) The Report of the U.S. delegation highlighted that as a matter of international comity, the convention sought to construct a process that was “tolerable” to the authorities of the country where the evidence was located.\(^{90}\) Further, the Report

---


84. See Message from the President, supra note 58; see also Rapport de la Commission spéciale, 4 Conférence de La Haye de droit international privé: Actes et documents de la Onzième session 55 (1970) (Actes et documents).


87. Comment, The Hague Convention, supra note 85, at 1464; see also Message from the President, supra note 58.

88. Message from the President, supra note 58, at 324.


90. Id. at 806.
emphasized that “the doctrine of ‘judicial sovereignty’ had to be constantly borne in mind.”91 That is, in civil law countries the courts take evidence, unlike in the United States where litigants conduct discovery and depositions. International comity influenced the thinking of the U.S. drafters and the Convention as a whole, and it energized the negotiations.92

The Convention created the system of Letters of Request as the primary vehicle for the production of information abroad, allowing parties to seek evidence in a more regulated manner. The system placed national authorities of both the requesting country and the target country as gatekeepers. For example, if an American litigant sought evidence located in France, he would have to adhere to the following procedure: (1) litigant files a proposed Letter of Request with the American court describing the information sought and the parties involved; (2) the court reviews and approves or rejects the letter; (3) litigant obtains a translation and sends the letter, with judicial approval, to the French Justice Ministry; (4) French Justice Ministry refers the request to the District Attorney for the particular location in France; and (5) internal French evidence procedures take effect.93 As described, the procedure gives French authorities a gatekeeping role where they can evaluate evidence requests and decide whether to comply with them. Moreover, in deference to local law, the evidence is actually obtained through the host country’s evidentiary procedures.

Although the Convention delegates did not explicitly agree to the primacy of the Hague Convention over transnational discovery procedures offered by local courts, a Commission gathered in 1989 to review the functioning of the Convention stated that “the [Special Commission Report on the Operation of the Hague Service Convention] thought that in all Contracting States, whatever their views as to its exclusive application, priority should be given to the procedures offered by the Convention when evidence located abroad is being sought.”94

The United States ratified the Hague Convention in 1972.95 Despite its apparent potential, litigants hardly used its procedures over the next decade.96

91. Id.
Federal courts split over whether the Hague Convention provided a mandatory or an optional alternative to the Federal Rules. For example, in *Compagnie Francaise*, the court noted that “extraterritorial discovery has been standard for some time and there is no evidence that the United States, in agreeing to comply with the Hague Convention, intended to abandon this practice.” On the other hand, some U.S. courts enforced the Hague Convention, seeing it as the “preferable” means for international discovery, and even instituted a rule of “first resort” to the Convention. By 1988, transnational discovery through the Hague Convention was a developing area of the law. There was much promise that Hague Convention procedures could be a method for redeeming international comity, but in 1988 the Supreme Court ended any hope that the Convention might displace the Federal Rules as the primary method of transnational discovery.

II. COMITY IN RETREAT: AÉROSpatiale AND THE DEFEAT OF THE HAGUE CONVENTION

102. Id. at 526.
104. Société Nationale Industrielle Aérospatiale, 482 U.S. at 547.
105. Id. at 542–43.
106. Id. at 534–39.
analysis of the respective interests of the foreign nation.”

Crucially, in footnote twenty-eight, the Court endorsed a revised version of the Restatement balancing test focusing on the interests of the countries at issue, the importance of the documents, the specificity of the request, the place of origin of the documents, and the availability of alternative means of obtaining the information. The Court otherwise noted that objections to discovery “that foreign litigants advance should . . . receive the most careful consideration. In addition, we have long recognized the demands of comity in suits involving foreign states, either as parties or as sovereigns with a coordinate interest in the litigation.” Although it rejected the Hague Convention and confined it to an optional procedure, the Supreme Court was careful to pay lip-service to international comity—a feature that became common in the cases that followed Aérospatiale.

The Court’s holding represented a victory for broad U.S. discovery and a momentous defeat for international comity. The decision peremptorily dismissed the arguments in favor of the Convention put forth by the United Kingdom, France, and Germany as amici, leading a prominent commentator to declare it “loosely-reasoned.” As Justice Blackmun wrote in dissent, “the needs of the international commercial system and the accommodation of those needs . . . [are] embodied in the Convention,” and yet the Court seemed content to reject it. The decision was suffused with a kind of judicial chest-thumping because, as some have put it, the Court was concerned with “reaffirming the sovereignty of our judicial system.”

Aérospatiale unleashed a new wave of expansive foreign discovery under the Federal Rules. Most courts considering requests for discovery paid lip service to the Convention, just like Aérospatiale had, cautioning that foreign sovereign interests had to be taken into account. But despite this false deference, the vast majority of cases involving requests for discovery of documents or assets located abroad rejected proceeding through the Convention.

107. Id. at 543–44.
108. See id. at 544 n.28.
109. Id. at 546.
111. Société Nationale Industrielle Aérospatiale, 482 U.S. at 568 (Blackmun, J., concurring in part and dissenting in part).
Aérospatiale damaged the Convention to such an extent that even in circumstances where the Federal Rules seemed inappropriate, such as ordering discovery from foreign non-parties, courts nonetheless rejected the Convention.\textsuperscript{115} The Third Circuit went as far in \textit{In re Automotive Refinishing Paint Antitrust Litigation} as refusing to carve out a rule of first resort to the Convention for jurisdictional discovery.\textsuperscript{116} Similarly, the Ninth Circuit refused to honor the Convention even in clear instances of foreign criminal laws prohibiting discovery.\textsuperscript{117}

Despite courts ostensibly evaluating foreign interests during this period, the Restatement standard seemed muddled, unworkable, and purely for show. Some criticized the Aérospatiale model as misguided because U.S. courts could not accurately take into account foreign interests due to U.S. judges' lack of experience with foreign laws.\textsuperscript{118} Others criticized the broad discovery powers given to district courts, which consequently did not allow for proper oversight by appellate courts.\textsuperscript{119} Perhaps most importantly, the balancing test allowed courts to discount international comity in favor of domestic interests without proper scrutiny. In the words of one court, “regrettably, the [Aérospatiale] Court declined to set forth specific rules” for the international comity analysis.\textsuperscript{120}


117. Richmark Corp. v. Timber Falling Consultants, 959 F.2d 1468, 1478 (9th Cir. 1992).
118. Fields, \textit{supra} note 112, at 308.
119. \textit{Id}.
Convention and foreign criminal laws only because a U.S. plaintiff was involved. In the Southern District of New York, courts routinely dismissed concerns with foreign laws and international comity, including in *NML Capital Ltd. v. Republic of Argentina*, where the court rejected possible conflicts with the laws of nine countries and ordered the production of documents located therein. Not only did U.S. courts refuse to comply with the Hague Convention, they also made it increasingly difficult for the targets of discovery requests to invoke it—placing on them the burden of demonstrating “that it is more appropriate for the Court to follow the Hague Convention” than the Federal Rules.

These cases became exemplary of the trend of *Aérospatiale*-inspired cases, followed almost unanimously in lower courts, that can be described as nothing less than a wholesale and total rejection of both international comity and the Hague Convention. This rejection of the Hague Convention after *Aérospatiale* presents an odd denial of a treaty that the United States sponsored. Courts have ruled against discovery through the Federal Rules in relatively few cases since *Aérospatiale* and have invoked the Hague Convention in even fewer cases. Instead, courts have typically relied on the Federal Rules paradigm defined since the 1960s.


125. See Report of the United States Delegation, supra note 86, at 786. As Justice Blackmun wrote in his *Aérospatiale* dissent, the “Convention was drafted at the request of the United States, which sought to broaden the techniques available for the taking of evidence abroad.” Société Nationale Industrielle Aérospatiale v. U.S. Dist. Court for S. Dist. of Iowa, 482 U.S. 522, 549 (1987) (Blackmun, J., concurring in part and dissenting in part).

It is important to emphasize that various countries entered into the Hague Convention in the hope of limiting U.S. discovery and safeguarding their judicial sovereignty.\textsuperscript{127} Broad U.S. discovery continues to challenge this hope. For that reason, civil law countries have been vociferous about their rejection of \textit{Aérospatiale} and its progeny.\textsuperscript{128} France has been among the most emphatic in its

\textsuperscript{127} See infra Part I (c); Graco, Inc. v. Kremlin, Inc., 101 F.R.D. 503, 519 (N.D. Ill. 1984) (“It cannot be denied that foreign displeasure with American discovery procedures played some part in shaping the Convention”).

\textsuperscript{128} Roth, supra note 57 at n.33 (citing “Letter from Edwin R. Alley, Esq., Carpenter, Bennett & Morrissey to Judge Joseph F. Weis, Jr., Senior United States Circuit Judge, United States Court of Appeals for the Third Circuit (Apr. 11, 1990) (noting that at least one nation contemplated acceding to the Evidence Convention before deciding otherwise because it viewed the Court’s \textit{Aérospatiale}...”)}
rejection of the Federal Rules as a legitimate avenue for foreign discovery.\textsuperscript{129} Germany and Switzerland have emphasized in amici, after \textit{Aérospatiale}, that “discovery of [their] nationals pursuant to the Federal Rules of Civil Procedure constitutes an intrusion into [their] sovereignty.”\textsuperscript{130} China has threatened the stability of its bilateral relations with the United States over the issue of discovery.\textsuperscript{131} Other nations have similarly voiced their concern with U.S. discovery.\textsuperscript{132} The effects of \textit{Aérospatiale} were clear: the comity system built by the Hague Convention was destroyed.

In sum, the \textit{Aérospatiale}-era consisted of a feeble tripartite arrangement to accommodate international comity. Whenever difficulties between U.S. discovery procedures and foreign laws surfaced, courts had to first determine whether there was a “true conflict” between domestic and foreign law.\textsuperscript{133} This was usually a low bar that could be met through expert submissions.\textsuperscript{134} Once that threshold was met, courts only needed (1) personal jurisdiction over a foreign party; (2) a finding that the party had possession, custody, or control over the documents; and (3) an analysis of the countries’ interests through the \textit{Aérospatiale}-endorsed Restatement balancing test. It was this \textit{Aérospatiale} arrangement that unleashed a wholesale repudiation of international comity by circuit and lower courts. It ended the promise of the Convention as a way to redeem international comity, and created difficulty for countries that saw broad U.S. discovery and exercise of U.S. jurisdiction as a threat to their sovereignty.

\section{III.
THE RETURN OF INTERNATIONAL COMITY: \textsc{Daimler}, \textsc{Gucci}, AND \textsc{Motorola} ESTABLISH A NEW PARADIGM}

Three recent cases have altered the landscape of transnational discovery and call into question the \textit{Aérospatiale} paradigm. It appears as if, in the span of a few years, international comity is experiencing a revival. These cases prominently featured the interests of foreign nations, and courts responded by

\begin{itemize}
  \item \textbf{129.} \textit{In re} Perrier Bottled Water Litig., 138 F.R.D. 348, 355 (1991) ("Although not all civil-law countries have expressed their disfavor of private litigants’ use of the Federal Rules’ procedures within its borders, of those which have, France has been among the most emphatic.").
  \item \textbf{134.} \textit{See} id.
\end{itemize}
refusing to subject foreign entities to jurisdiction in the United States or otherwise impose on them overbroad duties, including those in conflict with foreign laws. These cases open the door to renewed respect for international comity.

*Daimler*, *Gucci*, and *Motorola* are significant for several reasons, including their concern for foreign retaliatory laws, overbroad application of U.S. discovery procedures, and the possible effects of such broad jurisdiction on the international economy. Although previous courts have voiced similar concerns, these three decisions are notable because they espouse a consistent rejection of *Aérospatiale*-era jurisprudence and come from the Supreme Court, the Second Circuit, and the New York Court of Appeals.

Before undertaking an analysis of these cases, one point of clarification is in order: *Daimler*, *Gucci*, and *Motorola* have not come out of thin air. There has been a movement in the past few years towards cabining the extraterritorial application of U.S. law, and this Article argues that *Daimler*, *Gucci*, and *Motorola* extend this movement into the realm of discovery. In 2004, the Supreme Court referred to certain hypothetical extraterritorial applications of U.S. law as "legal imperialism" that did not align with principles of comity. The movement to cabin U.S. law has been reinforced by the post-financial crisis era of austere U.S. domestic and foreign policy. Two cases signaled the beginnings of renewed respect for international comity: *Kiobel v. Royal Dutch Petroleum Co.* (*Kiobel*) and *Morrison v. National Australia Bank Ltd.* (*Morrison*).

In *Morrison*, the Court found that certain provisions of the Securities Exchange Act did not apply extraterritorially, chiding the Second Circuit for "excis[ing]" the presumption against extraterritoriality of U.S. laws. The Court noted the incompatibility of the Act with foreign laws, including foreign rules about "what discovery is available in litigation," and concluded the Act did not apply to conduct that occurred outside the United States. Likewise, in *Kiobel*, the Court rejected the application of the Alien Tort Statute to events in Nigeria, reasoning that the presumption against the extraterritorial application of U.S. laws prevented such overbroad use of the statute and "serves to protect against unintended clashes between our laws and those of other nations which could result in international discord." The Court specifically warned against the "danger of unwarranted judicial interference in the conduct of foreign


136. Other cases around this time also emphasized similar concerns. *See*, e.g., *Figueiredo Ferraz E Engenharia de Projeto Ltda. v. Republic of Peru*, 665 F.3d 384, 392–93 (2d Cir. 2011) (dismissing on forum non conveniens grounds foreign party’s action for recognition of arbitral award in connection with foreign controversy).


138. *Id.* at 269.

Most importantly, the Court emphasized the deleterious foreign policy consequences of “impos[ing] the sovereign will of the United States onto conduct occurring within the territorial jurisdiction of another sovereign.” The concerns voiced in Kiobel and Morrison are precisely the kinds of concerns involved in international discovery.

Daimler, Gucci, and Motorola are an extension of Kiobel and Morrison because they voice concerns with international comity in the context of personal jurisdiction and discovery—areas previously dominated by the narrow focus of the Aérospatiale paradigm. These three cases are animated by several factors that will be discussed below, including (1) changes brought by the modern globalized economy and (2) the danger of retaliatory laws in international relations.

A. Daimler: The Supreme Court Revives International Comity

In Daimler, plaintiffs filed various claims under the Alien Tort Statute and the Torture Victims Protection Act against the Daimler Corporation in the Northern District of California. The claims alleged that Daimler’s Argentine subsidiary “collaborated” with the Argentine government in perpetrating murder, kidnappings, torture, and other crimes against plaintiffs’ relatives. The action had no connection to the United States, as plaintiffs were foreign parties, the defendant was a foreign corporation, and the situs was Argentina. The case had deep implications, however, for Germany and Argentina, as Daimler is headquartered and registered in Germany, and Argentine officials were implicated in the claims. Nevertheless, plaintiffs asserted that the court had general personal jurisdiction because Daimler’s subsidiary, Mercedes-Benz USA, LLC, distributed Daimler vehicles to dealerships in California and maintained an office and other facilities in the state. The district court refused to find jurisdiction, holding that Daimler’s “affiliations with California” were insufficient. The Ninth Circuit reversed based on its finding that Mercedes was Daimler’s agent and as such provided continuous activity in California sufficient for general jurisdiction.

In reversing the Ninth Circuit, a unanimous Supreme Court fundamentally altered the traditional test for general personal jurisdiction and announced that a

140. Id.
141. Id. at 1667.
142. Although not a topic discussed in this paper, the rise of e-discovery has certainly affected and increased instances of conflict of laws. See generally William R. Maguire, Current Issues in Federal Civil E-Discovery, Proportionality, International Discovery and Deposition Practice and Changes to the Federal Rules of Civil Procedure, in CURRENT DEVELOPMENTS IN FEDERAL CIVIL PRACTICE 2012 (William P. Frank & Jonathan L. Frank eds., 2012).
144. Id. at 751.
145. Id. at 752.
146. Id. at 753.
court could exercise general jurisdiction over a corporation only when it is “essentially at home” in the forum state. 147 This new “at home” test replaced the previous “continuous and systematic general business contacts” test, which the Court called “unacceptably grasping.” 148 Discarding the idea that a large corporation can be “at home” in all of the different places where it operates, Daimler pointed to the place of incorporation and principal place of business as the “paradigm[atic] . . . bases for general jurisdiction” because they “have the virtue of being unique—that is, each ordinarily indicates only one place—as well as easily ascertaintable.” 149 Both of these places for Daimler were located in Germany. Applying its new rule to the case, the Court held that there was no jurisdiction because “neither Daimler nor [Mercedes] is incorporated in California, nor does either entity have its principal place of business there.” 150 Moreover, it found that this was not a case of “exceptional” circumstances that would warrant finding general jurisdiction.

Daimler fundamentally changed the corporate personal jurisdiction analysis, and it did so based partly on grounds of international comity. In a decisive sentence, the Court reproached the Ninth Circuit for “pa[ying] little heed to the risks to international comity its expansive view of general jurisdiction posed.” 151 Rather than expressing mere generalities about foreign laws, the Court cited specific European Union rules, noting that, “[i]n the European Union, for example, a corporation may generally be sued in the nation in which it is ‘domiciled,’ a term defined to refer only to the location of the corporation’s ‘statutory seat,’ ‘central administration,’ or ‘principal place of business.’” 152 Additionally, to show that its reasoning was based on concrete concerns, the Court cited the Solicitor General’s opinion that “foreign governments’ objections to some domestic courts’ expansive views of general jurisdiction have in the past impeded negotiations of international agreements on the reciprocal recognition and enforcement of judgments.” 153 The Court concluded that such comity concerns guided its decision regarding personal jurisdiction.

