ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA
DEPARTMENT OF LEGAL STUDIES

VOL.6
ISSUE 1
2021

UNIVERSITY OF BOLOGNA LAW REVIEW

AlmaDLJournals
Open Access Scientific Journals
by AlmaDL University of Bologna Digital Library

E-ISSN 2531-6133

Doi: 10.6092/issn.2531-6133/v6-n1-2021

UNIVERSITY OF BOLOGNA LAW REVIEW

E-ISSN 2531-6133

GENERAL INFORMATION

The University of Bologna Law Review is a student-run, peer-reviewed, open access journal published by the University of Bologna, Department of Legal Studies. The Journal is published twice a year (all year around). The Journal is a Scopus Q3 indexed journal. Publication number E-ISSN 2531-6133. The Journal is hosted and mantained by ASDD-AlmaDL.

SUBMISSIONS

The University of Bologna Law Review accepts articles, essays, notes and conference papers. Authors can submit their manuscripts via email to bolognalawreview@unibo.it, via Scholastica, or via website.

SYSTEM OF CITATION

The University of Bologna Law Review follows The Bluebook: A Uniform System of Citation (20th Edition), copyright by The Columbia Law Review Association, The Harvard Law Review Association, the University of Pennsylvania Law Review, and the Yale Law Journal.

ADDRESS

University of Bologna, Department of Legal Studies, Via Zamboni 42, 40126, Bologna, Italy.

WEBSITE

www.bolognalawreview.unibo.it

ADVISORY BOARD

UNIVERSITY OF BOLOGNA LAW REVIEW

TABLE OF CONTENTS

Vol. 6.1

VOLUME 6 ISSUE 1 OBITER DICTUM

UNIVERSITY OF BOLOGNA LAW REVIEW

https://doi.org/10.6092/issn.2531-6133/12618

What COVID-19 does to our Universities

No mechanical means, however expeditous,
no materialism, however triumphant,
can eradicate the day break we experience
when we have understood a Master.[1]

Across the globe, COVID-19 is affecting multiple dimensions of our lives, ranging from the most existential aspects to the more mundane ones. In this brief article, I wish to reflect upon the impact the pandemic has on our Universities. Let me stress from the outset that I do not intend to downplay the terrible losses people all over the world have suffered. Nor do I wish to discuss the proportionality of the protective measures imposed by the governments, including those taken regarding research and teaching activities on University campuses. What I do want to address, however, is that COVID-19 currently functions as a fire accelerant, stimulating highly detrimental processes that flagrantly run counter to the idea of a University per se. It is vital to prevent COVID-19 from altering the University experience permanently. To accentuate my point, I will first lay down what I mean by the "idea of a University per se". Second, I will present a few thoughts on why COVID-19 has negative effects on our Universities.

The idea of a "University per se" goes back, *inter alia*, to the classical 19th-century theories of Alexander von Humboldt and John Henry Newman. According to these theories, Universities distinguish themselves from other kinds of professional instruction and training, by providing a liberal education. Understood in this way, education is a public good that pursues intellectual, cultural, and scientific aims. Education prepares and enables students to develop genuine autonomy, by introducing them to the open-ended search for deeper understanding and by stimulating them to form their minds and attitudes in a protected space. Universities also enable advanced scholarship. University scholars research fundamental questions, which do not exclusively answer to the immediate needs of current practical problems.

[1] George Steiner, *Lessons of the Masters*

By answering those more fundamental questions, University research is often the motor for grander and sustainable transformations in society. Further, both education and research are pursued not just in one single discipline or in otherwise narrow limits, but in an open-ended and unlimited discourse, which spans across multiple disciplines and approaches. Universities provide unique spaces for societal self-reflection, as well as an unparalleled room for the growth of individuals as persons. To achieve all this, Universities require as much autonomy as possible, so that they can live by their own intellectual rationale. External influence, stemming from political aims, social purposes, or economic pressure, is not only alien but also ruinous to it. The currency of Universities is truth, not money. Numerous other functions of a University have been added to this original idea of a University per se ever since. Not only were Universities transformed into degree-emitting factories for the greater part of any generation. Nowadays, they also engage in a wide range of activities, such as lifelong learning programmes, marketing, third party funding and collaborating with local and regional businesses and social institutions. In result, considering the pluralism of functions and this wide variety of activities, we work in a "multiversity". Yet, all these additional functions and activities are of a secondary character only. They are indirect consequences and contingent byproducts of their primary purpose. In view of the increasing pressure on scholars to engage in a preposterous race for external funding, it needs to be stressed that money is a precondition of good research – and not its aim.

This primary function, however, has been in decline for several decades now. Vast areas in our contemporary Universities are characterized by a highly "distracted, numbers-swamped, audit-crazed, grant-chasing" (Stefan Collini) atmosphere. Political influence and measures of the so-called "new public management" have turned many, if not most Universities, into organizations that are so far removed from the idea of a University that one wonders whether the use of the label "University" is still justified. Academic endeavours are commercialized and counted, and university leaders act as if they were leading economic enterprises. Research output is expected to be directly useful, allowing for an immediate implementation. In this way, Universities are pulled into the vortex of purposive research. Various external pressure groups, boards of governors, and artificially instituted competition are now influencing, if not imposing upon, strategic decisions in our Universities. Accountability measures have destroyed the institutional autonomy our Universities once enjoyed. Constant audit activities and chase for grants distract and hinder the scholars from living by their primary rationale.

These destructive processes are now augmented with COVID-19 and the way how University leaders, rectors, presidents and department chairs have reacted to the pandemic. Before I elaborate on these detriments, let me at least briefly acknowledge the paradox that, as far as research is concerned, at least in the humanities and social sciences, the pandemic turned out to have some positive impact. A significant number of talks, workshops, and conferences had to be cancelled during the lockdown. When not substituted by an online format, these cancellations gave scholars the most precious gift the world has for them: Time. Time to read. Time to write. As a reviewer, I already notice the higher quantity of research output following the lockdown period. One could doubt, of course, whether a higher amount of publication is good in itself. But still: more time for research is something we scholars regard highly, but experience rarely. This positive aspect apart, we are confronted with a deplorable picture: vacant campuses, closed libraries, empty seminar rooms. Exchange programmes have come to a halt, international student mobility has almost entirely subsided. Students and scholars report the terrible experience of the COVID-19 term. They miss the everyday realities of campus life: no-tech classes face-to-face, social and cultural diversions, extracurricular activities, in-person office hours, social interactions. The pandemic has triggered one of the most enormous disruptions in the history of University education. The muting of so many aspects of University life, closes down vital communication rooms and opportunities for actual face-to-face encounters, not just of knowledge and ideas, but of persons. Now, one may say that thanks to the various technologies we nowadays have at our disposal, we have alternative communicative spaces: online rooms. Many commentators praise digitalization as the ultimate benefaction, both of our lives in general, and of higher education in particular. The euphoria of digitalization is, however, fundamentally mistaken. The benefits of digitalization are overrated, while its downsides are grossly underestimated. This is a fact that both students and professors report alike after their first lockdown term. In a similar vein, I recently learnt that in my city, around 80 per cent of 3- to 6-year-olds are using tablets or mobile phones. I cannot bring myself to see this as an advantageous development. On the contrary, I wish these kids would browse storybooks instead. Many would want to convince us that a fully digitized University is the pinnacle of a movement toward an ever more inclusive and modernized research and learning environment. I find this idea utterly ridiculous. Reducing students to small images on video conference checkerboards does not support real education, but it hinders it. Surveys demonstrate that during the last COVID-term, students got ever more frustrated with their learning experience, the longer the online-teaching lasted. They miss the interactive, socializing life on campus.

Students report a significantly higher workload during the digitalized term, as well as problems to focus and to motivate themselves at home. They reproach the missing orientation when assessing their learning materials. Guidance must be given by professors in real presence. While it may be that some aspects of professional training can as well be achieved in an online format, University education is different. Education requires, what George Steiner named "real presence". "Real" means real, and not some zooming-skyping-webexing fake-type of real. A technophile fetishization of digital teaching should not deceive us here. One can certainly acquire a lot of knowledge, alone, in front of electronic devices. But when it comes to how this knowledge shapes the learner and affects society as a whole, it is impossible to replace the direct personal encounters of real education. Remote learning is nothing but a poor substitute.

It may be true that we cannot avoid the fake-type of real education for the time being, due to the pandemic. What I firmly believe, however, is the following: We must maintain our awareness of the differences between real and fake-real. University education is not a serial information programme that can be switched on and off as one pleases. And here is what alerted me so much in this respect: the phrasing Universities used to announce lockdown measures, was highly revealing. Many University leaders seemed to cherish that the pandemic now forces the "old dog" type of professors to learn the new tricks of digitized teaching. They conveyed the distorted image of online teaching as a special gift to the students. Allegedly supported by this image, it seemed as if they had finally found a justification for interfering with the professor's freedom of education.