Emphasizing the import of international comity in the decision, amici submitted to the Supreme Court in Daimler persuasively highlighted the dangers

---

147. Id. at 769.
148. Id. at 761.
149. Id. at 760 (internal citations and quotes omitted).
150. Id. at 761.
151. Id. at 763.
152. Id.
to the economy posed by broad U.S. jurisdiction and weakened international harmony. Various foreign corporations, including German banks and the Swiss Chamber of Commerce, argued as amici that “[t]he extraterritorial reach of U.S. laws—including U.S. courts’ exercise of personal jurisdiction over non-U.S. businesses with respect to those companies’ activities outside the United States—creates tremendous uncertainty that deters investment in and trade with the United States.”¹⁵⁴ Moreover, they highlighted that a rule granting general jurisdiction over foreign corporations with affiliates in the United States “has significant implications for international comity.”¹⁵⁵ The amici pointed specifically to problems with broad U.S. discovery, arguing that “given the uniquely expansive procedural rules governing civil litigation in the United States—including broad discovery . . . there is no doubt that foreign enterprises would revamp their operations to avoid subjecting themselves to general jurisdiction in U.S. courts, even if that would require relocating or significantly reducing their U.S. operations.”¹⁵⁶ The amici also recognized that broad assertions of extraterritorial jurisdiction “will inevitably injure . . . international comity.”¹⁵⁷ Finally, the amici concluded that withdrawal of foreign companies “would inflict significant harm upon the U.S. economy [and] would decrease foreign direct investment, which contributes significantly to [the U.S.] economy.”¹⁵⁸ Undoubtedly, according to the amici, the economic and international comity effects of personal jurisdiction and discovery are closely linked.

Defendants in Daimler also highlighted the effects on the international system of America’s uninhibited judicial power, noting that Judge O’Scaannlain had sought a rehearing of the Ninth Circuit’s decision en banc because the decision could “have unpredictable effects on foreign policy and international comity . . . as well as on our nation’s economy.”¹⁵⁹ Defendants warned that foreign corporations would withdraw their investments from the United States and the possible damage this could cause to “U.S. consumers and the U.S. economy.”¹⁶⁰ Justice Sotomayor explicitly recognized these arguments, noting in her concurrence that “[w]hat has changed since International Shoe is not the due process principle of fundamental fairness but rather the nature of the global economy.”¹⁶¹

In sum, Daimler’s significance for transnational discovery rests on its direct attack on the foundations of the Aérospatiale paradigm: broad personal

¹⁵⁵. Id. at *3.
¹⁵⁶. Id. at *10.
¹⁵⁷. Id. at *4.
¹⁵⁸. Id. at *11.
¹⁶⁰. Id. at *35.
jurisdiction and disregard for international comity. Without these two fundamentals, the *Aérospatiale* paradigm is weakened, and the Hague Convention remains the only alternative for transnational discovery. *Daimler* also shows that the Supreme Court has begun to grapple with international trade and comity in a new way. The decision intimates that the *Aérospatiale* paradigm is outdated. *Daimler* threw down the gauntlet for future courts, urging them to consider international comity as a crucial concern, rather than as a formality that should be dismissed through a contrived balancing test.

**B. Gucci: The Second Circuit Takes International Comity a Step Further**

*Gucci America Inc. v. Weixing Li*, decided in September of 2014, followed directly in the footsteps of *Daimler*. In *Gucci*, defendants sold counterfeit luxury goods over the Internet, labeled as Gucci and other brands, and wired the proceeds of their sales to Bank of China accounts. Plaintiffs, manufacturers of the real luxury products, filed an action in the Southern District of New York seeking to protect their intellectual property from the alleged counterfeiters. During pre-trial discovery, plaintiffs sought to freeze defendants’ assets to ensure recovery and safeguard evidence of the unlawful conduct. Accordingly, plaintiffs served Bank of China (“BOC”) with an asset freeze injunction and a subpoena for documents at its New York City branch, seeking an asset freeze and information from “any and all Bank of China accounts associated with [defendants].” In response to the subpoena and asset freeze, the BOC, a bank that is majority owned by the Chinese government, produced documents from its New York City branch but stated that it could not search for records located in China. The district court held BOC in contempt, and the bank appealed.

Applying *Daimler*, the Second Circuit concluded that there was no general jurisdiction over BOC because the bank was not “at home” in New York. The court dismissed plaintiffs’ arguments that *Daimler* did not apply to non-parties, stating:

> BOC’s nonparty status does not alter the applicability of these cases to the question presented here. The essence of general personal jurisdiction is the ability to entertain ‘any and all claims’ against an entity based solely on the entity’s activities in the forum, rather than on the particulars of the case before the court.

---

162. *Gucci Am., Inc. v. Weixing Li*, 768 F.3d 122, 126 (2d Cir. 2014).
163. *Id.*
164. *Id.*
165. *Id.*
166. *Id.* at 127.
167. *Id.* at 128.
168. *Id.* at 134–35.
169. *Id.* at 134 n.13.
The Second Circuit further held that BOC did not waive its personal jurisdiction defense because *Daimler* had created an entirely new test.\textsuperscript{170} Therefore, the Second Circuit found general jurisdiction was lacking because BOC “has branch offices in the forum, but is incorporated and headquartered elsewhere,” and it “has only four branch offices in the United States and only a small portion of its worldwide business is conducted in New York.”\textsuperscript{171}

Crucially, the court emphasized the importance of international comity, remanding with specific instructions for the district court to consider “whether, assuming the necessary [specific] jurisdiction is present, such an order is consistent with principles of international comity.”\textsuperscript{172} The court stressed that its decision was wholly based on *Daimler*, and noted that the Supreme Court had “expressly warned against the ‘risks to international comity’ of an overly expansive view of general jurisdiction inconsistent with the fair play and substantial justice due process demands.”\textsuperscript{173} The court found important a BOC declaration from a Chinese law expert showing a direct conflict between plaintiffs’ demands and Chinese banking laws.\textsuperscript{174} Emphasizing its overriding concern with international comity, the court cited *Aérospatiale* for the principle that “[t]he doctrine of international comity ‘refers to the spirit of cooperation in which a domestic tribunal approaches the resolution of cases touching the laws and interests of other sovereign states.’”\textsuperscript{175} Moreover, the court held that an international comity analysis under the Restatement (Third) of Foreign Relations\textsuperscript{176} is appropriate even in the context of asset freeze injunctions, a

\textsuperscript{170}. Id. at 135.

\textsuperscript{171}. Id.  Gucci also intimated that personal jurisdiction might be found where an entity has “consented to personal jurisdiction in New York by applying for authorization to conduct business in New York and designating the New York Secretary of State as its agent for service of process.” Id. at 136 n.15. However, this seems unlikely to develop into a feasible jurisdiction avenue because foreign banks register with state banking authorities under Section 200 of the New York Banking Law, which grants only specific jurisdiction. See Gliklad v. Bank Hapoalim B.M., No. 155195/2014, 2014 WL 3899209, at *1 (N.Y. Sup. Ct. Aug. 4, 2014). In addition, the New York State Senate recently rejected a bill (S. 7078) that would have made consent to do business a grant of general jurisdiction. But see Acorda Therapeutics, Inc. v. Mylan Pharm. Inc., 78 F. Supp. 3d 572, 587 (D. Del. 2015) (finding that registration to do business in Delaware conferred the court with general jurisdiction).

\textsuperscript{172}. Gucci Am., Inc., 768 F.3d at 129.

\textsuperscript{173}. Id. at 135 (quotation marks omitted).

\textsuperscript{174}. Id. at 138.

\textsuperscript{175}. Id. at 139.

\textsuperscript{176}. RESTATEMENT (THIRD) OF FOREIGN RELATIONS LAW § 403 (1987). The Restatement provides the following:

(1) Even when one of the bases for jurisdiction under § 402 is present, a state may not exercise jurisdiction to prescribe law with respect to a person or activity having connections with another state when the exercise of such jurisdiction is unreasonable.

(2) Whether exercise of jurisdiction over a person or activity is unreasonable is determined by evaluating all relevant factors, including, where appropriate: (a) the link of the activity to the territory of the regulating state, i.e., the extent to which the activity takes place within the territory, or has substantial, direct, and foreseeable effect upon or in the territory; (b) the connections, such as nationality, residence, or
situation that did not previously require such an analysis. Simply stated, the court made international comity a central consideration in the case.

To further emphasize the renewed importance of international comity, the Second Circuit peremptorily dismissed plaintiffs’ waiver argument in the asset freeze context, ruling that international comity arguments cannot be waived. The court stated: “given the important role that comity plays in ensuring the recognition which one nation allows within its territory to the legislative, executive or judicial acts of another nation, having due regard both to international duty and convenience, we do not deem the issue forfeited.” Finally, the court instructed the district court to conduct a comity analysis and give “due regard to the various interests at stake, including: (1) the Chinese Government’s sovereign interests in its banking laws; [and] (2) the Bank’s expectations, as a nonparty, regarding the regulation to which it is subject in its home state and also in the United States.”

In renewing the importance of international comity, Gucci follows directly from Daimler. Both courts found it vital to highlight the important interests of foreign countries and the necessary limits that international relations impose on U.S. courts. Several issues of foreign relations came to the forefront in Gucci that became imperative to the case and raised the issue of international comity.

First, the Bank of China introduced a letter written by an official from the China Banking Regulatory Commission, raising concerns about the impact the case could have on China-U.S. relations. The letter argued that Chinese bank privacy and secrecy laws were essential to the country’s sovereignty, and that a

---

177. Gucci Am. Inc., 768 F.3d at 140 (“Ordering compliance with an asset freeze, however, implicates different concerns from those implicated by an order for the production of documents”).
178. Id.
179. Id. (internal quotation marks and citations omitted).
conflicting order from a U.S. court would place the bank in an untenable position. 

Second, the High People’s Court of Beijing Municipality ordered the BOC to resume regular services to the defendants, in effect overturning the asset freeze imposed by the district court. This conflicting ruling gave international comity more than a speculative role in the case. The Chinese government and its courts were serious about protecting their interests, which should have been expected given the BOC’s status as a state-owned entity of vital importance to the Chinese government.

Third, the U.S. government as amicus curiae argued strongly for vacatur so that the lower court could perform a “thorough international-comity” analysis and carefully weigh the sovereign interests at stake. The U.S. government argued that when the extraterritorial application of U.S. laws implicate sovereign interests of foreign countries, “submissions from interested governments that address comity issues should be given serious consideration.” Moreover, the amicus made an important declaration: comity is not “mere courtesy and good will,” but involves serious consideration of “international duty and convenience.” The amicus announced the government’s official policy that the U.S. legal system should promote harmony, coordination, and respect for foreign sovereigns. Accordingly, the government lambasted the district court for not taking into account China’s interests, arguing that the “Gucci court should have been more mindful . . . and should not have summarily dismissed representations describing the national importance of China’s banking secrecy laws.” The U.S. position clearly recognized that deep issues of Chinese sovereignty and trade relations were at stake.

In sum, international comity was integral in the Second Circuit’s Gucci decision. Building on Daimler, the Second Circuit and U.S. amicus emphasized the renewed importance of international comity and its central role in foreign relations. Regardless of the reasons for the renewed importance of comity, the Second Circuit wanted to make clear to lower courts that comity should be seriously considered in every decision affecting foreign countries.

182. Id.
184. Id.
185. Id. at *17.
186. Id. at *25.
C. Motorola: The New International Comity Paradigm is Established

In Motorola Credit Corp. v. Standard Chartered, the New York State Court of Appeals upheld the continued applicability of the separate entity rule, which dictates that a U.S. bank branch is not concerned or responsible for assets held in foreign branches and thus cannot be forced to restrain or turnover assets held abroad. The separate entity rule is the embodiment of international comity; it exists to avoid forcing foreign bank branches to comply with U.S. orders and as recognition of foreign sovereign power over banks located in their countries. Daimler and Gucci provided intellectual support for the revival of the rule.

Motorola involved the prolonged litigation of Motorola Credit Corporation against the Uzans, a Turkish family involved in a sprawling web of businesses. The Uzans borrowed billions of dollars from Motorola before diverting and misappropriating those funds. In 2003, the District Court for the Southern District of New York awarded Motorola $2.1 billion in compensatory damages and noted that the Uzans were criminals who had “perpetrated a huge fraud.” Since 2003, Motorola has endeavored to enforce its award, as well as $1 billion in additional punitive damages, by serving subpoenas on non-party banks and attempting to attach Uzan-related property around the world.

In 2013, the district court ordered the restraining of Uzan assets anywhere in the world by anyone with notice of the order. Thereafter, Motorola served restraining orders and subpoenas on a variety of banks, including Standard Chartered Bank (“SCB”), a bank incorporated and headquartered in the United Kingdom, through service on its New York branch. Pursuant to the order, SCB searched and located Uzan assets in its United Arab Emirates branches. After SCB froze $30 million worth of Uzan assets, the U.A.E. central bank retaliated against SCB by debiting around $30 million from an SCB account therein. Likewise, the Central Bank of Jordan seized documents from SCB’s Jordan branch to punish the bank. These actions prompted SCB to seek relief from the restraining order and the subpoena, arguing that the separate entity rule

---

187. The separate entity rule is a creature of New York common law, providing that even if a bank is subject to personal jurisdiction in New York, “its other branches are to be treated as separate entities for certain purposes,” including prejudgment attachments and “postjudgment restraining notices and turnover orders.” As the Motorola court noted, the rule has been justified on three grounds: (1) the importance of international comity and respect for foreign sovereigns’ power over banks located in their countries; (2) the danger of double liability that banks may face at home and abroad; and (3) the high burden and cost of requiring banks to “monitor and ascertain the status of bank accounts in numerous other branches.” Motorola Credit Corp. v. Standard Chartered Bank, 21 N.E.3d 223, 226-27 (2014).

188. Id. at 149.

189. Id. at 156.

190. Id. at 156–57.

191. Id. at 157.

192. Id.

193. Id.
and international comity confined the restraining notice to its New York branch and not its foreign branches.\textsuperscript{194} Motorola moved to compel the requests.

In August of 2013, the district court sided with SCB and refused to compel the subpoenas because of, among other things, international comity. In its analysis, the court noted that the subpoenas implicated the criminal laws of Jordan and the U.A.E., which were being enforced to the detriment of SCB. Although the court rejected the applicability of the separate entity rule, it ultimately found that international comity weighed against the production of SCB documents in the U.A.E. and Jordan. This was a victory for international comity. Motorola appealed this decision, and the Second Circuit certified the question of the separate entity rule to the New York Court of Appeals.

A five-member majority of the Court of Appeals upheld the separate entity rule as a crucial part of New York common law. The court noted that “[c]ourts have repeatedly used it to prevent the postjudgment restraint of assets situated in foreign branch accounts based solely on the service of a foreign bank’s New York branch.”\textsuperscript{195} The majority concluded that a “judgment creditor’s service of a restraining notice on a garnishee bank’s New York branch is ineffective under the separate entity rule to freeze assets held in the bank’s foreign branches.”\textsuperscript{196}

In reaching this conclusion, the majority analyzed the history and purpose of the rule and emphasized that one of the most important justifications for the rule was “the importance of international comity” and foreign laws and regulations.\textsuperscript{197} The majority observed that the same justifications that led to the creation of the rule continued to resonate, including the avoidance of “competing claims and the possibility of double liability” and “the practical constraints and costs associated with conducting a worldwide search for a judgment debtor’s assets.”\textsuperscript{198} The majority thus rejected Motorola’s argument that new technological developments rendered the rule anachronistic.

In the context of discovery and comity, the court made various relevant findings that follow directly from \textit{Daimler}. First, to emphasize that it was following the Supreme Court, the New York Court of Appeals cited \textit{Daimler} as “recognizing the importance of considering ‘the risks to international comity’” and supporting the proposition that “the separate entity rule promotes international comity and serves to avoid conflicts among competing legal systems.”\textsuperscript{199} Second, the court noted the costs of worldwide discovery, commenting that “courts have continued to recognize the practical constraints and costs associated with conducting a worldwide search for a judgment debtor’s assets.”\textsuperscript{200} The prospect of burdensome discovery and its implications

\textsuperscript{194}. \textit{Id.}
\textsuperscript{195}. \textit{Id.} at 162.
\textsuperscript{196}. \textit{Id.} at 163.
\textsuperscript{197}. \textit{Id.} at 159.
\textsuperscript{198}. \textit{Id.} at 159–62.
\textsuperscript{199}. \textit{Id.} at 162.
\textsuperscript{200}. \textit{Id.}
for international comity was an important factor animating the decision. Third, the court recognized that the rule provided benefits to international banks and to New York’s “status as the preeminent commercial and financial nerve center of the Nation and the world.”

Finally, the majority noted that in this specific case SCB faced clear repercussions in Jordan and the U.A.E., placing it in an impossible situation and risking double liability. In sum, Motorola’s holding, much like Daimler and Gucci, was based on practical considerations of international comity, the global economy, and the United States’ place in it.

Writing for the dissent, Judge Abdus-Salaam echoed Aérospatiale’s embrace of broad U.S. power over foreign parties and defended broad jurisdiction as appropriate in the modern world where discovery should not be burdensome. After laying out its view of the world as different than the majority’s, the dissent dismissed the court’s concern with comity as “akin to using a cannon to kill a fly” because many countries did not have conflicting laws and thus a case-by-case approach would be more appropriate. Judge Abdus-Salaam otherwise criticized the separate entity rule as anachronistic and misguided. First, the dissent complained that the decision allowed the criminal Uzan family to evade enforcement proceedings in New York and shielded judgment creditors who could “make a mockery of our courts’ duly entered judgments.”

A general fear that judgments will go unenforced is a common concern among courts that support broad U.S. discovery. Second, the dissent relied on statutory construction of New York’s Civil Practice and Rules, concluding that under Section 5222 foreign bank branches were not exempt from complying with a restraining notice.

Third, Judge Abdus-Salaam emphasized that technology had rendered the rule obsolete because “[i]n this day of centralized banking and advanced technology, bank branches can communicate with each other in a matter of seconds.” Plaintiffs seeking broad discovery typically argue it is not as burdensome as defendants or non-parties claim it is. Finally, the dissent stressed that banks faced increasingly complex regulations and were not deterred from conducting business in New York and would thus adapt to the abolishment of the separate entity rule. In support of this conclusion, Judge Abdus-Salaam quipped that:

Banks have apparently adjusted to the societal expectation that they will be responsible corporate citizens, presumedly by using modern technology and a reasonable share of their resources to shoulder the burden of compliance with a

---

201. \textit{Id.} (internal quotation marks omitted).