What an absurd idea! Instead, Universities should have apologized to their students. They should have labelled digitized teaching as a poor copy and cheap imitation of real education. They should have mentioned humbly, that for the time being, they could, sadly, "only" provide online teaching. They should have promised to return to real education as soon as the situation would allow it. But instead, what we saw were some high-gloss marketing announcements, signalling that "all is fine" and "everything is under control". Some Universities even seemed to be tempted to prolong their lockdowns a bit longer, to make the digitalization more permanent. In short, the pandemic was abused to hinder the type of education students have a right to experience, and professors have a right to provide, if we take the idea of a liberal education seriously. A German minister even praised the digitalization of teaching as the ultimate solution to the tense situation of the housing market in many University cities: after all, online-students no longer need to live near their Universities. This statement reveals a profound misunderstanding of the function of Universities: an education presupposes the whole spectrum of horizon-broadening experiences that come with

living in a University city. And the statement also displays a cynical vision of how we should develop and improve the social structure in our cities. The educational losses of the pandemic now threaten to extend beyond the current generation, erasing decades of progress, as far as the opportunities for the most vulnerable young adults from marginalized groups are concerned. Persons living in rural or impoverished areas, refugees or persons with disabilities rarely have access to a fully equipped home-office with a high-speed internet connection. And even if they had, distance learning is second class learning from the start. I do not deny the necessity to build resilient educational systems and the need to be adaptive and innovative in improving our means to provide high-quality education and research. Instead, I want to remind all of us that it is at least as important, if not more, to rediscover the original idea of a University. Universities must proactively improve their institutional abilities to change. But they must most vigorously avoid doing so at the expense of their academic core mission. While the members of Universities nearly perish by drowning in the flash floods of endless calls for adaptation, agility, and attitude reframing, I urge all of us to vividly remember our origins. The invocation for education and research is alarmingly silent in the ever-growing, excessive hullabaloo of innovation and change. In this respect, the pandemic, with all its detriments, may at least give us a forceful reminder of what Universities are for, and what conditions they need to flourish. The highly detrimental influence of ministers of higher education, boards of governors, and external funding providers needs to be stopped. The pandemic could help us revitalize our understanding of the nature and importance of Universities. This understanding is so much in danger of being lost sight of. While the immediate effects of COVID-19 have been dramatic and detrimental, mid-term, it also brings with it the opportunity to refocus on higher education and advanced research as a public goods, and to rediscover our Universities' true mission.

In a major weekly German newspaper, one could read a miniature recently written by a University professor. She reported that she was photocopying on a late Friday afternoon, just before the term started. When the job was done, and the machine fell silent, she first felt isolated and alone in the vast and utterly empty building. Yet suddenly, she heard distant music. Going about to find out, she discovered a violinist playing in a seminar room. In her concluding memento, she praises the place of a University campus: "People not only work here – they live here". If only we could revive this spirit again soon. Until then, let us at least remember it with all our forces.

Matthias Klatt
Chair of Jurisprudence
University of Graz, Austria

VOLUME 6 ISSUE 1 ARTICLES & ESSAYS

UNIVERSITY OF BOLOGNA LAW REVIEW

https://doi.org/10.6092/issn.2531-6133/12880

Received: 12 Sept. 2020 | Accepted: 30 Sept. 2020 | Published: 5 May 2021

What Is Next for Digital Trade in a Post-Brexit Britain? – Examining the Regulation of Data Flows Under G.A.T.S. & Possible Implications of G.D.P.R. on Britain as a Third Country

ABHISHEK RANA* & RISHABHA MEENA**

* Abhishek Rana is a Lecturer at the Jindal Global Law School, O.P. Jindal Global University, India. He is also a visiting faculty in W.T.O. law at the Indian Society of International Law.
** Rishabha Meena is a Research Fellow at the Centre for Trade and Investment Law at the Indian Institute of Foreign Trade.
The authors would like to thank Neha Mishra and Svetlana Yakovleva for their precious comments on earlier drafts. They also wish to specify that all errors remain theirs alone.

@ * abhishekrana@jgu.edu.in ** rishabh.149620@gmail.com

ABSTRACT

Data, much like other currencies, flows cross-border -from one jurisdiction to the other. However, it is hard to regulate the privacy aspects surrounding such free-flowing data by rules strictly based on jurisdiction. This article thereby begins by discussing the importance of data protection regulations like the General Data Protection Regulation (G.D.P.R.), followed by a brief analysis of the General Agreement on Trade in Services' pivotal role in regulating data flows and digital trade, and how it can be further used in checking the World Trade Organisation consistency of various data protection requirements resorted by the European Union (E.U.) so far under the G.D.P.R.. Lastly, the note examines how, post the Brexit transition period, the situation will change for the United Kingdom (U.K.) as it has become a third country for the E.U. data protection regime, with the authors critiquing the various models, including the recent Draft U.K.-E.U. Comprehensive Free Trade Agreement, that may help the U.K. in attaining an "adequacy" status, which is requisite for the continuation of an unconstrained digital trade with the E.U. .

KEYWORDS

World Trade Organization; General Data Protection Regulation; Brexit; General Agreement on Trade in Services; Data

TABLE OF CONTENTS

INTRODUCTION

In the modern era, the privacy of a person is an essential facet of human rights law which necessitates legal protection,[1] and it has been enshrined in multiple international instruments covering fundamental human rights.[2] For instance, under Article 17 of the International Covenant on Civil and Political Rights, a person cannot be subjected to "arbitrary or unlawful interference with his privacy, family, home or correspondence, nor to unlawful attacks on his honour and reputation".[3] In furtherance of the same, several countries have either amended their existing legislations or have brought in new regulations to deal with the issue of privacy.[4] The debates around data privacy, apart from being addressed from a human rights perspective, can also be understood from an international economic law lens as privacy aspects heavily influence the international trade of data; although the domestic data protection regimes of the World Trade Organisation's [hereinafter W.T.O.] Members have tried to act as barriers to such free flow of digital trade.[5]

[1] Samuel D. Warren & Louis D. Brandeis, *The Right to Privacy*, 4 Harv. L. Rev. 193 (1890).

[2] International Covenant on Civil and Political Rights, art. 17, Dec. 16, 1966, UNTS vol. 999, 171 (hereinafter ICCPR); *See* Human Rights Committee, General Comment No. 16 (Thirty-second session, 1994), Compilation of General Comments and General Recommendations Adopted by Human Rights Treaty Bodies, UN Doc. HRI/GEN/1/Rev.1, 21 (1994).

[3] ICCPR, Art. 17.

[4] Australia is governed by "Australian privacy principles", see Privacy Act 1988 (Cth) sch 1 ; Canada is governed by the Personal Information Protection and Electronic Documents Act, S.C. 2000, c. 5; India is governed by the Information Technology Act, 2000, No. 21, Acts of Parliament, 2000 while the Personal Data Protection Bill, 2019, No. 373 is yet to be become an Act.

[5] Svetlana Yakovleva & Kristina Irion, *The Best of Both Worlds? Free Trade in Services and EU Law on Privacy and Data Protection*, 2 Eur. Data Prot. L. Rev.191 (2016). *See* Rolf H. Weber, *Regulatory Autonomy and Privacy Standards under the G.A.T.S*, 7 Asian J. WTO & Int'l Health L & Pol'y 25 (2012).

Even the "global organisations" that are operating their businesses in multiple countries, the privacy aspect of their business is generally governed by the domestic legislation(s) of the state in which their individual offices are based.[6] However, situations requiring a transfer of data from one country to another, in due course, may lead to different legal systems coming into conflict with each other, especially in determining the adequate standard of data protection.[7] This issue might entail another analogous debate on whether data protectionism hinders globalisation[8] but the authors have not dealt with that question in this article.

Amongst the various existing data protection laws around the world, the European Union's [hereinafter E.U.] General Data Protection Regulation [hereinafter G.D.P.R.] is one of the most comprehensive ones, and has become a global standard for most of the countries.[9] The G.D.P.R. goes far beyond merely being domestic legislation as it has implications on other countries as well, wherein the underlying detailed and specific regulatory standards can be imposed on non-E.U. or European Economic Area [hereinafter E.E.A.] based companies, involved in gathering or transfer of data. While the EU and E.E.A. member states enjoy an unrestricted flow of data, it becomes a tricky situation when one Member from the Union decides to pull out. On March 29, 2017, the United Kingdom [hereinafter U.K.] decided to withdraw its membership from the E.U. . Brexit saw the fall of two experienced Conservative Party leaders who were initially pro-E.U. . After Britain officially left the E.U. ("Brexit") on January 31, 2020, it entered a transition period – as per which it remains a part of the E.U. customs union; although it no longer is a part of the political institutions of the E.U. . The concerned Withdrawal Agreement has allowed E.U. law to be implemented in Britain until December 31, 2020.