202. It bears emphasis that the relevance of foreign criminal laws providing the prospect of double liability is one of the foundational reasons for the existence of international comity.

203. \textit{Motorola Credit Corp.}, 24 N.Y.3d at 164 (Abdus-Salaam, J., dissenting).

204. \textit{Id.} at 165–66. The dissent criticized the Separate Entity Rule as a “judicially created doctrine” that is not “tethered to the CPLR’s text.” Because the CPLR did not limit the reach of restraining notices, the dissent concluded that restraining notices applied abroad and the Separate Entity Rule should therefore be “rejected, not embraced.”

205. \textit{Id.} at 167.
This sentence encapsulated the dissent’s theory: banks will adapt and comply, as they have with other regulations, as long as courts require them to do so. The dissent concluded that the rights of judgment creditors outweighed the concerns of foreign banks.

On the whole, *Motorola* is a remarkable case that explores two vastly different views of the modern globalized economy and the United States’ place in it. The majority strained to maintain New York’s privileged position as the financial capital of the world while the dissent dismissed these concerns as overblown and irrelevant. In so doing, *Motorola* relied on *Daimler* and Gucci’s renewed appreciation for international comity, emphasizing respect for foreign countries and the need to limit the uninhibited nature of U.S. judicial power. *Motorola* is crucial because it will impact the way banks operate in New York, and it signals the extension of *Daimler* into matters of state law.

* * *

In conclusion, *Daimler*, Gucci, and *Motorola* represent a new paradigm of respect for international comity. These cases signify a break from the past because they made international comity a prominent reason for refusing to hear cases implicating foreign interests. These cases also weakened the *Aérospatiale* paradigm that relied on personal jurisdiction and control. Although many other courts had considered international comity in the past—through a pretense balancing test—no three cases from such prominent courts had so thoroughly and so quickly linked respect for other sovereigns to decisions about jurisdiction and discovery.207

One could argue that *Daimler*, Gucci, and *Motorola* are outliers because of their unique facts: the involvement of the Argentine government and a German corporation, the state-owned Bank of China, and the threat of actual criminal punishment confronting SCB in other countries. Yet the *Aérospatiale*-era was littered with cases involving even greater foreign interests. For example, in *Bodner v. Paribas*, plaintiff Holocaust survivors and defendant French Banks implicated the deep and historical interests of a variety of European countries.208 Similarly, *Wiwa v. Royal Dutch Petroleum Co.* involved torture, imprisonment, and murder by Nigerian government officials and English and Dutch oil companies, somewhat akin to the facts in *Daimler*.209 But in both of those cases

206.  *Id.* at 169.

207.  Recently, in *Motorola Credit Corp. v. Uzan*, No. 02 Civ. 666(JSR), 2015 WL 5613077, at *3 (S.D.N.Y. Sept. 9, 2015), Judge Rakoff, perhaps recognizing the new paradigm, granted a protective order in favor of two banks seeking to avoid discovery in violation of Swiss law. Similarly, the district court for the Northern District of Illinois refused to order two foreign banks to produce documents located abroad due to concerns with international comity, holding that “the interests of international comity weigh against ordering these foreign non-party banks to comply with Plaintiffs’ broad discovery requests.” *Leibovitch v. Islamic Republic of Iran*, No.08 C 1939 (N.D. Ill. May 19, 2016) Dkt. 203.


district courts ignored the overriding importance of international comity. Moreover, much like *Gucci*, *Aérospatiale* itself and various cases thereafter involved state-owned companies,210 and, like *Motorola*, foreign countries have threatened to punish companies who produce documents in U.S. courts.211 There simply is no clear way to factually distinguish these cases.

It seems likely that the newfound respect for international comity will revive the Hague Convention.212 The *Aérospatiale* paradigm relied on two pillars: control and personal jurisdiction. *Daimler* and *Gucci* have severely limited the existence of general personal jurisdiction over foreign companies with affiliates in the United States. As long as the foreign companies involved in any particular case are based and registered in foreign countries, they are not “at home” in the United States and cannot be subjected to general jurisdiction.213 Therefore, barring an extraordinary expansion of specific jurisdiction or the concept of “consent” to general jurisdiction, the *Aérospatiale* paradigm has been weakened.214 Accordingly, the Hague Convention will, at the very least, become

212. Requests for documents held abroad can be targeted at (1) foreign companies with U.S. affiliates or at (2) U.S. companies with foreign affiliates. These cases directly attack the first kind of request, and indirectly weaken the second kind. In either case, the Hague Convention should be substantially strengthened.
213. Gucci Am., Inc. v. Weixing Li, 768 F.3d 122, 135 (2d Cir. 2014).
214. An expansion of specific jurisdiction is possible, and perhaps likely. But the extent of any such expansion is unclear. If the target of a discovery request is a defendant and they are found to have “purposefully directed” their activities at the forum, then why would the documents be located abroad? This could happen in certain cases but it should not be the norm. Similarly, if the target of a request is a non-party, then surely plaintiffs cannot argue that their claims are based on the foreign company purposefully directing its activities at the forum and that litigation arises out of or relates to those activities. Moreover, in either case courts will have to evaluate the “fair play and substantial justice” necessary for any assertion of jurisdiction and, as explained above, may often find that this factor is not satisfied.

On the other hand, on September 29, 2015, the *Gucci* district court, on remand from the Second Circuit, found specific jurisdiction over Bank of China records in China related to New York transactions. Gucci Am., Inc. v. Weixing Li, No. 10 Civ. 4974(RJS), 2015 WL 5707135, at *7 (S.D.N.Y. Sept. 29, 2015). The court held that the Bank of China had purposefully availed itself of New York by establishing an office there, owning various properties, initiating lawsuits, and generally doing business therein. As such, Bank of China gained access to “New York’s dependable and transparent banking system, the dollar as a stable and fungible currency, and the predictable jurisdictional and commercial law of New York and the United States.” Id. at *3. The court also rejected comity concerns, but mostly because it had already conducted a pre-*Daimler* comity analysis. Id. at *11–12. It is unlikely that other courts will follow this precedent to order the production of documents held abroad because the documents at issue in *Gucci* were directly related to the litigation and had, at one point, touched New York. See also Strauss v. Credit Lyonnais, S.A., 2016 WL 1305160, at *17-19 (E.D.N.Y. Mar. 31, 2016) (accepting a broad theory of specific jurisdiction).

With regards to consent to general jurisdiction, although some courts have been receptive to arguments that foreign corporations consent to jurisdiction when they register to do business in a particular state, the Second Circuit recently rejected this theory. Brown v. Lockheed Martin Corp.,
necessary in cases dealing with foreign corporations with a presence in the United States. If Daimler and Gucci are applied consistently by lower courts in transnational discovery disputes, litigants will have to familiarize themselves with the requirements of the Hague Convention in order to request documents held abroad, as it may be the only vehicle available.215

Looking to the future, if Hague Convention procedures are invoked more often, this could initiate a self-reinforcing mechanism whereby growing familiarity with the Convention encourages more litigants and courts to employ it. As legal actors become more familiar with the Convention, they might develop the institutional knowledge necessary to make applying the Convention a smoother and more routine process. That would, in turn, encourage further use of the Convention to the detriment of the Federal Rules. This process might even catalyze the development of a new treaty that could replace the Convention and become even more useful. In short, the cases and doctrines described above may have far-reaching consequences. 216

IV. ECONOMICS AND DIPLOMACY: THE MAIN FACTORS BEHIND DAIMLER, GUCCI, AND MOTOROLA

If Daimler, Gucci, and Motorola represent a change in the law, as this Article has thus far argued, then it is worth exploring the reasons behind such a change. There are two common elements in these decisions that are worth briefly discussing here: (1) the need for reciprocity in international relations and the danger of retaliatory laws, and (2) the importance of international comity to international trade and the global economy.


215. Gap, Inc. v. Stone Int’l Trading, Inc., No. 93 Civ. 0638 (SWK), 1994 WL 38651, at *1 (S.D.N.Y. Feb. 4, 1994) (“As a practical matter, in many cases the Hague Convention provides the only means to request documents or testimony from foreign non-parties over whom the court has no personal jurisdiction and who are beyond the subpoena power of the court.”). See Torreblanca de Aguilar v. Boeing Co., 806 F. Supp. 139, 144 (E.D. Tex. 1992) (case dismissed on forum non conveniens grounds where nonparty witness beyond subpoena power of court and Hague Convention thus provided only means of obtaining testimony), aff’d, 11 F.3d 55 (5th Cir. 1993).

216. On the other hand, requests for documents held abroad by American companies face an uncertain future. Although “personal jurisdiction and control” will almost always be present, the requests will still involve the interests of foreign countries in the litigation. Motorola may be instructive in this regard as it shows that Daimler means more than just a limitation on general jurisdiction; it also means greater respect for international comity. Motorola involved the freezing of a British bank’s property in the U.A.E. and Jordan. Although it was not a Jordanian or Emirates company, these countries were still deeply interested in the litigation and the New York Court of Appeals recognized that fact. Motorola therefore indicates that Daimler’s consequences will reverberate in the context of American companies with documents abroad. Courts will have to conduct a rigorous analysis of the foreign interests at stake before deciding, as they have for thirty years, to order production.

http://scholarship.law.berkeley.edu/bjil/vol34/iss1/1
1. Reciprocity, Retaliation, and Foreign Relations

Concerns with retaliation and reciprocity animated the three decisions. Foreign companies often argue in front of U.S. courts that exercising overbroad jurisdiction will cause foreign sovereigns to treat U.S. companies the same way. If a court orders production of documents by a French company in New York, the thinking goes, then a French court is more likely to order transnational discovery against a U.S. company in Paris. 

Daimler, Gucci, and Motorola took this argument into serious consideration and followed Kiobel’s advice that courts should be wary of the “danger of unwarranted judicial interference in the conduct of foreign policy.”

Echoing one of the most important justifications for international comity, the defendant in Daimler argued that broad U.S. jurisdiction would “encourage foreign nations to enact retaliatory jurisdictional laws that threaten U.S. companies with subsidiaries abroad.” This was not an empty claim. It was substantiated by Ninth Circuit Judge O’Scannlain’s finding that other countries had already enacted retaliatory laws against the United States. Indeed, renowned academic Gary Born found that:

[S]everal civil law countries have enacted “retaliatory” jurisdictional provisions. These provisions empower national courts to exercise jurisdiction over foreign persons in circumstances where the courts of the foreigner’s home state would have asserted jurisdiction. For example, Belgian domiciliaries can bring actions in Belgian courts against foreign defendants if they can demonstrate that the courts of the foreigner’s domicile would entertain a comparable action against a Belgian defendant. Likewise, Italian courts will exercise jurisdiction over actions by Italian nationals against foreigners, provided that the foreigner’s courts would entertain claims against Italians in like circumstances. Austria and Portugal also have comparable retaliatory statutes.

Taking these concerns into account, the Supreme Court explicitly noted the Solicitor General’s statement that foreign sovereign objections to broad general jurisdiction had “in the past impeded negotiations of international agreements on the reciprocal recognition and enforcement of judgments.” Even plaintiffs in Daimler recognized in their brief that “international opposition to American legal tradition” caused “international friction.” In short, the danger of retaliatory laws or impediments to international negotiations was an important factor for the Supreme Court in Daimler.

---

219. Id.
In Gucci, the Second Circuit was similarly concerned about the consequences of its ruling for China-U.S. relations. Chinese regulators submitted a letter ominously warning that the case would impact “the conduct of China-U.S. relations, and, in particular, our countries’ Strategic and Economic Dialogue.”223 The court took this threat seriously and, as a result, invited the U.S. government to weigh in on the case.224 In response, the government urged the court to take greater account of China’s sovereign interests and criticized the district court for not doing so.225

Likewise, the Motorola court did not have to consider a hypothetical punishment under foreign law because Jordan and the U.A.E. had already proven their willingness to challenge U.S. discovery rules when they held SCB criminally liable under local law. The New York Court of Appeals specified that one of the main reasons the separate entity rule, itself an embodiment of international comity, was applicable in the case was because SCB faced repercussions abroad.226 Ordering the bank to comply with the turnover order would have placed it in the difficult position of obeying conflicting rules by three sovereigns—exactly the kind of situation the separate entity rule and international comity principles attempt to avoid.

In sum, concerns of reciprocity and retaliation played important roles in all three decisions. The important feature to notice is that all three courts dealt with these concerns in the same way—by emphasizing the importance of international comity and foreign sovereigns’ interests in the litigation. Unlike the Aérospatiale-era, these three courts took foreign interests seriously, keeping in mind the Kiobel advice that courts should avoid “clashes between our laws and those of other nations which could result in international discord.”227 Although many courts during the Aérospatiale-era paid lip service to this idea, they often concluded that U.S. interests were more important and justified broad international jurisdiction and discovery. Daimler, Gucci, and Motorola are revolutionary in concluding the opposite, that foreign interests are important enough to limit U.S. general jurisdiction, discovery, and restraining notices.

2. The International Economy

International comity has always been, to a large extent, about the international economy and commercial system. Justice Blackmun emphasized in

223. Letter from Huai Peng Mu, supra note 2.
226. Motorola Credit Corp. v. Standard Chartered Bank, 21 N.E.3d 223, 229 (2014) (“Indeed, as the District Court observed, the facts of this case aptly demonstrate that the policies implicated by the separate entity rule run deeper than the ability of a bank to communicate across branches. In seeking to comply with the restraining order, SCB faced regulatory and financial repercussions abroad.”).
his Aérospatiale dissent that the Convention “serves the long-term interests of the United States in helping to further and to maintain the climate of cooperation and goodwill necessary to the functioning of the international legal and commercial systems.”

Commercial interdependence across the globe is one of the most important factors behind the need for international legal harmony. This is why the U.S. Chamber of Commerce has prudently noted that “[a]s globalization advances, the potential for conflicting judicial processes to harm U.S. commercial and other interests increases, and in reaction the United States has consistently promoted international comity as a means to harmonize international legal systems.”

The Supreme Court recently agreed with this spirit, stating that taking account of the “legitimate sovereign interests of other nations . . . helps the potentially conflicting laws of different nations work together in harmony—a harmony particularly needed in today’s highly interdependent commercial world.”

Courts generally recognize the interaction between global commerce and international comity but have only recently begun to view transnational discovery in this context. Just a few years ago, one court expressly recognized that “transnational discovery requests are increasing due to the global nature of ‘international commerce.’”

Beyond diplomacy and paying heed to the concept of “harmony,” international comity affects the country’s bottom line: the economy.

In Daimler, Gucci, and Motorola, the foreign parties highlighted the cases’ possible repercussions on international commerce. The three courts responded by finally affirming the importance of international comity in this context. In Daimler, defendant’s brief highlighted that Judge O’Scannlain had sought a rehearing en banc because the Ninth Circuit’s decision could “have unpredictable effects . . . on our nation’s economy.”

The defendant also warned that foreign corporations would withdraw their investments out of the United States and that this could cause possible damage to “U.S. consumers and the U.S. economy.”

Justice Sotomayor explicitly recognized these arguments, noting in her concurrence that “[w]hat has changed since International Shoe is not the due process principle of fundamental fairness but rather the nature of the global economy.” It seems that Daimler, Gucci, and Motorola recognized
that, by extension, what has changed since *Aérospatiale* is the nature of the
global economy.

*Gucci* and *Motorola* also faced these concerns. In *Gucci*, Chinese
regulators wrote a letter outlining that cooperation was crucial “in the arenas of
banking and finance.” Broad U.S. jurisdiction over the BOC threatened this
stability and the U.S.-China economic dialogue. Plaintiff argued that Chinese
data protection laws were part of China’s strategy in developing its financial
system “which is of no small importance to the world economy.” The U.S.
amicus even asked the court to weigh the interests of China in opposing the asset
freeze in accordance with a wider balancing test provided by Section 403 of the
Restatement (Third) of Foreign Relations Law, which takes into account “the
importance of the regulation to the international political, legal, or economic
system.” The Second Circuit duly obliged, ordering the district court to use
that test.

*Motorola* prominently featured a profound debate between the majority and
the dissent discussing New York’s place as the financial capital of the world.
The majority noted that “[u]ndoubtedly, international banks have considered the
[separate entity rule’s] benefits” when opening branches in New York, which
“has played a role in shaping New York’s status as the preeminent commercial
and financial nerve center of the Nation and the world.” The court also
emphasized that it believed abolition of the rule “would result in serious
consequences in the realm of international banking to the detriment of New
York . . .” In espousing these ideas, the majority evinced a profound concern
for New York’s economy and its relationship with international comity. In
contradistinction, the dissent, embracing the style of *Aérospatiale* progeny,
emphasized repeatedly the rights of U.S. plaintiffs to enforce their judgments,
concluding that “any burden imposed on the banks is far outweighed by the
rights of judgment creditors to enforce their judgments.” Moreover, the
dissent rejected any arguments that banks relied on the rule, noting that they
faced increasingly complex government regulations and “[y]et banks continue to
do business in this country.” The dissent went on to emphasize that the U.S.
government had recently imposed severe fines on foreign banks, but nonetheless
banks had “adjusted to the societal expectation that they will be responsible
corporate citizens.”


11-3934-cv), 2013 WL 1790984, at *37.

237. Brief for United States as Amicus Curiae, Gucci Am., Inc., v. Bank of China, 768 F.3d
122 (2d Cir. 2014) (Nos. 11-3934), 2014 WL 2290273, at *20.

citations and quotations omitted).

239. Id. at 163.

240. Id. at 170.

241. Id. at 168.

242. Id. at 234.
Ultimately, the *Motorola* dissent rejected the majority’s concern with international comity as overbroad because it protected restraining assets in countries without conflicting laws. However, an overbroad rule of comity seems preferable to one that is underinclusive. As shown *infra* Part II, the *Aérospatiale* paradigm was insufficient to quell foreign-sovereign concerns with U.S. discovery. It was also inadequate for the modern international economy. The *Motorola* majority’s take on international comity sought to foster an international environment that promotes cooperation and international trade. The dissent did not truly take this into account. It instead focused on the hardship to American plaintiffs. But this ignores the fact that American plaintiffs seeking to enforce their judgments abroad will not suffer significant difficulties as they may use local procedures.