Post this transition period, the digital trade that is currently being overseen by the G.D.P.R. for data protection purposes, is going to get hampered. There will be an automatic ban on any "default transfer" of personal data from the E.U. members to the U.K.. Consequentially, the legal entities existing in different E.U. member states will no longer be able to transfer personal data as per the earlier U.K.-E.U. relationship. The transfer of data will only become possible when the U.K. is able to show adequate level of protection; or there are appropriate protection measures like Binding Corporate Rules [hereinafter B.C.R.s] or Standard Contractual Clauses [hereinafter S.C.C.s]; or other conditions which are part of Chapter V of G.D.P.R. which deals with transfer of personal

[6] Robert L. Totterdale, *Globalization and Data Privacy: An Exploratory Study*, Int'l J. Info. Sec & Priv., Spring 2010, at 19.
[7] *Id.*
[8] Dan Gunderman, *Are Data Privacy Regulations Hindering Globalization*, Cyber Security Hub (Dec. 7, 2018) https://www.cshub.com/data/news/are-data-privacy-regulations-hindering-globalization.
[9] Paul M. Schwartz, *Global Data Privacy: The EU Way,* 94 N.Y.U. L. Rev. 771 (2019).

data to third countries or international organisations. Some of these requirements which form part of Chapter V, in the long run, are speculated to be a significant obstruction in digital trade flows for all the third countries, and not just the U.K., as it would require the third countries to align their data protection laws[10] as per the G.D.P.R., and also factor in other privacy related aspects that are part of the legal framework of the E.U. .[11]

This note, in Part II, traces the importance of G.D.P.R., and reflects on how it will impact the digital trade flows between the E.U. and third countries as some of its provisions impose obligations relating to the E.U. data subjects even on the companies based outside the E.U. . Part III provides a basic understanding of the General Agreement on Trade in Services' [hereinafter G.A.T.S.] framework responsible for regulating the trade in services and provisions essential to govern the transboundary flow of data. It assesses the G.A.T.S. consistency of various G.D.P.R. provisions providing for the usage of either data protection measures such as S.C.C.s and B.C.R.s or availing an "adequacy status" for the third countries. Part IV delves into the options available to the U.K. for regulating and protecting the transfer of data post-transition period with the E.U./E.E.A. . The authors analyse the two possible options that the U.K. is currently exploring for a smoother future exchange of data, while also taking into account other non-trade concerns that might come with these options. The two options that have been critically analysed include: (i) the recent Draft E.U.-U.K. Comprehensive Free Trade Agreement [hereinafter E.U.-U.K. C.F.T.A.] released in May 2020, which while excludes data adequacy, interestingly includes "digital trade" under the draft agreement; or (ii) U.K. showing an adequate level of data protection under its domestic legislation, which can later meet the equivalence of data standards set by the E.U. . Finally, the note recommends the best option that the U.K. might have moving forward, and what steps the private parties can take to protect their data in case there is a delay in finalising a possible arrangement between the post-transition period U.K. and the E.U. .

[10] The requirement of an equivalent level of data protection or alignment of law has been imposed by G.D.P.R. because a lower-level data protection in the transferee country poses threat and risks to the privacy of the individual. Lee A. Bygrave, Data Protection Law: Approaching Its Rationale, Logic and Limits 79 (2002).

[11] Graham Greenleaf, *"European" Data Privacy Standards Implemented in Laws Outside Europe* 149 Privacy Laws & Business International Report 21-23 (2017) Additionally, it is not only G.D.P.R. with which the UK has to comply with. Under Article 45(2), the UK is also required to take into account the national security laws, defence laws, human rights laws, etc; C-311/18, Data Protection Commissioner v. Facebook Ireland Ltd., 2020 E.C.R.

1. DATA PROTECTION & DIGITAL TRADE GOVERNED BY G.D.P.R.

The E.U.'s data protection regime since the 1990s has been extensively regulated and was always applied uniformly for all types of personal information; unlike in the United States [hereinafter U.S.] where generally sector-specific data privacy guidelines exist for the regulation of personal information - varying from industry-to-industry.[12] So, with the introduction of the revolutionary G.D.P.R.,[13] replacing 1995 Data Protection Directive,[14] all the legal entities and undertakings[15] (majorly businesses and public bodies) covered under this regulation were given two years to prepare themselves for the changes,[16] with the G.D.P.R. finally entering into force on May 25, 2018.

G.D.P.R. led to an introduction of specific new rules along with the augmentation of the already existing regulations,[17] with heavy penalties if the requirements as per the regulations are not complied with. The Data Protection Authorities have the power to impose strict sanctions in the form of progressive fines up to €20m or up to 4% of an undertaking's worldwide annual turnover.[18] Another vital distinction to note here is that the E.U.-wide data protection law is in the form of a regulation,[19] and not a directive, making it directly applicable to the member states without a need to be incorporated into their domestic laws.[20] Thus, G.D.P.R.'s enforcement affected nearly every company in the E.U., but the ones who are affected by it the most are responsible for holding, controlling, and processing[21] the data of the consumers,[22] and it majorly includes the parties in a digital economy, viz. the technology firms and the marketers, along with the

[12] Kurt Wimmer & Joseph Jones, *Brexit and Implications for Privacy*, 40 Fordham Int'l L.J. 1553 (2017). However, in case of California, there is no sector specific laws for data protection. The California Consumer Privacy Act 2018 is single legislation governing data protection which applies to all types of business upon fulfilment of certain criteria.

[13] Council Regulation 2016/679/EC of the European Parliament and of the Council on the Protection of Natural Persons with regard to the Processing of Personal Data and on the Free Movement of such Data and Repealing Directive 95/46/EC, 2016 O.J. (L119) 1 [hereinafter G.D.P.R.].

[14] Directive 95/46/EC of the European Parliament and of the Council of 24 October 1995 on the protection of individuals with regard to the processing of personal data and on the free movement of such data, 1995 O.J. (L 281).

[15] Mario Rosentau, *The General Data Protection Regulation and Its Violation of EU Treaties*, Juridica Int'l, March 2018, at 36, 38.

[16] Matt Burgess, *What is GDPR? The summary guide to GDPR compliance in the UK*, wired.co.uk (Mar. 24, 2020), https://www.wired.co.uk/article/what-is-gdpr-uk-eu-legislation-compliance-summary-fines-2018.

[17] Wimmer & Jones, *supra* note 12, at 1555.

[18] Alex Hern, *What is GDPR and how will it affect you?*, The Guardian (May, 21, 2018), https://www.theguardian.com/technology/2018/may/21/what-is-gdpr-and-how-will-it-affect-you.

[19] Section 2 of the European Communities Act 1972 provides that all the EU regulations and treaties are applicable to the U.K. without express incorporation under a domestic law.

[20] Wimmer & Jones, *supra* note 12, at 1555.

[21] G.D.P.R. makes a differentiation between "controller" and "processor" of data.

[22] Wimmer & Jones, *supra* note 12, at 1554.

data brokers who help in connecting such technology firms to marketers.[23] These companies are presumed to incur the maximum amount of legal and technical costs.[24]

G.D.P.R. continues to allow for a free transfer of personal data between the E.U./E.E.A. member states, insofar as the rules for the protection of such data are followed. Data can also be transferred to third countries, but only in certain circumstances, which generally includes either (i) transfer of data to a third country whose data protection regime has been considered by the European Commission to be "adequate" (deciphered to mean "essentially equivalent" to the data protection regime in the E.U.[25]) or (ii) relying on other safety measures (like S.C.C.s or B.C.R.s)[26] with a business organization existing in a third country that has not yet obtained 'adequacy status'.[27] Personal data can be transferred to third countries outside the E.U./E.E.A. without these usually relied upon options too, in special situations and single cases.[28]

The E.U. members have long enjoyed the benefits of having an unimpeded transfer of data, especially with the E.U. proactively signing different arrangements with third countries ranging from the concept of sharing of law enforcement data[29] (e.g., under the E.U.-U.S. "Umbrella" Agreement)[30] to other data sharing courses of action (e.g., the E.U.-Canada Passenger Name Record Agreement).[31] However, after the transition period, the U.K. will turn into a "third country" for data regulation in the E.U., which implies that the personal data is bound to be severely confined.[32] Therefore, to guarantee the transfer of such personal data with few limitations or restrictions between the E.U. and the U.K., the U.K. government will have to assess whether it can accede to