These two views of the global economy and international comity embraced by the majority and the dissent could not have been more different. The majority pinpointed economics as the most important factor in the case and saw the court’s role as the protector of New York’s position as the financial capital of the world. On the other hand, the dissent approached the case from the point of view of litigants seeking to enforce their judgments who nonetheless face difficulties due to state law artificial restraints. One point to note is that this debate over economics was inextricably linked to value judgments about the importance of litigant rights and state interests, and involved deep questions of international comity and its relationship to global economics.

In conclusion, *Daimler*, *Gucci*, and *Motorola* challenge the *Aérospatiale* paradigm of international comity partly because of the necessities of the modern global economy. Taking this factor into account, the New York Court of Appeals saw the separate entity rule as necessary for maintaining New York’s place as the global financial center. The Supreme Court in *Daimler* recognized the wide-ranging effects on the U.S. economy and global trade of an “uninhibited” rule of general personal jurisdiction. Likewise, the Second Circuit saw the danger of upsetting the crucial U.S.-China economic relationship. These three cases did not deal with exceptional facts. It is the reality of the modern global economy that cases dealing with jurisdiction and discovery will inevitably involve important foreign interests. In the face of such a reality, a broad rule of international comity is necessary.

V.

**HOW DAIMLER SHOULD RESHAPE COMITY: IMPOSING LIMITS ON U.S. INTERESTS AND REQUIRING INPUT FROM FOREIGN SOVEREIGNS**

The downfall of broad general jurisdiction over foreign parties with affiliates in the United States and the renewed respect for international comity will have a deep impact on future case law. It will affect the way courts evaluate international comity concerns in the context of transnational discovery and the Hague Convention. It seems clear that courts should take heed of *Daimler*, *Gucci*, and *Motorola* when evaluating international comity and should no longer
shirk their international comity responsibilities by reciting the elements of a pretense test that almost always allows courts to order discovery of documents held abroad. *Daimler*, *Gucci*, and *Motorola* require a more serious and rigorous consideration of foreign countries’ interests.

As explained *infra* Part II, at least two arguments have been leveled against the *Aérospatiale* paradigm: (1) the *Aérospatiale* balancing test overvalues U.S. interests; and (2) U.S. courts are unable to adequately address the interests of foreign nations because judges have little experience with foreign laws or international law. These are legitimate concerns with the way U.S. courts treat international comity. Scholars and judges writing before *Daimler* proposed improvements to the comity analysis in ways that still resonate today. One way to support the changes heralded by *Daimler* is to alter the existing balancing test towards a presumption in favor of the Hague Convention.243 Another way would be to shift the burden of persuasion on plaintiffs to prove that the Convention should not be employed.244 Although both of these approaches are appropriate, they do not tackle the overvaluing of U.S. interests responsible for the excesses of the *Aérospatiale*-era.

For that reason, this Article advocates a third, more novel approach: courts should engage in a qualitative limitation on the kinds of U.S. interests that are significant and should require frequent input from foreign sovereigns.245 As

243. Société Nationale Industrielle Aérospatiale v. U.S. Dist. Court for S. Dist. of Iowa, 482 U.S. 522, 547 (1987) (Blackmun, J., concurring in part and dissenting in part). This is the approach advocated by Justice Blackmun’s dissent in *Aérospatiale*. This approach would greatly aid the adoption of Hague Convention procedures and will limit the uninhibited exercise of U.S. power. Such an approach would limit the obstacles currently in place on the use of the Hague Convention and would influence the balancing test in favor of the Convention. Justice Blackmun’s presumption would shift this burden and force plaintiffs, instead of defendants, to prove that the Federal Rules should trump the Convention.

244. This burden shifting scheme has been proposed before, Matthew B. Kutac, *Reallocating the Burden of Persuasion Under the Aérospatiale Approach to Transnational Discovery*, 24 REV. LITIG. 173, 199 (2005). Cases that ratified the current burden scheme: *In re Auto. Refinishng Paint Antitrust Litig.*, 358 F.3d 288, 305 (3d Cir. 2004); *Wultz v. Bank of China Ltd.*, 910 F. Supp. 2d 548, 552 (S.D.N.Y. 2012); *In re Vitamins Antitrust Litig.*, 120 F. Supp. 2d 45, 51 (D.D.C. 2000); *Valois of Am., Inc. v. Risdon Corp.*, 183 F.R.D. 34, 346 (D. Conn. 1997); *In re Perrier Bottled Water Litig.*, 138 F.R.D. 348, 354 (D. Conn. 1991). See also *In re Vivendi Universal, S.A. Sec. Litig.*, 02-CIV-5571(RJH)(HBP), 2006 WL 3378115, at *2 (S.D.N.Y. Nov. 16, 2006). Arguably, plaintiffs are in a better position than defendants to meet this burden, as they know the kind of information they are seeking, the locations where they are seeking it, and are thus better equipped to argue for one procedure over another. Plaintiffs should therefore meet their burden by specifically outlining why the Federal Rules provide a more efficient procedure in a particular case, and crucially, by showing that the foreign country at issue has cooperated in the past with such requests and is likely to do so in this case. Plaintiffs should also be required to evaluate the possible conflict of laws and outline why U.S. procedures should nonetheless overcome foreign law. The most important effect of such a requirement would be to incentivize plaintiffs to truly consider the Hague Convention before launching into broad subpoenas and requests under U.S. law. This would, in accord with *Daimler* and its progeny, limit the raw and uninhibited exercise of U.S. courts’ power.

245. I use the word “qualitative” here to refer to changes in the substantive scrutiny of the current comity test. In short, I argue that courts should improve the test by changing the quality of their legal analysis. This “qualitative” argument can be distinguished from a second argument I offer below: that courts should limit the number of interests recognized by the comity analysis.
many courts have noted, the “most important” factor in the comity analysis is the balance of national interests. Targeting this factor could ameliorate the problem of overvaluing U.S. interests above foreign laws and is a particularly attractive way to nudge U.S. law towards a more international approach, as it does not necessitate a broad makeover of current law.

A. The Problems with the Categorical Approach that Overvalued U.S. Interests

The single most important problem with the Aérospatiale paradigm was the pretense of a balancing test, which almost always discounted foreign interests in favor of U.S. judicial power. In his Aérospatiale dissent, Justice Blackmun correctly predicted this development, noting:

> Experience to date indicates that there is a large risk that the case-by-case comity analysis now to be permitted by the Court will be performed inadequately and that the somewhat unfamiliar procedures of the Convention will be invoked infrequently. I fear the Court’s decision means that courts will resort unnecessarily to issuing discovery orders under the Federal Rules of Civil Procedure in a raw exercise of their jurisdictional power. . .

Justice Blackmun did not see the three decades of case-by-case comity evaluations that followed Aérospatiale, where courts invariably affirmed the Federal Rules. However, he understood that a balancing test was inappropriate in this context because courts would, even if subconsciously, always favor U.S. interests and an exercise of the courts’ “raw power.” A balancing test in this context was always misplaced because the question for courts was, in effect, whether to balance away their power in favor of foreign sovereigns. Given the expected effects of pro-forum bias, the test always weighed against comity. One court perfectly encapsulated this problem:

> Despite the real obligation of courts to apply international law and foster comity, domestic courts do not sit as internationally constituted tribunals. Domestic courts are created by national constitutions and statutes to enforce primarily national laws. The courts of most developed countries follow international law only to the extent it is not overridden by national law. Thus, courts inherently find it difficult neutrally to balance competing foreign interests. When there is any doubt, national interests will tend to be favored over foreign interests. This partially explains why there have been few times when courts have found foreign interests to prevail.

---

248. Id. at 553 (Blackmun, J., concurring in part and dissenting in part) (“pro-forum bias is likely to creep into the supposedly neutral balancing process and courts not surprisingly often will turn to the more familiar procedures established by their local rules.”).
Sure enough, case law shows that courts are unwilling to reduce their own power, particularly in the context of discovery, where court supervision can determine the evidence to be used in an entire case.\textsuperscript{250}

The effect of pro-forum bias reached such heights that before \textit{Daimler}, courts had nearly stopped their exacting examination of U.S. interests and instead had developed certain categories of U.S. interests that automatically overcame concerns for foreign laws. These categories included cases involving patent law,\textsuperscript{251} antitrust law,\textsuperscript{252} criminal law,\textsuperscript{253} and anti-terrorism laws.\textsuperscript{254}

\textsuperscript{250}. See notes 252–62 for cases where courts rejected international comity.


\textsuperscript{254}. E.g., Linde v. Arab Bank, PLC, 706 F.3d 92, 112 (2d Cir. 2013) ("[W]e find no clear abuse of discretion in the District Court’s conclusion that the interests of other sovereigns in enforcing bank secrecy laws are outweighed by the need to impede terrorism financing as embodied in the tort remedies provided by U.S. civil law and the stated commitments of the foreign nations."); Wultz v. Bank of China Ltd., 910 F. Supp. 2d 548, 556 (S.D.N.Y. 2012) ("But in light of the significant U.S. interest in eliminating sources of funding for international terrorism, and the other factors discussed below, the law governing discovery disputes in this case must ultimately be the broad discovery rules of the Federal Rules of Civil Procedure."); Rubin v. Islamic Republic of Iran, No. 03 C 9370, 2008 WL 192321, at *20 (N.D. Ill. Jan. 18, 2008) (“Congress has also decided that the ‘grace and comity’ generally extended to foreign sovereigns should be limited in specific ways, particularly when those sovereigns promote terrorist acts that injure U.S. nationals."); Strauss v. Credit Lyonnais, S.A., 242 F.R.D 199, 214 (E.D.N.Y. 2007) (U.S. interest in “combating
Although these categories concerned legitimate interests, this Article attempted an exhaustive review of all court acknowledged U.S. interests and found that courts had developed wildly uninhibited categories, where U.S. interests were seen as paramount without much explanation. These interests included: “fully and fairly adjudicating matters before [U.S.] courts;” protecting U.S. nationals merely because they were involved in a case as defendants or vindicating the rights of American plaintiffs; ensuring the integrity of the terrorism”); Weiss v. Nat’l Westminster Bank, PLC, 242 F.R.D. 33, 47 (E.D.N.Y. 2007) (“Executive and Congressional interests in freezing terrorist financing.”).


financial markets or securities laws; 258 personal injury or products liability cases; 259 the administration of bankruptcy cases; 260 and ad hoc categories, such as “assuring restitution to holocaust victims,” “regulating the economy,” “ensuring that [U.S.] laws are enforced,” and ensuring “the solvency of [American] insurance companies.” 261

Although courts could argue that all of these interests are compelling, two complicating consequences weaken that position: (1) there is no discernible limiting principle to such a laundry list of interests; and (2) similar foreign interests should be equally compelling. Lacking a proper limiting principle, the laundry list would continue to grow and swallow the Hague Convention (as it arguably did), rendering the entire comity analysis illusory. Extending the recognition of these interests to foreign nations—as courts would have to do to retain a semblance of balance—would otherwise lead courts back to step one: how to compare compelling sovereign interests. The longer the list of

---


compelling interests grows, the more difficult such a competing analysis would become. Because of this, it is much simpler, more predictable, more effective, and more loyal to Hague Convention treaty obligations to enumerate as few compelling interests as possible and avoid the complications of a laundry list.

The Aérospatiale paradigm utterly failed to do this. This laundry list of categories represented the excesses of Aérospatiale. In effect, it meant that no U.S. interest was speculative enough to overcome foreign law concerns. Such a haphazard categorical analysis even led the Seventh Circuit to comment that considering national interests seems a “ridiculous assignment.”

262. Reinsurance Co. of Am. v. Administratia Asigurarilor de Stat, 902 F.2d 1275, 1280 (7th Cir. 1990).
Admittedly, such a categorical approach can be efficient for lower courts as it allows them to fit the facts of each case into pre-developed groupings that have been determined to overcome foreign interests in the transnational discovery context. This approach has the added benefit of being consistent and predictable to a certain extent. However, lower courts took this approach too far by relying on inchoate categories and avoiding a scrutiny of the specific facts of each case. Relying on broad categories—for example, a U.S. interest in “fully adjudicating matters before U.S. courts”—can be unfair for defendants and foreign countries because it ignores the factual nuances of each case and overvalues U.S. interests. Moreover, loosely defined categories can become entrenched and allow courts to forfeit their duty to properly evaluate comity.

This table is not exhaustive. It includes cases found in Westlaw and Lexis using broad search terms designed to capture citations to the Hague Convention and Aérospatiale.
This is precisely what happened after Aérospatiale. Lower courts developed loosely defined yet insurmountable interests that unfairly ignored foreign laws and, in the words of Daimler, became “uninhibited.”

For example, an important problem with a categorical interest in “fully and fairly adjudicating matters before [the U.S.’s] own courts” is that, by definition, it forecloses any possibility of the United States not having a substantial interest in any case. This is so because the comity-balancing test is a feature of U.S. jurisprudence and therefore will always be evaluated in front of a U.S. court. If the United States always has an interest in fully and fairly adjudicating matters before its courts—an interest embraced by more than twenty-one cases—then its interests will almost always overcome foreign interests, thus rendering international comity a dead concept. Such an interest is wildly uninhibited.

In the same manner, an “interest in vindicating the rights of American plaintiffs” will settle any case involving an American plaintiff regardless of whether genuine national interests are involved or not. The United States could conceivably, given the right set of facts, prefer that an American plaintiff proceed through the Hague Convention. But an interest in vindicating the rights of American plaintiffs is so broad that it does not allow for a detailed analysis to accommodate other, perhaps more valid, national interests.

These are not hypothetical concerns; they are the natural consequence of the flawed categorical analysis that district courts conducted for thirty years. This wildly uninhibited approach to discovery was embraced in In re Air Cargo Shipping,

264 where plaintiff sued Société Air France (“Air France”) for antitrust violations due to an alleged price-fixing scheme. During discovery, plaintiff moved to compel Air France to produce information held in France. Defendants rejected this request because of the risk of criminal sanctions for violating the French blocking statute. Applying the comity balancing test, the district court noted that the balance of national interests was the “most important” factor and that the United States had “several strong national interests” implicated, namely, an interest in “antitrust laws whose enforcement is essential to the country’s interests in a competitive economy” and “a substantial interest in fully and fairly adjudicating matters before its courts.”

265 With those two generalities uttered and nothing more, the court found it necessary to compel the production of documents held in France.

266

Similarly, in Pershing Pacific West, LLC v. MarineMax, Inc., an American company brought claims against a German defendant, among others, for the faulty manufacture of a yacht.

267 When plaintiff demanded the production of documents related to the manufacturing process, defendant objected that

---

265. Id. at 54.
266. Id. at 55.
German law did not permit such discovery except through the Hague Convention. In its analysis of the national interests involved, the district court specified that "[w]e must assess the interests of each nation in requiring or prohibiting disclosure, and determine whether disclosure would affect important substantive policies or interests of either the United States or [Germany]." Then, however, the court engaged in a perfunctory three-sentence analysis of U.S. interests, noting that "[t]he United States has a[n] interest in vindicating the rights of American plaintiffs," and that "[t]his interest has been described as vital." Without any further scrutiny, the court then evaluated German law, which penalized the discovery at issue. Nevertheless, the court concluded that discovery should proceed through the Federal Rules because, among other things, the national interests “factor weighs slightly in favor of the United States’ interest and disclosure under the Federal Rules.” The court apparently did not find it necessary to note that the Hague Convention was an available “avenue” that could satisfy the sovereign interests involved.

_In re Air Cargo Shipping_ and _Pershing Pacific West_ are representative of a wider jurisprudence that engaged in a meaningless analysis of national interests lacking any rigor. The courts merely surmised that the antitrust violations at issue in _In re Air Cargo Shipping_ and the interests in vindicating the rights of American plaintiffs in _Pershing Pacific West_ involved actual U.S. interests in those particular cases. The _In re Air Cargo Shipping_ court made no findings regarding the specific interests of the United States in that particular case. Both courts were satisfied with making general unsubstantiated statements about national interests. Clearly, both courts failed to properly take international comity into account.

It is difficult to reconcile the existence of a balancing test and the Hague Convention with the perfunctory dismissal of foreign interests. These decisions exemplify the problems with _Aérospatiale_-inspired judicial chest-thumping, concerned only with “reaffirming the sovereignty of our judicial system.” Categorical interests have rendered the Convention inapplicable in almost all cases. Such an approach must change in light of _Daimler_ and its progeny.

_B. Daimler Demands a More Individualized Analysis of U.S. Interests_

The limits to U.S. jurisdiction recognized in _Daimler_, _Gucci_, and _Motorola_ and the renewed importance of international comity strike directly at the categorical analysis that allows an overvaluing of U.S. interests and the overbroad exercise of raw judicial power. The main problem noted by these courts was the “uninhibited” reach of U.S. jurisdiction, and by extension, discovery. Because of this, as explained below, _Daimler_, _Gucci_, and _Motorola_
teach that U.S. interests have to be specific to each case and cannot be based on speculative generalities.

In *Daimler*, the Ninth Circuit affirmed jurisdiction over a case that had no relationship to the United States, but instead involved foreign officials and a foreign corporation. The Supreme Court rejected the Ninth Circuit’s holding due to the threat to international comity posed by the Ninth Circuit’s “uninhibited approach to personal jurisdiction.” The Supreme Court did not speak in generalities about U.S. interests. Instead, it directly quoted the Solicitor General’s opinion that uninhibited personal jurisdiction had “in the past impeded negotiations of international agreements on the reciprocal recognition and enforcement of judgments.” This approach at least valued non-speculative interests. In *Gucci*, the lower court held the State-owned Bank of China in contempt for actions wholly unrelated to their New York branch. In reviewing this decision, the Second Circuit did not overvalue speculative U.S. interests. It instead asked the United States to express its interests in the case through an amicus. Likewise, in *Motorola*, the plaintiff sought to restrain assets of a British bank wherever located regardless of any foreign criminal laws penalizing such actions or of any jurisdictional concerns. The *Motorola* court described at length why the separate entity rule was incredibly important for New York.

Applying the teachings of *Daimler* means that U.S. interests need to be more than merely speculative or stated in broad terms. This can be done by following two steps: (1) limiting the categories included as “important U.S. interests” and (2) evaluating the interests of the United States in each specific case rather than through mere generalities. Courts may use these two methods to begin engaging in a proper and legitimate balancing of U.S. interests against foreign law concerns. These two methods can work in the following ways:

*Daimler* and its progeny embraced a standard that counsels the following: U.S. interests should be specific to the case and not just broad, uninhibited, and vague. Excessive interests, such as “the United States[s] . . . interest in fully and fairly adjudicating matters before its courts” or vindicating the rights of an


273. *Id.*

American plaintiff must be rejected. A good example of this constrained approach was embraced in Weiss v. National Westminster Bank, PLC, where survivors of terrorist attacks sued terrorist group Hamas, seeking assets held by U.K. based National Westminster Bank. When plaintiffs sought banking documents in the United Kingdom, defendant refused and demanded compliance with the Hague Convention. In its comity analysis, the court analyzed U.S. interests in-depth. After noting that the United States had an interest in "combating terrorism," it scrutinized how that interest applied against Hamas and related defendants (Interpal), noting that "[t]he American interest in disrupting terrorist networks with global assistance from American allies is particularly apparent here," because the Department of the Treasury had designated the particular groups involved in the case (Hamas charities) as terrorists. Moreover, the court highlighted that "not only does the United States have a demonstrated interest in halting terrorist financing, both domestically and internationally, but the United States has also explicitly found that NatWest’s client, Interpal, is a ‘principal’ conduit for those funds." This represented a laser-like focus on the defendant at issue and not on the general interest in combating terrorism.

The most important point to notice about Weiss is that the court narrowed the specific interests of the United States to those involved in that particular case. The court refused to rely on uninhibited platitudes about U.S. interests in combating “terrorism” generally, but instead focused on NatWest’s client, Interpal, who was involved in the case.

Another example of this salutary approach was adopted in Doster v. Schenk, where defendants claimed that subpoenas to produce information located in Germany were intrusive under Germany’s constitutional principle of proportionality “pursuant to which a judge must protect personal privacy, commercial property, and business secrets.” Those were speculative interests expressed in broad terms, much like those embraced by the Aérospatiale-era courts. The Doster court, however, rejected this argument because “[e]ven if the Court were to recognize those principles as significant sovereign interests, defendant must show that the specific discovery in these cases would compromise those interests” by a resort to the Federal Rules of Civil Procedure.

Courts should focus on the “specific discovery” requested in each case and how it would “compromise” U.S. interests. They should not answer these

---

275. See, e.g., In re Vivendi Universal, S.A. Sec. Litig., No. 02CIV5571RJH/HBP, 2006 WL 3378115, at *1 (S.D.N.Y. Nov. 16, 2006).
277. Id. at 47 (emphasis added).
278. Id. at 48.
280. Id. (emphasis added).
questions through generalities. This approach is important because, as some courts have recognized, “[U.S.] interests diminish the less closely a case is related to the United States.”281 A court can only determine how “closely a case is related to the United States” through an individualized analysis. For example, in a case involving antitrust laws, a court should not effuse that the enforcement of those laws “is essential to the country’s interests in a competitive economy.”282 That kind of general statement is uninhibited, ignores the particularities of the case, and fails to rigorously evaluate the interests at stake. Instead, courts should analyze why the United States has a specific interest in the particular case by referring to previous instances where the United States submitted amici in similar cases, or policy statements from the U.S. Justice Department antitrust division showing an interest in the particular antitrust violations. That is exactly what the court did in Weiss where it highlighted that the Department of the Treasury had designated the particular groups involved in that case as terrorists. Daimler and its progeny demand this approach.

Although employing a categorical analysis is not always inappropriate, it should only be the beginning of an international-comity scrutiny. Thus, for example, cases involving patents can be analyzed by courts through recognition that the U.S. has an important interest in patents generally; but, in addition, they should find that the patents involved in the case are relevant to the U.S. government. This method combines the benefits of a categorical approach with an individualized case-by-case analysis. A list of legitimate interests would include cases that involve U.S. criminal laws, antitrust laws, or terrorism laws, as well as U.S. government agencies or possible impact on a broad segment of the population. The most important benefit of this approach is that it forces courts to conduct a more individualized analysis that will improve respect for international comity.

One possible negative effect of employing a more specific analysis of U.S. interests is that it would require more specialized knowledge from courts and might weigh down dockets. To deal with these problems, this Article advocates a more proactive approach by the federal government in these cases. Courts and litigants should seek direct input from the government. As described below, this approach would be in line with prior case law and would be relatively easy to employ.

Courts have previously recognized that the United States has a “right to make its position known in cases with important foreign policy ramifications.”283 Some courts have even emphasized the importance of this type of input.284 Although it is not realistic to expect letters or amici from the

government in every case, both Gucci and Daimler show that when these kinds of input are present, courts are much better informed. Government input is beneficial because the executive is much better equipped than courts to evaluate the interests of the nation in each particular case. As noted by Justice Blackman in his Aérospatiale dissent, “[it] is the Executive that normally decides when a course of action is important enough to risk affronting a foreign nation or placing a strain on foreign commerce. It is the Executive, as well, that is best equipped to determine how to accommodate foreign interests along with our own.”285 The executive should, through letters and amici, supplement judicial determinations by emphasizing the interests of the United States in discovery cases. After Daimler, it is no longer legitimate to continue the judicial policy of determining U.S. interests without the input of the executive because, “[u]nlike the courts, diplomatic and executive channels are, by definition, designed to exchange, negotiate, and reconcile the problems which accompany the realization of national interests within the sphere of international association.”286

In short, whenever a comity issue arises in the context of discovery, courts should request that the executive intervene in some way.

Government intervention in discovery cases might dramatically burden the executive’s workload. However, there are various practical ways in which the executive can assert the government’s interest without significant added burden. This might include the submission of department letters of interest, re-used amici, policy papers, or regulatory statements. For example, as described above, the Weiss court relied on a Department of the Treasury designation of a terrorist group. The court did not need further input from the government. Below, this Article discusses other methods that are currently used by foreign governments to achieve this. Although this may be difficult in lower court proceedings, courts should experiment with requesting different types of input.

C. The Need for Input from Foreign Countries

For over two decades, courts have recognized the importance of direct input from foreign governments in transnational discovery cases.287 At various times, however, courts have also ignored letters from foreign ministers and ordered the production of documents located abroad despite foreign protests.288

Executive Branch has described the interests of the United States in this matter through its Statement of Interest . . . the general interests it describes also bear on the Court’s overall exercise of its discretion in connection with Plaintiffs’ discovery request.”).


286. Id. (quotations and citations omitted).


Daimler and its progeny urge courts to increase their openness to direct foreign input in transnational discovery cases.

There is little doubt that direct foreign input can strengthen a foreign party’s arguments. In *Gucci America, Inc. v. Curveal Fashion*, a case involving the production of documents located in Malaysia, the court noted that, “the Malaysian government has not voiced any objections to disclosure in this case, which the Second Circuit has found militates against a finding that strong national interests of the foreign country are at stake.”289 Similarly, in *In re Honda America Motor Co.*, the court recriminated the Japanese government for failing to intervene in the transnational discovery question at issue, highlighting that “[t]he failure of the Japanese government to weigh in as amicus curie on this matter is further evidence that its sovereignty is not implicated” in the case.290 These statements can be so important that some courts have declared that without foreign input it is difficult for courts to weigh sovereign interests at all. In *British International Insurance Co. v. Seguros La Republica, S.A.*, where plaintiff sought documents in Mexico, the court noted that “[t]he level of [sovereign] interest is difficult to gauge in this case since the Mexican government has not taken any steps to object to the discovery sought . . .”291

Requiring input from foreign sovereigns should be institutionalized to complement the reduced emphasis on U.S. interests. Courts should routinely require input from foreign countries, either through government officials or agencies. The U.S. amicus in *Gucci* advocated this approach, arguing that when the extraterritorial application of U.S. laws implicate foreign sovereign interests, “submissions from interested governments that address comity issues should be given serious consideration.”292 Requiring input from foreign governments serves three goals: (1) avoiding the “danger of unwarranted judicial interference in the conduct of foreign policy;”293 (2) tackling judges’ lack of experience with foreign law by giving them reliable information about those laws; and (3) continuing *Daimler*’s call for “inhibiting” U.S. judicial power.

At least one Supreme Court justice has explicitly voiced support for this approach. In *Republic of Argentina v. NML Capital, Ltd.*—involving the worldwide discovery of Argentine sovereign property under the Foreign Sovereign Immunities Act—Argentina argued during oral argument that foreign

---

countries opposed such discovery of Argentine property in their countries. Justice Scalia responded by asking,

> Why haven’t foreign countries protested? Why aren’t they here as amici? Is there a single foreign state that has taken your position? ... They file amicus briefs all the time and if this is as horrific as [Argentina is] painting it, we would have had some briefs from them.

Justice Scalia’s eagerness to hear from other countries should be shared by all courts. If foreign countries do not intervene through amici or letters, then courts should feel free to make adverse inferences. This would encourage countries to intervene whenever possible. In *Daimler*, the Supreme Court took special notice of the Solicitor General’s opinion in amici and European Union general jurisdiction rules. The Court quoted specific foreign law language provided by foreign organizations’ amici. In *Motorola*, the court also took special heed of Jordan and the U.A.E.’s reactions to the restraining notices.

These cases are exhorting courts to seek more information about foreign interests. Indeed, these cases seem to be saying that whenever a litigant faces requests for discovery of documents or property located abroad, they should find documentary support to counteract the request. Even if this means submitting amici from other cases or any other policy statement. Switzerland has been particularly active in this area, submitting amici even in district court proceedings. The current system relies on expert submissions by the parties, which the U.S. government took seriously in *Gucci*, arguing that the “*Gucci* [lower] court should have been more mindful ... and should not have summarily dismissed representations describing the national importance of China’s banking secrecy laws.” This can be supplemented by letters as China submitted in *Gucci*, where Chinese regulators expressed China’s objections to the overbroad discovery requests. Letters provide an efficient alternative to amici. Parties should try to obtain letters from local government officials that detail limitations under local law, express their preference for the Hague Convention, and describe problems with U.S. style discovery.

295. Id. at 22.
296. Residential Funding Corp. v. DeGeorge Fin. Corp., 306 F.3d 99, 106–07 (2d Cir. 2002) (stating that district court has “broad discretion” in dealing with breaches of discovery obligations, including the power to draw an adverse inference).
300. See Letter from Huai Peng Ma, supra note 2.
301. Even if courts welcome input from foreign sovereigns and begin to cabin U.S. interests, they may still have to use a balancing test. A great concern with *Aérospatiale* is that it
In sum, Daimler, Gucci, and Motorola counsel that discovery of documents abroad can no longer be “uninhibited.” A good starting place would be to qualitatively limit the kind of U.S. interests that are considered “substantial.” By focusing on the specific interests involved in each case and by soliciting foreign input whenever possible, courts can begin to honor their comity duties.

CONCLUSION

Daimler and its progeny suggest that the thirty-year Aérospatiale paradigm is coming to an end. The United States Supreme Court, the Second Circuit, and the New York State Court of Appeals have revived international comity with decisions that prominently denied the uninhibited exercise of raw judicial power over foreign parties. Influenced by concerns with international commerce and retaliation by foreign sovereigns, these courts have imposed limits on judicial power. Even if courts continue to reject the Hague Convention and alternatives to overbroad U.S. discovery, we should never again see the excesses of the

“regrettably. . . declined to set forth specific rules” for the international comity analysis. Scarminach v. Goldwell Gmbh, 531 N.Y.2d 188, 189 (1988). Adding to this difficulty is that lower courts felt compelled to analyze all of the factors expressed in the Restatement test, instead of engaging in a more qualitative analysis that can be both simpler but also more rigorous. Daimler is a good example of a simpler international comity analysis. The Supreme Court warned of the dangers to international comity and then analyzed only four factors: (1) European Union laws, (2) the Solicitor General’s views, (3) the Defendant’s arguments, and (4) the fair play and substantial justice concerns involved in the case. Daimler AG v. Bauman, 134 S. Ct. 746 (2014). This basic test is different from the current balancing test in that it is both simpler but more substantive. Instead of attributing speculative interests to the United States, the Court focused on a specific interest voiced by the executive in that particular context. Moreover, the Supreme Court took into account “fair play and substantial justice” in the context of international comity, not just as another element in a personal jurisdiction test. Id. at 763. This should, without doubt, be incorporated by lower courts.

A rigorous analysis of the fair play and substantial justice concerns of requiring foreign companies to produce documents in accordance with the Federal Rules would cut against the excesses of the Aérospatiale era. For example, the Second Circuit should overturn its finding in First Am. Corp. v. Price Waterhouse LLP, that a foreign non-party can be ordered to produce documents in the same manner as a litigant. Fair play and substantial justice after Gucci counsel a change of law in this context. As a general matter, courts should adopt the rule that “an order compelling production should be imposed on a nonparty . . . only in extreme circumstances.” Minpeco, S.A. v. Commodity Serv., Inc., 118 F.R.D. 331, 332 (S.D.N.Y. 1988). And this rule should be even more important in the context of foreign non-parties who may have no reason to expect being hauled to U.S. courts for actions wholly unrelated to their operations. Moreover, an emphasis on fair play and substantial justice would overturn the Third Circuit’s finding in In re Auto. Refinishing Paint Antitrust Litig., that the Hague Convention should not be employed for jurisdictional discovery. 358 F.3d 288 (3d Cir. 2004). If anything, international comity means that the Convention should always apply in cases of jurisdictional discovery where a court has not even established jurisdiction. Beyond this, courts should also incorporate the approach advocated by the U.S. amicus in Gucci that urged courts to take into account a wider balancing test provided by Section 403 of the Restatement (Third) of Foreign Relations Law, which evaluates “the importance of [a foreign] regulation to the international political, legal, or economic system.” Such an approach would allow courts to consider the possible effects of any case on the global economy and consider a wider array of foreign interests.
Aérospatiale-era which discarded the Convention as a useless treaty and affirmed U.S. transnational discovery rules in almost all cases.

If Daimler and its progeny are taken to their logical conclusion, the balancing of foreign interests should change dramatically. Courts should focus on the specific interests of the United States involved in each case, and should refrain from ruling based on mere generalities. Courts should also create a more routine procedure for acquiring executive and foreign input in particular cases. Finally, the comity balancing test should be reworked to emphasize this input from foreign countries and the U.S. government. Generally, lower courts should begin to experiment with ways to change the current comity analysis to accommodate the Daimler emphasis on international comity.

Taking everything into account, international comity in the discovery context is making a comeback. Following in the footsteps of Morrison and Kiobel, the U.S. judiciary seems ready to adopt a more diplomatic stance in the face of foreign interests. The changing nature of the modern global economy and the danger of retaliation by foreign countries should continue to influence these developments. What is clear is that U.S. courts are more attuned to international norms than they have been in the recent past, especially in the context of discovery.
Cooperating Alone: The Global Reach of U.S. Regulations on Conflict Minerals

Remi Moncel*

ABSTRACT

In 2010, the United States Congress adopted the Dodd-Frank Wall Street Reform and Consumer Protection Act. The Act includes an unprecedented provision to curb the mining in the Democratic Republic of the Congo (DRC) of so-called conflict minerals: components found in many consumer electronics that are sometimes the source of human rights abuses in the mines and regions from which they originate. Companies traded on the U.S. Stock Exchange are now required to conduct due diligence assessments of their supply chains and disclose the presence of such conflict minerals.

The mining of conflict minerals is a global problem for which international cooperation among States and companies seems the necessary solution. However, the United States acted alone; it unilaterally adopted regulations that focused on only one country—the DRC—and one set of targets—companies publicly traded in the United States. These regulations likely required less time to adopt and implement than traditional State-to-State cooperation. Critics might argue that conflict minerals originate not just from the DRC but also from other politically unstable nations, and companies publicly traded in the United States are not the only ones to integrate these minerals into their products. Yet, this Article argues that Dodd-Frank’s influence likely extends far beyond its stated geographical scope.

This Article is the first to ground the U.S. rules on conflict minerals in the literature on unilateral regulatory globalization. That literature posits that, under the right conditions, a country’s unilateral regulations can unleash a “California Effect” that causes companies outside its jurisdiction and other States to voluntarily align with those regulations. By analyzing the conflict minerals regulations through the lens of unilateral regulatory globalization, this Article reveals the Dodd-Frank Act’s potential to reach beyond its stated goals and

DOI: http://dx.doi.org/10.15779/Z388565

* J.D., 2015, U.C. Berkeley, School of Law. The author acknowledges with great appreciation Jamie O’Connell and Jerome Hsiang for their comments on earlier drafts, Katerina Linos for introducing me to concepts critical to this piece, and the editors of the Berkeley Journal of International Law for their diligent reviews and edits. Any remaining errors are mine.
enriches the existing literature by examining when regulations focused on business and human rights might trigger a California Effect.

Abstract

Introduction

I. Conflict Minerals: Overview of the Problem

A. Human Rights Violations

B. The Electronics Industry and the Market for Conflict Minerals


A. Overview of Dodd-Frank Conflict Minerals Regulations

B. Range of Possible U.S. Policy Responses

C. Implementation of Dodd-Frank

III. The Global Pull of an Enticing Market and Powerful Regulator

A. The Theory of Unilateral Regulatory Globalization

B. Can Dodd-Frank’s Conflict Minerals Provisions Unleash a “California Effect”?

1. Market Power

2. Regulatory Capacity

3. Preference for Strict Rules

4. Target Elasticity

5. Nondivisibility of Standards

6. Remaining Uncertainties

Conclusions

INTRODUCTION

The market continues to expand for consumer electronics, many of which contain metals partially sourced in conflict-rife zones. Armed factions, including those in the Democratic Republic of the Congo (DRC), control some of the mines that feed the global electronics market. These groups have committed human rights violations by exploiting workers in the mines and using the revenues to buy weapons and finance wars.