[23] *Id.*
[24] Rosentau, *supra* note 15, at 38.
[25] C-362-14, Max Schrems v. Data Protection Commissioner, 2015 E.C.R. 73,74.
[26] Jeffrey N. Berman, *GDPR – What are the Model Contract Clauses?*, BFVLAW.COM (Dec. 20, 2018), https://www.bfvlaw.com/gdpr-what-are-the-model-contract-clauses/.
[27] As of now, Andorra, Argentina, Canada (commercial organisations), Faroe Islands, Guernsey, Israel, Isle of Man, Japan, Jersey, New Zealand, Switzerland, Uruguay and the United States of America (limited to the Privacy Shield framework) have been granted the adequate standard.
[28] Such special situations and single cases are applicable when the data subject has given consent to such transfer even after knowing all the risks; usually in circumstances where such a transfer is necessary like in the instance of public interest. However, these cases become applicable only when (i) there is a single occasion; (ii) it concerns limited data subjects; and (iii) takes place after weighing of interest .GDPR art. 49; *Transfer of data to a third country*, DATAINSPEKTIONEN.SE, https://www.datainspektionen.se/other-lang/in-english/the-general-data-protection-regulation-gdpr/transfer-of-data-to-a-third-country/.
[29] Wimmer & Jones, *supra* note 12, at 1556-1557.
[30] Agreement between the United States of America and the European Union on the Protection of Personal Information Relating to the Prevention, Investigation, Detection, and Prosecution of Criminal Offences, 2016 O.J. (L 336/3) 3-13.
[31] Draft agreement between Canada and the European Union on the transfer and processing of Passenger Name Record data. *See* Proposal for a Council Decision on the Conclusion of the Agreement Between Canada and the European Union on the Transfer and Processing of Passenger Name Record Data, COM (2013) 528 final (July 18, 2013).
[32] Stuart Anderson, *Brexit Data Confusion*, INT'L J. DATA PROT. OFF., PRIV. OFF. & PRIV. COUNS, 2018, Volume 2 at 10.

the existing arrangements or manage to get a separate bilateral agreement. It is vital to maintain and potentially reproduce the current arrangements as it would let the critical flow of personal data to continue, and would further help in settling any international legal disputes that may arise from time to time when the companies have to honour the requests for data to be shared for law enforcement purposes.[33]

2. REGULATING DATA FLOWS & DATA PRIVACY MEASURES UNDER G.A.T.S.

The W.T.O.'s establishment, on 1 January 1995[34], led to an emergence of an effective administrative and juridical system in the rules of global trading. The multilateral agreements under the W.T.O., which are attached as Annexes to the Marrakesh Agreement, not only dealt with trade in goods[35], but also in services[36] and intellectual property rights[37]. It was the modernisation, ubiquitous digitisation, and free flow of data through the internet which unravelled the need for regulating trade in digital economic sectors[38] with the help of G.A.T.S. .[39] Thus, G.A.T.S. has dealt with such issues primarily through five obligations and commitments, viz. (i) violation of the Most-Favored Nation [hereinafter M.F.N.] treatment[40], (ii) violation of the Market Access [hereinafter M.A.] commitments[41], (iii) violation of National Treatment [hereinafter N.T.][42], (iv) domestic regulations[43] and (v) general[44] and security[45] exceptions to M.F.N., N.T. and M.A. commitments.

[33] Wimmer & Jones, *supra* note 12, at 1559.
[34] Marrakesh Agreement Establishing the World Trade Organization, Apr. 15 1994, 1867 UNTS 3, 1868 UNTS 3, 1869 UNTS 3.
[35] General Agreement on Tariffs and Trade 1994, Apr. 15, 1994, Marrakesh Agreement Establishing the World Trade Organization, Annex 1A, 1867 U.N.T.S. 187, 33 I.L.M. 1153 (1994) [hereinafter GATT 1994].
[36] *See* General Agreement on Trade in Services, Apr. 15, 1994, Marrakesh Agreement Establishing the World Trade Organization, Annex 1B, 1869 U.N.T.S. 183, 33 I.L.M. 1167 (1994) [hereinafter G.A.T.S].
[37] Agreement on Trade-Related Aspects of Intellectual Property Rights, Apr. 15, 1994, Marrakesh Agreement Establishing the World Trade Organization, Annex 1C, 1869 U.N.T.S. 299, 33 I.L.M. 1197 (1994) [hereinafter TRIPS Agreement].
[38] Margaret Byrne Sedgewick, *Transborder Data Privacy as Trade*, 105 Cal. L. Rev 1513, 1542 (2017).
[39] G.A.T.S. in its Article I:2 mentions about governing "trade in services" through four modes of supply, which are "cross border supply" (mode 1), "consumption abroad" (mode 2), "commercial presence" (mode 3), and "presence of natural persons" (mode 4); *See* Weber, *supra* n. 5.
[40] G.A.T.S., *supra* note 36.
[41] *Id.* Art. XVI.
[42] *Id.* Art. XVII.
[43] *Id.* Art. VI.
[44] *Id.* Art. XVI.
[45] *Id.* Art. XVII.