In many ways, “conflict minerals” present a familiar puzzle. Similar to garments, diamonds, oil, or coffee, the conflict minerals tin, tungsten, tantalum, and gold are globally traded. These raw materials originate in developing countries and end up in the consumer markets of wealthier nations. In many cases, the mining and harvesting of raw commodities takes place under politically unstable regimes. Large companies headquartered in wealthier countries rely on other corporate entities along their supply chains to source the minerals, integrate them into their products, and sell them to the consumer base.
Despite the similarities between conflict minerals and other raw materials, this Article examines one intriguing difference: the policy response to conflict minerals departs from traditional approaches to address challenges at the intersection of business and human rights. For example, labor rights violations in the agricultural sector and garment industry in developing countries have led nongovernmental organizations (NGOs) to develop fair trade certification schemes and governments to push for the implementation of the International Labour Organization’s Core Conventions. In response to concerns regarding the corruption in extractive industries, such as oil and gas, governments and stakeholders have established a voluntary reporting mechanism under the umbrella of the Extractives Industry Transparency Initiative (EITI). To constrain the trade in “conflict diamonds,” governments, industry, and NGOs have developed a global certification scheme through the Kimberley Process. These conventional approaches are not as prominent in the conflict minerals movement.

The Dodd-Frank Wall Street Reform and Consumer Protection Act (Dodd-Frank) takes a different approach. This U.S. policy is an isolated State response, not an interstate initiative. In addition, the U.S. Congress targeted only one region, the DRC and its neighbors, even though the African Great Lakes Region is not the sole source of production of conflict minerals. This U.S. conflict minerals rule seems to represent a new mode of intervention against human rights and sustainability challenges in global supply chains. Dodd-Frank’s drafters had ambitious yet limited goals: while aiming to curb the trade in conflict minerals, they focused on minerals from only one region, and they required only companies trading on a U.S. Stock Exchange to disclose the content of their supply chain. In a sense, the United States “cooperated alone.”

1. See generally Raluca Dragusanu et al., The Economics of Fair Trade, 28 J. ECON. PERSP. 217 (2014) (discussing the mechanisms of various fair trade standards and whether they in fact improve the working conditions of farmers in developing countries).
6. The African Great Lakes Region is the area surrounding Lake Victoria, Lake Tanganyika, and nearby smaller lakes. The countries generally considered part of this region are the DRC, Burundi, Rwanda, and Uganda. See generally About the Great Lakes Region, U.S. DEP’T OF ST., http://www.state.gov/s/greatlakes_drc/191417.htm (last visited Nov. 18, 2015).
8. Id. § 78m(a).
It acknowledged a global problem and joined an international movement to address it, but rather than take part in an international collaborative initiative, it adopted domestic regulations unilaterally.

This Article argues that the Dodd-Frank’s reach is potentially far greater than the drafters’ purported ambition. To explain Dodd-Frank’s significance, this Article draws from the literature on unilateral regulatory globalization, the phenomenon by which one State entices businesses outside its jurisdiction as well as other States to follow its regulations. This phenomenon occurs when a regulator oversees a large market that companies have a strong desire to enter. One recent example concerns the European Union’s (EU) regulations of household chemicals: since 2007, EU rules known as “REACH” impose on manufacturers who sell to EU consumers strict safety standards “to ensure a high level of protection for human health and the environment.” These standards are higher than those required by the United States, yet American companies have altered their operations and products to align with the higher EU standards when selling both to EU and U.S. markets.

I argue that Dodd-Frank represents a form of unilateral regulatory globalization with the potential to promote, on the issue of conflict minerals, a global convergence of regulations and corporate behavior. Although conflict minerals are present in a range of products, this Article focuses primarily on the electronics market to understand the likely effect of various policy interventions and Dodd-Frank in particular. The Article seeks to show that the size of the consumer electronics market, the global span of its supply chain, and the allure of U.S. capital markets all likely combine to unleash a “California Effect,” which induces companies outside the United States and other countries to follow the Dodd-Frank regulations. Importantly, I take no view on the ultimate effectiveness of Dodd-Frank. Although the regulations are affecting corporate behavior around the world, some have argued that the U.S. regulations were misguided. Rather, I assess whether the regulations, regardless of their merit, are likely to affect State and corporate behavior beyond the law’s stated geographical scope.

This Article is novel in two respects. First, it is the first to assess the U.S. conflict minerals regulations through the lens of unilateral regulatory globalization theory. Other articles have addressed the plight of the Congolese,

---


12. *See, e.g.*, Dan Fahey, “Conflict Minerals” in Ituri, TEX. AFR. (Aug. 4, 2010), http://texasinafrica.blogspot.com/2010/08/conflict-minerals-in-ituri.html (suggesting that the conflict minerals provisions of Dodd-Frank were misguided because, by the time the law was adopted, the war had ended in large parts of the DRC where minerals were produced). Part III.B. *infra* also discusses some of the negative, unintended effects of Dodd-Frank, particularly a de facto embargo on all minerals from the African Great Lakes Region.
the problem of conflict minerals in general, the details of the Dodd-Frank regulations, and their implementation. By contrast, this Article uses an understudied theory to explain why this seemingly modest U.S. regulation is likely to have an impact stretching beyond the United States’ jurisdiction and the DRC. Second, this Article adds to the literature on unilateral regulatory globalization by suggesting ways to complete the existing analytical framework, and by testing its application in a new area. While the existing literature discusses unilateral regulatory globalization in relation to environmental, antitrust, health, privacy, and tax policy, it does not study the theory’s relevance for issues at the intersection of business and human rights.

The Article proceeds as follows. Part I provides an overview of the conflict minerals problem, including the human rights violations Dodd-Frank seeks to remedy and the structure of the supply chain in the electronics industry. Part II contextualizes the U.S. policy response by describing the Dodd-Frank regulations on conflict minerals and their implementation thus far, and by comparing Dodd-Frank to more traditional policy interventions the United States might have adopted instead. Part III describes and builds on the literature on unilateral regulatory globalization before applying available theories to Dodd-Frank. The conclusion explains the significance of this Article’s findings for businesses, governments, and human rights advocates.

I. CONFLICT MINERALS: OVERVIEW OF THE PROBLEM

A. Human Rights Violations

Tin, tantalum, tungsten, and gold—the four “conflict minerals”—are components of many consumer electronics items, including cell phones and laptops. These minerals often are the source of human rights abuses in the communities where they are mined. The DRC has been one such “hot spot,” where natural resource extraction has fueled conflict between rebel groups and the national army in the country’s eastern, mineral-rich areas. Tragically,

13. See infra notes 52–53 and accompanying text.
14. See infra note 94 and accompanying text.
violence and conflict have long been part of the DRC’s history, not only during the colonial era, but also since its independence in 1960. In 1993, a particularly deadly conflict erupted in the Eastern DRC, in part due to the influx of over seven hundred thousand refugees fleeing the Rwandan Genocide. The conflict in the DRC has led to an estimated five million deaths, and both army and rebel representatives have committed a range of human rights violations, including mass murder, mass rape, systematic shelling of refugee camps, and the enlistment of child soldiers.

The Eastern DRC is also where many conflict minerals have been mined, and these minerals have been a “key factor” in the violence in the region. Rebel groups and national army commanders have controlled those mines and sold the minerals for millions of dollars every month to refineries and smelters as a way to finance the conflict. A 2010 report commissioned by the United Nations Security Council observed at the time that, “[i]n the Kivu provinces, it appears, almost every mining deposit is controlled by an armed group.” A 2012 companion United Nations report documented the national legislative efforts in the DRC to improve the industry certification schemes, but insecurity around certain mining sites remained a serious problem, as was the smuggling of minerals into and out of the country. The U.S. regulations examined in this


21. STEARNS, supra note 19, at 4.

22. Arimatsu, supra note 20, at 158–59 (characterizing the violence in eastern DRC as “unprecedented” and recounting the systematic shelling of refugee camps and the displacement of 250,000 to 500,000 people); U.N. Final Report, supra note 177, ¶ 147 (describing sexual violence by armed groups, including “mass rapes” and rapes against minors); id. ¶ 148 (describing “indiscriminate killings of civilians”); id. §§ 153–58 (documenting recruitment of child soldiers).


Article may be partly responsible for the decrease in rebel group mining of tin, tungsten, and tantalum. But such reforms have had more limited impact on gold trade, which presents singular challenges. In addition to the human rights abuses of the conflict itself, rebels and military officers have imposed excruciating labor conditions on miners and employed children in the mines. For example, one report described child labor and work shifts of “two or more days at a time in dark, damp holes pervaded by the smell of human sweat and excrement.” Another report documented forced labor and deaths from harsh working conditions.

B. The Electronics Industry and the Market for Conflict Minerals

Tin, tungsten, tantalum, and gold are present in a range of products, including medical devices, industrial tools, and jewelry. But these minerals’ presence in consumer electronics has received particular scrutiny. In our cell phones, laptops, tablets, and other electronic devices, these components fulfill several purposes, such as coating other metals, storing electricity, and conducting electricity and heat.

The global consumer electronics supply chain involves multiple business entities and countries, impeding the ability to trace conflict minerals. Minerals are extracted in one country; sold to trading houses and other intermediaries in the region; exported to countries with smelters to be melted; exported again in refined form to countries where manufacturing plants use them to assemble finished products; and finally sold to the consumer as part of the finished product, usually in yet another country. Advocacy groups, businesses, and consulting groups have identified smelters as the point in the supply chain where

---

32. U.N. Final Report, supra note 17, at 170.
the number of actors is the smallest, and audits are, therefore, more feasible.\footnote{37} Not only are there relatively few smelters globally, but they are also geographically concentrated. For example, in 2012, China accounted for approximately 46\% of global primary smelter production of tin.\footnote{38} And 45\% of all U.S. imports of tungsten between 2009 and 2012 came from China.\footnote{39} Still, the small amount of minerals contained in each final product makes their traceability particularly difficult: one consulting group explained that “a 2 kilogram (4.5 pound) laptop contains 10 grams of tin, 0.6 grams of tantalum, 0.3 grams of gold, and 0.0009 grams of tungsten.”\footnote{40}

Despite being an important producer of tin, tungsten, tantalum, and gold, the DRC does not dominate global production of the minerals. The DRC accounted for only 1.7\% of global tin mine production in 2013,\footnote{41} 0.004\% of global tungsten production in 2011,\footnote{42} 18.6\% of global tantalum production in 2013\footnote{43} (another 25.4\% was produced in neighboring Rwanda\footnote{44}), and 0.10\% of global gold production in 2011.\footnote{45} Yet this comparatively modest share of global production has not discouraged human rights campaigners from focusing their advocacy efforts on the DRC because of the disproportionate potential of mining in that country to fuel armed conflict.\footnote{46}

The United States is a major consumer of the aforementioned minerals, yet U.S. consumption is concentrated in only a few sectors and companies. For example, in 2013, twenty-five U.S. companies accounted for approximately 90\% of domestic primary tin consumption, 17\% of which was used for electrical purposes.\footnote{47} As another example, a single company, Intel, manufactures 80\% of the world’s semiconductors.\footnote{48} This concentration explains in part the possible significance of ripple effects from U.S. regulations across global supply chains in the electronics sector. The Securities and Exchange Commission (SEC)

\footnote{37} A.T. Kearney, supra note 333, at 4; Global Witness, ‘The Hill Belongs to Them,’ supra note 25, at 19 (“[T]he number of major international smelters of tin and tantalum . . . is fairly small and they represent a key bottleneck in the global supply chain.”); Apple’s Conflict Mineral Policy, Actio (Sept. 16, 2014), http://blog.actio.net/supply-chain-management/apples-conflict-mineral-policy/ (“We believe the only way to impact the human rights abuses on the ground is to have a critical mass of smelters verified as conflict-free, so that demand for the mineral supply from questionable sources is affected.”).
\footnote{38} James F. Carlin, Jr., Tin, in 2012 MINERALS YEARBOOK 77.9 tbl.10 (2014).
\footnote{40} A.T. Kearney, supra note 333, at 4.
\footnote{41} Mineral Commodity Summaries 2014, supra note 39, at 169.
\footnote{43} Mineral Commodity Summaries 2014, supra note 39, at 161.
\footnote{44} Id.
\footnote{46} See, e.g., Baflemba et al., supra note 288, at 4 (estimating that before Dodd-Frank, conflict minerals generated an estimated $185 million per year for armed groups and the army).
\footnote{47} Mineral Commodity Summaries 2014, supra note 39, at 168.
estimates that its regulations on conflict minerals will affect approximately six thousand U.S. and foreign companies.\textsuperscript{49} The EU estimates that “150,000–200,000 EU companies—mostly downstream operators—are involved in the supply chains” of the six thousand affected U.S. companies.\textsuperscript{50} After the first deadline in 2014 for companies to file reports to the SEC documenting their reliance on conflict minerals, 1,313 companies filed such reports, and many of these companies were from the semiconductor, broadcasting, electronic components, and computer communications equipment sectors.\textsuperscript{51}

\section{II. A SINGULAR U.S. POLICY RESPONSE: THE DODD-FRANK CONFLICT MINERALS PROVISIONS}

Part I described the problem of conflict minerals generally, the conflict in the DRC, and the structure of the consumer electronics value chain. Part II focuses on the U.S. policy response to the problem of conflict minerals: the Dodd-Frank Wall Street Reform and Consumer Protection Act. This Part provides an overview of the Act’s conflict minerals provisions, compares this policy response to more traditional approaches for addressing cross-border human rights challenges, and summarizes the implementation of the U.S. law thus far.

\subsection{A. Overview of Dodd-Frank Conflict Minerals Regulations}

Detailed descriptions of the U.S. regulations on conflict minerals have been provided elsewhere.\textsuperscript{52} Others have also analyzed and debated in detail the law’s effectiveness.\textsuperscript{53} In contrast, this Article provides a brief overview of the U.S.

\begin{itemize}
\item \textsuperscript{49} Jim Low, "Dodd-Frank and the Conflict Minerals Rule," KPMG DIRECTORS & BOARDS 44 (4th Quarter 2012).
\item \textsuperscript{50} Joint Communication, supra note 166, at 7.
regulations to assess whether they represent a new kind of policy intervention likely to spur global action.

In 2010, in response to the financial crisis, the United States enacted Dodd-Frank, primarily to "improv[e] accountability and transparency in the financial system." This major reform contained several "Miscellaneous Provisions," including one on conflict minerals. Section 1502 of Dodd-Frank amended the Securities Exchange Act of 1934 to address "the exploitation and trade of conflict minerals originating in the Democratic Republic of the Congo," which was financing "conflict characterized by extreme levels of violence in the eastern Democratic Republic of the Congo, particularly sexual- and gender-based violence, and contributing to an emergency humanitarian situation therein."

The statute and accompanying SEC regulations require all companies with stock traded on the U.S. Stock Exchange to report yearly to the SEC the existence of conflict minerals in their supply chains originating from "the Democratic Republic of the Congo or an adjoining country." Companies subject to the regulations must file a special form (known as Form SD), in which they detail the steps taken to conduct due diligence along their supply chains; indicate which products, if any, contain conflict minerals; and identify the suppliers and mines from which the minerals originated. With limited exceptions, companies’ due diligence must “conform to a nationally or internationally recognized due diligence framework,” and companies’ Conflict Minerals Reports must be subject to a “private sector audit” to the extent that companies wish to declare their products as “DRC conflict-free.”

Importantly, Dodd-Frank does not ban companies from using conflict minerals. Rather, the U.S. policy assumes that the presence of conflict minerals in supply chains is a material risk to companies’ bottom lines that merits shareholder scrutiny. The U.S. Congress thus embraced a transparency-based regulation theory, according to which exposing a problem to the public can

foster public action against it. The result may be the same: several technology companies subject to the rule, including Apple, HP, Intel, and SanDisk, have already committed to removing all conflict minerals from their supply chains.

## B. Range of Possible U.S. Policy Responses

In order to assess the significance of Dodd-Frank, this section compares it to other possible policies the United States could have adopted to address the problem of conflict minerals. Dodd-Frank departs from the traditional policy approaches deployed by the United States and other countries to tackle international human rights challenges involving the private sector. One approach could have been to strengthen domestic regulatory frameworks and enforcement mechanisms in the countries where the violations occur. The multipronged international response to the 2013 collapse of the Rana Plaza building in Bangladesh, which resulted in the deaths of more than one thousand garment workers, exemplifies such an approach. The world’s major retailers and brands, governments, and international organizations, including the EU and the International Labor Organization, established a partnership to strengthen the Bangladeshi labor laws, implement oversight mechanisms, and compensate victims. In addition, the U.S. government suspended some of Bangladesh’s trade benefits until the Bangladeshi government could demonstrate an improvement in workers’ rights and conditions. The U.S. Senate considered various policy responses to address the Rana Plaza disaster, but none resembled the corporate disclosure required by Dodd-Frank.

---


64. MAJORITY STAFF OF S. COMM. ON FOREIGN RELATIONS, 113TH CONG., WORKER SAFETY AND LABOR RIGHTS IN BANGLADESH’S GARMENT SECTOR 1–2 (Comm. Print 2013) (recommending as policy interventions temporary suspension of trade benefits against Bangladesh; increased funding for technical assistance to Bangladesh; education of corporate suppliers; stronger sanctions by the government of Bangladesh of companies that violate local law; and improvement of Bangladeshi labor laws).
By contrast, the U.S. government response to the problem of conflict minerals in the DRC has not involved attempts to amend Congolese laws. The United States’ reluctance to pursue such legal reforms may have been due to the DRC’s lack of technical capacity; although the DRC had enacted legislation in 2012 aimed at reducing the mining profits of armed groups, corruption and a scarcity of public resources have prevented effective enforcement of the new legislation.66

Dodd-Frank also departs from the U.S. policy response to conflict diamonds, in which the Kimberley Process established—with limited success—an international State-led certification scheme.67 Nor does Dodd-Frank model the EITI, a multi-stakeholder initiative aimed at combating corruption in the exploration of oil, gas, and minerals.68 Furthermore, rather than relying on a new treaty, Dodd-Frank relies on domestic legislation to spark global action regarding conflict minerals.

It is true that Dodd-Frank relies on familiar international strategies in some respects. For instance, the U.S. law relies in part on industry certification schemes. The London Bullion Market Association’s Responsible Gold Guidance69 and the Electronic Industry Citizenship Coalition’s and Global e-Sustainability Initiative’s Conflict-Free Smelter Program70 are both aimed at identifying conflict minerals along companies’ supply chains. Both certification schemes could support companies seeking to comply with the SEC’s disclosure requirements, which call on independent auditors to verify the presence of conflict minerals in companies’ supply chains in certain circumstances.71 Moreover, in requesting due diligence in corporate supply chains, the United States embraced the approach advocated by the Organisation for Economic Co-

65. Arrêté Ministériel N.0057.CAB.MIN/MINES/01/2012 du 29 Février 2012 Portant Mise en œuvre du Mécanisme Régional de Certification de la Conférence Internationale sur la Région des Grands-Lacs “CIRGL” en République Démocratique du Congo, art. 8 (requiring that all companies operating in the DRC conduct due diligence assessments in line with OECD standards).


67. See Holly Cullen, Is There a Future for the Kimberley Process Certification Scheme for Conflict Diamonds?, 12 MACQUARIE L.J. 61 (2013) (pointing out recent failures of the Kimberley Process and asking whether it still has a role to play in addressing the problem of conflict diamonds); KIMBERLEY PROCESS, supra note 4 (“The Kimberley Process Certification Scheme . . . imposes extensive requirements . . . on its members to enable them to certify shipments of rough diamonds as ‘conflict-free’ and prevent conflict diamonds from entering the legitimate trade.”);

68. EXTRACTIVE INDUSTRIES TRANSPARENCY INITIATIVE, supra note 3.


operation and Development (OECD), which developed a framework for companies to “respect human rights and avoid contributing to conflict through their mineral purchasing decisions and practices.”72 Taken as a whole, however, Dodd-Frank is a U.S.-centric response that relies on international certification methods to protect and inform U.S. consumers and investors, not to join an international cooperation effort.

C. Implementation of Dodd-Frank

Since Dodd-Frank’s enactment in 2010, the SEC has issued implementing regulations and companies have begun to comply with the new law. Legislative repeal of the provisions on conflict minerals appears unlikely at this stage, though congressional priorities are hard to predict. For example, in December 2014, large banks included an amendment to Dodd-Frank in an unrelated budget bill, repealing restrictions on risky derivatives trading.73 Opponents of the conflict minerals regulations could proceed similarly to dismantle those provisions.

In addition, some industry groups have already convinced a federal court to strike down a portion of the conflict-minerals rule, and additional legal challenges are possible. A federal appeals court in 2015 concluded that part of the rule violated the U.S. Constitution.74 The court held that the SEC may not require companies to state in their reports to the SEC or on their websites that their supply chains were “not found to be ‘DRC conflict-free.’”75 Requiring companies to do so would amount to compelled speech and thus infringe on their First Amendment rights, the court said.76

The other provisions of the rule, however, survived this legal challenge, and companies continue to submit conflict minerals report to the SEC.77 This decision reaffirmed a ruling by the same court a year earlier.78 After the first

---

75. Nat’l Ass’n of Mfrs., 800 F.3d at 530, 553 n.8.
76. Id. at 524.
77. Id. at 553 n.8; Dynda A. Thomas, SEC Conflict Minerals Rule Legal Challenge is Over – But Not For Good, CONFLICT MINERALS LAW (Apr. 12, 2016), http://www.conflictmineralslaw.com/2016/04/12/sec-conflict-minerals-rule-legal-challenge-is-over-but-not-for-good/.
78. The D.C. Circuit decided to rehear this case after the same court, en banc in another case, clarified the scope of the doctrine of protected commercial speech under the First Amendment. Am. Meat Inst. v. U.S. Dep’t of Agric., 760 F.3d 18 (D.C. Cir. 2014) (en banc). During the rehearing, the
ruling, the SEC issued guidance and a temporary stay of the conflict-minerals rule, both of which clarify the 2015 judicial decision’s effect on the regulation and companies’ obligations. The SEC has explained that companies with conflict minerals in their supply chains continue to have an obligation to disclose “the facilities used to produce the conflict minerals, the country of origin of the minerals and the efforts to determine the mine or location of origin.” Moreover, while no company is obligated to declare its products as “DRC conflict free,” “not been found to be ‘DRC conflict free,’” or “DRC conflict undeterminable,” companies may continue to apply the “conflict free” label voluntarily, so long as they conduct independent private sector audits to support that assertion.

The first year companies were required to report their use of conflict minerals to the SEC was 2014. One month after the deadline, a consulting group counted 1,313 reporting companies. Approximately 20% of these companies listed their supply chains as “conflict free,” while the other 80% said they were unable to make a final determination. In 2015, 1,272 companies filed forms SD to the SEC, with a similar proportion of filers listing their supply chains as “conflict free.”

The new regulations have prompted companies to map their supply chains to an unprecedented extent. The SEC estimates that the regulations will cost companies three to four billion dollars the first year and two hundred million

---

80. Higgins, supra note 58.
81. Id.
82. Usvyatksy, supra note 51.
83. Id. To ease implementation, the regulation had allowed companies in the first two or four years (depending on the size of the company) to declare as “DRC undeterminable” the status of products whose provenance could not be reliably ascertained after appropriate due diligence. FORM SD: SPECIALIZED DISCLOSURE REPORT, supra note 57, at 3. In 2015, the share of companies to list their status as “undeterminable” was approximately 80% as well. Coleman, supra note 51. As noted above, however, the SEC clarified in the aftermath of the D.C. Circuit’s decision in National Association of Manufacturers that companies were no longer required to use this label. See supra note 81 and accompanying text.
84. Coleman, supra note 51.
85. Bob Trebilcock, Source Intelligence: Mapping the Supply Chain and Monitoring for Risk, SUPPLY CHAIN MGMT. REV. (Nov. 18, 2014) (interview with Jess F. Kraus), http://www.scmr.com/article/source_intelligence_mapping_the_supply_chain_and_monitoring_for_risk (“There is no other requirement in the US that requires you to map your supply chain, determine where your materials are coming from, and have that audited by a third party. In fact, there’s never been anything like it.”).
dollars every year afterwards. In a review of Dodd-Frank, one group concluded that:

because the New York Stock Exchange and other U.S. capital markets are still an important destination for corporations worldwide, particularly for large, multinational companies that produce the final products that use minerals, the legislation has had a significant impact on the global supply chains of three of the four conflict minerals. The regulations are also expected to improve conditions in the DRC. After conducting field research in the Great Lakes Region following Dodd-Frank’s enactment, one NGO discovered that armed groups now experience more difficulty trading in tin, tungsten, and tantalum, but that the illegal trade in gold has been harder to curtail. These findings are supported by another group, although a clear cause-and-effect relationship is difficult to establish because of other simultaneous legislative enactments, such as a ban on artisanal mining in the Eastern DRC. Additionally, a 2012 United Nations report highlights the risk of smuggling: while exports of conflict minerals from the DRC decreased, smugglers located in nearby countries exported the minerals from Burundi, Rwanda, and Uganda. In addition, strict regulations and due diligence requirements can have unintended side effects, including a possible “de facto embargo” of the Great Lakes Region, as sourcing verified, conflict-free minerals becomes prohibitively expensive for companies. Instead, companies might choose to source the desired minerals from other regions.

In sum, the regulations have induced companies to take a closer look at their supply chains, develop systems to track their reliance on conflict minerals, embrace the OECD Due Diligence Guidelines, and establish partnerships with third-party auditors to certify some of their SEC reports and vet suppliers. Additionally, the regulations have impacted companies not listed on the U.S. Stock Exchange: since 2010, companies filing with the SEC have sought assurances from suppliers around the world that no conflict minerals were used in manufacturing their products.

86. Low, supra note 49, at 44.
87. BAFILEMBA ET AL., supra note 288, at 6.
88. Id. at 1, 8.
89. MANHART & SCHLEICHER, supra note 533, at 30.
91. MANHART & SCHLEICHER, supra note 53, at 33; accord Dan Fahey, “Congo Gold”: Three Problems with the 60 Minutes Story, AFR. ARGUMENTS (Dec. 11, 2009), http://africanarguments.org/2009/12/11/three-problems-with-60-minutes/ (“Cutting off Congo’s gold would be a social and economic disaster for areas like Ituri that are struggling to emerge from war.”).
92. For a critical view of the Dodd-Frank regulations emphasizing this point in particular, see Fahey, “Conflict Minerals” in Ituri, supra note 12 (“[It is easy for the producers of electronics destined for the USA to obtain their ‘conflict minerals’ from other sources.”).
93. MANHART & SCHLEICHER, supra note 53.
94. Id. at 26.
III.
THE GLOBAL PULL OF AN ENTICING MARKET AND POWERFUL REGULATOR

Part II.B. explained how the Dodd-Frank regulations on conflict minerals differ from conventional policy responses to international challenges; rather than investing in multilateral institutions, the United States unilaterally adopted domestic legislation targeting one region where the problem of conflict minerals was acute. In this Part, this Article, for the first time, grounds this U.S. policy response in the literature on unilateral regulatory globalization. That literature does not discuss the Dodd-Frank regulations or other efforts to tackle the problem of conflict minerals. In fact, problems at the intersection of business and human rights more generally have not been analyzed through the prism of unilateral regulatory globalization theory. Part III begins with an overview of the literature on unilateral regulatory globalization and continues with an application of the theory to Dodd-Frank’s rules on conflict minerals.

A. The Theory of Unilateral Regulatory Globalization

The literature on unilateral regulatory globalization studies the power of one State, or one group of States, to impose its regulatory policies on other States and on companies in a way that leads to a global harmonization of standards and practices.95 David Vogel, one of the first proponents of this theory, described California’s propensity to set environmental standards that would be subsequently followed by other states and the U.S. federal government.96 Vogel observed that from the 1970s to 1990s, California regularly set the country’s most stringent automobile emission standards, after which other states and the federal government would raise their own standards to California’s level.97 He termed this “ratcheting upward of regulatory standards” across political jurisdictions the “California Effect.”98

Building on Vogel’s theory, scholars subsequently began arguing that a similar phenomenon could take place across borders: one country’s regulations could influence another and lead to global harmonization. Scholars tested Vogel’s theory about automobile emission standards across borders and concluded that countries that exported cars to jurisdictions with more stringent automobile emission standards tended to adopt more stringent emission standards themselves.99 Furthermore, other scholars studied the conditions under which the EU was more likely to impose some of its stricter environmental,

95. For example, Anu Bradford explains that “[u]nilateral regulatory globalization occurs when a single state is able to externalize its laws and regulations outside its borders through market mechanisms, resulting in the globalization of standards.” Bradford, supra note 10, at 3.
97. Id.
98. Id.
Finally, others focused on the global reach of national regulations involving corporate taxation, international investment and banking, antitrust, health and safety, privacy, and the environment.

To illustrate the phenomenon, one author argued provocatively that “EU regulations dictate what kind of air conditioners Americans use to cool their homes and why their children no longer find soft plastic toys in their McDonald’s Happy Meals.” Another example may be familiar to American consumers: Canada’s laws likely explain why some products on the shelves of U.S. supermarkets are labeled both in English and French. Canadian law requires many consumer products sold in Canada to be labeled in both official languages. While Canadian regulations do not require products in U.S. stores to satisfy the same requirement (and the U.S. government does not mandate labeling in French), some American companies nevertheless label their products in both languages. They do so because exporting to Canada requires bilingual labeling, and it is more economical for a product to come off the assembly line with packaging ready for either of North America’s two largest markets. In a sense, some companies are voluntarily complying with Canadian law outside Canadian borders.

In her article The Brussels Effect, Anu Bradford focuses on the EU’s ability to “export” its regulations to other countries—a variation of the California
Effect. While her contribution echoes prior scholars’ analyses in certain sections, it provides a useful framework to understand the facets and conditions of the California Effect. I will therefore use it as a starting point to describe the phenomenon.

Bradford distinguishes de jure and de facto harmonization, explaining that the California Effect can result in either or both.\footnote{Bradford, supra note 10. The author refers to the process of global harmonization led by the EU as “The Brussels Effect.” For simplicity, and because this Article analyzes the potential of an American rule to unleash a similar effect, I will refer to the phenomenon as the “California Effect” throughout this Article.} De jure harmonization refers to the adoption by other States of the strict rules of the dominant regulator.\footnote{Id. at 8.} De facto harmonization, on the other hand, occurs when companies choose to follow the dominant regulator’s rules in their operations around the world even though other States have not adopted the dominant regulator’s stricter rules.\footnote{Id.} This de facto harmonization—at play in the Canadian labeling example above—takes place because businesses find it economically advantageous to standardize their practices globally to follow a single rule.\footnote{Id. at 6.} De facto and de jure harmonization often go hand in hand: once large companies have standardized their practices, they have an incentive to lobby their home governments to level the playing field with their domestic competitors who are not export-oriented and so do not need to comply with the foreign regulator’s stricter standards.\footnote{See id. at 6.}

Of course, global harmonization of standards can result from international cooperation as well. Countries could adopt a new treaty banning trade in conflict minerals and requiring each signatory State to enact legislation to that effect. The ban on trade in endangered species exemplifies this cooperative approach: 182 countries\footnote{Member Countries, CONVENTION ON INT’L TRADE IN ENDANGERED SPECIES OF WILD FAUNA AND FLORA, http://www.cites.org/eng/disc/parties/index.php (last visited Apr. 24, 2016).} are now party to the Convention on International Trade in Endangered Species of Wild Fauna.\footnote{Convention on International Trade in Endangered Species of Wild Fauna and Flora, Mar. 3, 1973, 27 U.S.T. 1087, 993 U.N.T.S. 243.} Harmonization through unilateral regulatory power, however, has distinct advantages. In particular, it is easier to adopt and enforce since the dominating country need not secure the consent or compliance of other States.\footnote{Bradford, supra note 10, at 44.}

Bradford lays out five conditions that give rise to the California Effect: (1) market power, (2) regulatory capacity, (3) preference for strict rules, (4) the regulation of inelastic targets, and (5) nondivisibility of standards.\footnote{Id. at 10–11.}
First, market power refers to a State’s ability to offer foreign companies access to a lucrative domestic market, preferably one with wealthy consumers, in exchange for compliance with the State’s regulations. Bradford argues that “the larger the market of the (strict) importing country relative to the (lenient) market of the exporter country, the more likely the Brussels Effect will occur.” According to the author, the EU, the United States, China, and Japan “possess domestic markets large enough to use access to their markets as leverage.”

Second, regulatory capacity is necessary for a State “to translate its market power into tangible regulatory influence.” Regulatory capacity requires regulatory expertise and the authority to impose harsh sanctions for noncompliance. According to Bradford, “[t]he U.S. administrative agencies’ capacity to promulgate and enforce rules in the United States is well understood,” and the EU is rapidly developing an equivalent regulatory order. But outside these two blocs, the author argues that regulatory capacity escapes other large economies, including China.

Third, preference for strict rules refers to the willingness of States with market power and regulatory capacity to deploy these attributes towards the adoption and enforcement of strict regulatory standards. Bradford argues that wealthier countries are more likely to adopt strict rules, as are countries that are more risk averse and more committed to “a social market economy.” According to Bradford, the EU’s adoption of the precautionary principle illustrates a general preference for strict rules, whereas U.S. agencies’ insistence on cost-benefit analysis to justify intervention reflects a relative aversion to strict rules.

Fourth, the inelastic-targets factor refers to the propensity of the regulation’s target to relocate to circumvent the strict regulations. Bradford explains that the EU regulations are likely to unleash the California Effect when they focus on consumer markets, such as product or food safety, because it is a sale to EU consumers that triggers companies’ obligation to comply with EU regulations, and consumers are unlikely to “relocate” outside EU borders to

116. *Id.* at 11–12.
117. *Id.* at 11.
118. *Id.*
119. *Id.* at 12.
120. *Id.* at 12–13.
121. *Id.* at 13.
122. *Id.* at 13 n.48.
123. *Id.* at 14.
124. *Id.* at 14–15.
125. *Id.* at 15–16. Vogel also analyzed the factors that drove the United States and the EU towards stricter or laxer rules over time. See David Vogel, *The Politics of Precaution: Regulating Health, Safety, and Environmental Risks in Europe and the United States* (2012) [hereinafter *Vogel, The Politics of Precaution*].
avoid strict regulations. Thus, if a company wishes to reach the lucrative EU consumer market, it has no choice but to comply with the EU regulations or risk sanctions. In contrast, corporations’ places of incorporation are more elastic, since a company wishing to avoid a high tax rate, for example, will typically be able to relocate to another jurisdiction without significant damage to its operations or profits. Regulations of capital are generally less likely to lead to global harmonization for a similar reason: companies can relocate their financial assets relatively easily without sacrificing market share or access to financial services. As a result, Bradford predicts that the United States’ recent regulatory pursuits in the financial sector are “less likely” to be “converted to global standards because of the relative elasticity of capital.”

Finally, nondivisibility of standards refers to companies’ incentives to standardize their products and operations across world markets. Bradford explains, “the exporter has an incentive to adopt a global standard whenever its production or conduct is nondivisible across different markets or when the benefits of a uniform standard due to scale economies exceed the costs of forgoing lower production costs in less regulated markets.” For example, EU privacy regulations concern only Google’s service offerings within the EU, but technical limitations sometimes force Google to amend its operations worldwide because it is unable, or finds it prohibitively expensive, to devise a version of its services or data collection systems just for the EU.

B. Can Dodd-Frank’s Conflict Minerals Provisions Unleash a “California Effect”?

Can the Dodd-Frank regulations on conflict minerals unleash a “California Effect” that would lead to de facto or de jure global regulatory convergence? On a theoretical level, all five factors discussed in Part III.A. arguably weigh in favor of such an effect, but there are important differences between Dodd-Frank and the environmental, privacy, and health measures discussed above to illustrate the phenomenon. In addition, empirical evidence—however limited since the recent enactment of Dodd-Frank—can also shed light on the extent to

127. Id. at 17.
129. Bradford, supra note 10, at 60.
130. Id. at 17.
131. Id. at 18. A counterexample would be the European Court of Justice decision on the “right to be forgotten,” according to which Google must comply with qualifying requests from EU citizens to remove content from Google’s search engine. Because Google can display different search results on different country pages (such as Google.de and Google.com), the technology giant can choose not to implement the EU’s “right to be forgotten” across all of its platforms worldwide. In this instance, the product regulated by the EU is divisible. See Mark Scott, ‘Right to Be Forgotten’ Should Apply Worldwide, E.U. Panel Says, N.Y. TIMES, Nov. 26, 2014. See also Alex Hern, Google Says Non to French Demand to Expand Right to Be Forgotten Worldwide, GUARDIAN (July 30, 2015), http://www.theguardian.com/technology/2015/jul/30/google-rejects-france-expand-right-to-be-forgotten-worldwide.
which companies not directly subject to Dodd-Frank and other regulators are embracing the U.S. standard.

Before turning to the five factors, we must determine, as an initial matter, what constitutes evidence of gradual regulatory convergence. While no company outside the SEC’s jurisdiction will voluntarily file a Form SD with the Commission, convergence could manifest itself in other ways, such as increased company due diligence, increased reliance on third-party certifications, a reduction in sourcing of minerals from the Great Lakes Region, increased disclosures on company websites of conflict minerals policies, or the adoption by other regulators of disclosure requirements similar to the United States’.

1. Market Power

The United States likely satisfies the first factor: market power. In the cases typically examined in the literature on unilateral regulatory globalization—such as health or environmental policies—market power is defined by the consumer base a company can reach if it complies with the market’s regulator. As an example, Johnson & Johnson will decide to comply with the EU’s REACH regulations on the safety of household chemicals because such compliance is a condition to reaching the lucrative EU consumer base.132

In the context of conflict minerals, however, two initial differences emerge that justify thinking about market power more holistically. First, Dodd-Frank imposes no content requirements on products entering the U.S. market. Unlike EU substantive regulations on imports of beef raised on growth hormones,133 for example, the U.S. rules on conflict minerals are procedural in nature: Dodd-Frank compels companies to disclose the presence of conflict minerals in their products.134 The United States thus leaves the choice to the consumer to purchase or shun the products. Note, however, that even if the regulator does not ban a product or component but merely compels its disclosure, the regulations may still lead to a California Effect if consumers in the target market are likely to shop based on those disclosures. Second, Dodd-Frank offers more than access to U.S. consumers in exchange for compliance with its rules: it offers access to U.S. investors. Companies whose stock is traded in the U.S. Stock Exchange must disclose their reliance on conflict minerals to the SEC,135 so access to U.S. capital markets becomes a major market incentive to comply with the conflict

---

132. See supra notes 9–10 and accompanying text for a discussion of the EU’s REACH regulations.


minerals rule. A Japanese, Korean, or European company can sell products containing conflict minerals to U.S. consumers without reporting it to the SEC as long as the company is publicly traded only outside the United States.\footnote{136} Similarly, non–publicly traded companies do not need to report to the SEC.

A comprehensive analysis of the share of the global electronics market served by companies listed on a U.S. Stock Exchange is outside the scope of this Article. But several of the world’s largest consumer electronics manufacturers are listed on the U.S. Stock Exchange, which suggests Dodd-Frank’s broad potential geographical reach. These manufacturers include Apple, Canon, HP, IBM, Intel, Microsoft, Philips, and Sony.\footnote{137} Notable absences include Dell, HTC, Hitachi, Lenovo, LG, Nikon, Nintendo, Panasonic, Samsung, and Toshiba.\footnote{138}

Still, even some of the companies that are not required to file with SEC have adopted and publicized conflict mineral policies—some of which are more ambitious than others—including Dell,\footnote{139} LG,\footnote{140} and Lenovo.\footnote{141}

---


\footnote{137} See id. at 56,287 (clarifying that, despite arguments to the contrary by two of the bill’s cosponsors, only issuers that file reports with the SEC under Section 13(a) or Section 15(d) of the Securities Exchange Act are required to file a Form SD).


\footnote{149} Addressing Conflict Minerals, DELL, http://www.dell.com/learn/us/en/uscorp1/conflict-minerals (last visited Nov. 21, 2015) (“Dell has been involved in many other efforts to bring us closer to a conflict-free supply chain.”).
relationship between cause and effect is difficult to establish, but it is possible that as some companies disclose more about their supply chains to comply with SEC regulations, other companies not subject to the SEC regulations will feel pressure to match this due diligence to alleviate suspicions from consumers or regulators that conflict minerals lie in their supply chains.152

2. Regulatory Capacity

The U.S. government has the regulatory capacity to trigger the California Effect, particularly in the financial sector where the SEC has vast powers to enforce federal securities laws.153 It is true that some have criticized the SEC for insufficiently enforcing securities laws after the 2009 financial crisis.154 Doubts about the SEC’s willingness to enforce Dodd-Frank’s conflict minerals provisions could lead some companies to file no report or incomplete reports. A more thorough analysis of companies’ SD-Form filings over time could test that hypothesis. But the nine-billion-dollar fine against France’s BNP Paribas in 2014 shows that, under some circumstances, the U.S. government is willing to flex its political and legal muscle, even against powerful foreign banks backed by their home governments.155

3. Preference for Strict Rules

A preference for strict rules, the third condition, is clear in this case. While the United States has advocated weaker rules than the EU in some areas over the
past two decades, including on consumer and environmental protection, the United States went further on conflict minerals with Dodd-Frank than any other country. One industry consultant argued, “[t]here is no other requirement in the U.S. that requires you to map your supply chain, determine where your materials are coming from, and have that audited by a third party. In fact, there’s never been anything like it.” Admittedly, the regulations are not as “strict” as they could be. For example, Dodd-Frank could require disclosure of conflict minerals coming from all “hot spots” rather than just the DRC. But the California Effect does not depend on one regulator adopting the strictest possible rule. Rather, as long as one powerful regulator’s rules are stricter than its foreign counterparts’, companies wishing to enter the powerful regulator’s market will consider aligning all of their operations with those stricter rules.

4. Target Elasticity

The targets the United States is regulating are probably inelastic, though the outcome is less clear on this factor. Assuming the regulation’s target is the economic actor whose behavior the regulation seeks to shape, then the SEC regulations’ targets are the estimated six thousand companies required to file a Form SD with the agency. Assessing the elasticity of these targets’ behavior means asking how likely these companies are to shift their activities to avoid being subject to the regulation. Since a duty to report to the SEC stems from a company’s registration on the U.S. Stock Exchange, elasticity exists if companies are likely to pull out of the U.S. Stock Market, or refuse to enter it, to avoid the conflict minerals rule.

It is too soon to determine empirically the elasticity of the companies’ stock exchange listing decisions to the Dodd-Frank rules. Companies commonly relocate to take advantage of lower tax rates, but the academic literature is less decisive on the impact of government regulations on companies’ decisions to list in a given country’s securities market. Capital is generally more elastic than individual consumers, and in this sense, the targets of the conflict minerals rule are elastic; some companies may find that being listed on another major economy’s stock exchange offers benefits similar to participation in the U.S.

156. Vogel, The Politics of Precaution, supra note 4 (explaining that while the United States used to impose on companies more stringent environmental and food-safety standards than did the EU, the reverse has been true since approximately 1990); Bradford, supra note 10, at 15 (“Since [the 1980s] . . . the EU has increasingly adopted tighter standards of consumer and environmental protection while the United States has failed to follow the EU’s lead.”).

157. Trebilcock, supra note 77.

158. See Global Witness, Tackling Conflict Minerals, supra note 16, at 10 (providing a map of “hotspots” where natural-resource extraction is fueling conflicts).

159. See Low, supra note 49, at 44 (estimating at six thousand the number of companies in the United States and abroad affected by the SEC regulation).

160. Chorvat, supra note 128.

Stock Exchange but without the regulatory costs. But companies likely will weigh the costs of compliance against the costs of exiting U.S. capital markets. After an initial investment of three to four billion dollars in the first year—which most companies made when filing their first SD Forms in 2014—the SEC estimates annual company compliance costs at two hundred million dollars per year. That cost is not trivial, but when weighed against the ability to raise financing on U.S. capital markets, most large companies are likely to absorb the expense. The SEC seems to have come to the same conclusion: a commentator on the proposed conflict minerals rule advised the SEC that “if the final rule would cause ‘more than an insignificant number of foreign private issuers to leave the U.S. markets or not to enter the U.S. markets,’ [the SEC] should consider exempting all or some foreign private issuers from the final rule.”

Despite this suggestion, the SEC chose to keep all foreign private issuers subject to the final rule.

5. Nondivisibility of Standards

Finally, the standards in this case are most likely nondivable. The question this condition poses is whether companies subject to Dodd-Frank must change their practices worldwide to comply with the U.S. regulations, or whether, in some geographical areas, those companies can decide not to track the presence of conflict minerals in their supply chains. Nondivisibility can be legal, economic or technical: a company may align its global practices with Dodd-Frank’s standards because it is legally required to do so, because it is economically rational to do so, or because not doing so is technically difficult. For example, will Philips—a European company publicly traded in the United States and hence subject to Dodd-Frank—limit its supply chain due diligence to minerals that end up in final products sold in the United States? Likely no.

The primary reason for this answer is legal indivisibility. Dodd-Frank does not limit the scope of a company’s due diligence requirements to minerals and products that end up in the United States: companies traded on the U.S. Stock Exchange must report on the existence of DRC conflict minerals across their entire supply chain. The rules’ reach is thus very broad. As soon as a company decides to publicly issue stock in the United States, it incurs an obligation to report to the SEC its possible reliance on conflict minerals worldwide.

162. See id. at 17 (“While not perfectly elastic, capital is significantly more mobile than consumer markets.”).
163. Low, supra note 49, at 44.
164. Id.
166. Id. at 56,288 (“[W]e are not exempting foreign private issuers . . . .”).
167. See Bradford, supra note 10, at 18 (distinguishing legal, technical, and economic nondivisibility).
Moreover, even if the regulation required reporting only on the components of electronic products sold in the United States, business practices would likely remain both economically and technically nondivisible. Economically, a company that makes substantial investments to improve its supply chain monitoring likely will draw on economies of scale to track the presence of conflict minerals across its products. Once a company makes the initial investment to develop processes to monitor a portion of its supply chain (for instance, in a given region), those same processes likely can be deployed at a comparatively low marginal cost across the rest of the company. In addition, there may be incentives for companies to conduct comprehensive assessments of their supply chains, since doing so allows them to market their products worldwide as “conflict-free,” a label consumers may come to value.

Technically, accurate monitoring of conflict minerals in a company’s supply chain in one region may actually require tracking across all regions and suppliers. Conflict minerals fulfill multiple technical functions in consumer electronics, and a myriad of suppliers use them as they manufacture and assemble component parts. So without a comprehensive audit, it is possible that a multinational company will miss a point at which conflict minerals enter its supply chain. In addition, as described in Part I.B., smelters are the most practical point of intervention in companies’ supply chains in the consumer electronics sector. As a result, companies increasingly seek certified, “conflict-free” smelters to avoid having to disclose to the SEC the presence of conflict minerals in their supply chains. In doing so, these companies are cleansing most if not all of their supply chain, since only conflict-free smelters channel materials to suppliers.

6. Remaining Uncertainties

In sum, Dodd-Frank’s geographical reach seems quite broad. Companies subject to the U.S. regulations must conduct due diligence across their supply chains worldwide to detect conflict minerals. Despite this onerous requirement, many companies are likely to consider this cost worthwhile in exchange for access to the U.S. capital markets.

In addition, we would expect the California Effect to lead other jurisdictions to adopt regulations similar to Dodd-Frank. The theory posits that multinational companies incorporated outside the United States but publicly traded on the U.S. Stock Exchange will lobby foreign governments to enact comparable due diligence requirements to level the playing field with competitors not subject to Dodd-Frank. While it is still difficult to state definitively whether this is happening, there are indications that this phenomenon is underway. One commentator to the proposed U.S. conflict minerals rule observed that requiring even foreign private issuers to report to the

168. Fitzpatrick et al., supra note 15, at 975–76.
169. See supra note 104 and accompanying text.
SEC on their reliance on conflict minerals “could actually motivate foreign companies to advocate for similar conflict minerals regulations in their home jurisdictions to reduce any competitive disadvantages they may have with companies from their jurisdictions that do not register with [the SEC].”\textsuperscript{170} The EU has been considering rules on conflict minerals since 2014\textsuperscript{171}—which could affect eight-hundred-thousand European companies\textsuperscript{172}—and some European companies subject to Dodd-Frank, such as Philips, have been engaged in the development of these counterpart EU regulations.\textsuperscript{173}

Apart from the five conditions discussed above, several other considerations merit discussion because they can influence the extent to which Dodd-Frank fosters global regulatory convergence on conflict minerals. First, market power can erode over time.\textsuperscript{174} The appeal of the U.S. market for electronics may decrease as developing countries consume a larger and larger share of the yearly consumer electronics output. Similarly, registration on the U.S. Stock Exchange is attractive today but this too could change. Lastly, another country could soon adopt stricter rules than the U.S.’s rules on conflict minerals, which would lead to global policy convergence towards those new, stricter regulations. In particular, the EU’s rule on conflict minerals may apply to minerals sourced from all conflict-prone areas around the world, not just the DRC.\textsuperscript{175} Global convergence towards the EU’s regulations rather than those of the United States would weaken the persuasive power of U.S. authorities. But from the point of view of advocates seeking to eradicate conflict minerals from the global trade in consumer electronics, this shift in power would not be a concern. On the contrary, human rights advocates would prefer convergence towards the more ambitious policies.

\textsuperscript{174.} Bradford, supra note 10, at 49.
Political opposition could also reduce Dodd-Frank’s global influence. Political opposition from the United States, China, and India to the EU’s regulations on greenhouse gas emissions from international flights landing or departing from the EU forced the EU to repeatedly delay its plans. 176 Similar complaints could weaken Dodd-Frank’s reach and potentially could force Congress to amend the law.

Finally, one should consider the possible unintended consequences of unilateral regulatory action. Several commentators have observed that by focusing exclusively on the Great Lakes Region, Dodd-Frank is ridding global supply chains of conflict minerals at the expense of economic development and stability in the DRC. 177 Instead of working as an incentive to normalize that country’s trade in minerals, the U.S. regulations may be imposing a de facto embargo on the DRC, as it is easier for companies to steer clear of the region altogether than to try to clarify chains of custody and establish relationships with trusted counterparts. The European Commission noted, for example:

There are indications that [Dodd-Frank] has worked as a deterrent to source minerals from the [Great Lake Region], regardless of whether the minerals are legitimately extracted or not. Some affected companies are pursuing a no-risk strategy and source from mines outside the region or even outside Africa. The remaining “conflict-free” minerals struggle to reach US or EU markets and are frequently traded at below market prices. Loss of trade means loss of local livelihoods in a setting where alternative employment opportunities are scarce, in particular in the case of artisanal and small-scale mining. 178

**CONCLUSIONS**

In adopting the Dodd-Frank regulations on conflict minerals, the United States opted to tackle, through unilateral regulations, a global problem that might have called for international cooperation. The United States chose to “cooperate alone.” While this approach lacks many analogs in the business and human rights field, where policy interventions have traditionally been more international and cooperative, a useful analytical framework exists elsewhere. The literature on unilateral regulatory globalization explains how, under the right circumstances, a powerful regulator can entice other States and foreign companies to follow the same procedures the regulator applies to domestic actors.

Dodd-Frank shares many of the attributes of unilateral regulatory globalization. This U.S. law is the product of a major market with the capacity to adopt and enforce strict rules. The regulation focuses on relatively inelastic targets—multinational corporations listed on the U.S. Stock Exchange—and the

177. See supra notes 84–85 and accompanying text.
standards and practices Dodd-Frank requires companies to adopt are nondivisible. The result is striking: one short “miscellaneous” provision in a statute in one country has the potential to change the behavior of businesses and their suppliers in an industry along the supply chain worldwide.

Preliminary evidence from Dodd-Frank implementation suggests that the regulations have caused a decrease in smuggling of conflict minerals in the Great Lakes Region for three of the four target minerals: tin, tungsten, and tantalum. Also to Dodd-Frank’s credit is the increase in interest from the EU in adopting a comparable regulation, as well as the rapid development of multi-stakeholder initiatives to certify smelters and allow companies to exchange best practices in the management of their supply chains.

However, Dodd-Frank also appears to have triggered some unintended side effects. In particular, by targeting one region—the DRC and its neighbors—the U.S. regulation is likely steering away from the Great Lakes economic activity that is badly needed to support local communities and lift the affected countries out of poverty. In addition, the DRC is not the only country in which mining fuels wars, yet Dodd-Frank seems on its face to have no concern for these other regions.

Still, despite its apparently limited scope, Dodd-Frank likely can count on the “California Effect” to achieve far wider impact. Companies are likely to gradually monitor their supply chains worldwide and rid them of conflict minerals from all sources. In 2010, advocates in Washington secured the adoption of a small provision against one specific country. This Article shows that this minor provision is likely the precursor, thanks to the California Effect, to global regulatory harmonization on conflict minerals. Such harmonization would no doubt be much more difficult to reach through more traditional forms of international cooperation, particularly an international convention banning the use of conflict minerals. While global regulatory convergence had been discussed mostly in the context of antitrust, tax, privacy, and the environment, little had been written about unilateral regulatory approaches to problems at the intersection of business and human rights. This Article begins to fill this gap by showing that the United States’ unilateral regulations on conflict minerals were likely easier to pursue than conventional international initiatives, but could potentially be just as influential, if not more.