In this part, the authors explore the consistency of various regulations of G.D.P.R. in light of the commitments undertaken by the E.U. under G.A.T.S., as these regulations form the basis of international data transfer with third countries.

2.1 ASSESSMENT OF G.D.P.R. WITH E.U.'S SCHEDULE OF COMMITMENT

The G.A.T.S. borrowed the basic principles of non-discrimination and market access, amongst others, from the General Agreement on Tariffs and Trade [hereinafter G.A.T.T.] 1947, and in so doing applied these principles in service sectors such as finance, communication, digital market, education, tourism, licensing, *et cetera*.[46] In principle, any limitations amounting to zero quota or total ban put forth on the transfer of any data to other Members or third countries will constitute a restriction on M.A. . [47]

However, services that are covered under G.A.T.S. are not automatically opened to competition, except in the case of general commitments like most favoured nation treatment. W.T.O. Members guarantee the access to their domestic markets only for those sectors or modes of supplies that have been specified in their "schedule of commitments", and further accepted by individual Members with or without limitations.[48] This schedule is legally binding and an integral part of the G.A.T.S.,[49] where each Member specifies in which sector, in relation to what service, and to what extent the commitment taken shall bind them. Members can select to either be fully bound[50] or be bound with limitations.[51] Therefore, it becomes vital to, firstly, classify the category in which the data transfer services would fall; secondly, identify the mode through which such services are supplied; and lastly, assess the consistency of the data regulation provisions in accordance with the commitments undertaken in the national treatment and market access categories. These requirements together "*create the tapestry of fundamental principles used by most nations to regulate international trade*" in data services.[52]

[46] *Id.* Art. I; *See also The most dynamic segment of International Trade: Trade in Services*, WTO.ORG (2015) https://www.wto.org/english/thewto_e/20y_e/services_brochure2015_e.pdf.

[47] *Understanding on commitments in financial services*, https://www.wto.org/english/tratop_e/serv_e/21-fin_e.htm. This principle can be applied to all types of services involving data transfer.

[48] *See Schedules of specific commitments and lists of Article II exemptions*, WTO.ORG, https://www.wto.org/english/tratop_e/serv_e/serv_commitments_e.htm.

[49] Svetlana Yakovleva & Kristina Irion, *The Best of Both Worlds - Free Trade in Services and EU Law on Privacy and Data Protection*, 2 EUR. DATA PROT. L. REV. 191, 192 (2016).

[50] A full commitment of market access means a prohibition to maintain, predominantly qualitative, market access barriers included in the exhaustive list of Art. XVI: 2. Yakovleva & Irion, *supra* note 5, at 191.

[51] If a party wants to preserve certain market access barriers banned by Art. XVI:2 in sectors and in relation to modes of supply where it undertook a specific market access commitment, such limitations should be included in the Services Schedule in the column 'Limitations on Market Access'. *See Guide to reading the GATS schedules of specific commitments and the list of Article II (MFN) exemptions*, WTO.ORG, https://www.wto.org/english/tratop_e/serv_e/guide1_e.htm.

[52] Joshua D. Blume, *Reading the Trade Tea Leaves: A Comparative Analysis of Potential United States WTO – GATS claims against Privacy, Localization, and Cybersecurity Laws*, 49 GEO. J. INT. L. REV. 801, 807 (2018).

So, to determine the type of service under which data transfer through digital trade would fall under, we take a cue from the 1993 Scheduling Guidelines[53] that illustrated the need for a classification of sectors and sub-sectors based on the Services Sectoral Classification List,[54] which is further based on the United Nations Central Product Classification[55]. While classifying a particular service, the focus must be laid on the teleological interpretation of the character of that service, the purpose that is being achieved,[56] and any other component that would give an essential characteristic to that particular service.[57] From the explanatory notes for the services in the U.N. Central Product Classification, it is evident that the data & message transmission services consist of network services that transfer data, receive and send electronic messages, and manipulate the information in databases.[58] The critical feature of telecommunication services is the transmission and reception of signals by electromagnetic waves.[59] Since data transfer requires the transmission and reception of signals,[60] data transfer services are classified under the telecommunication ones.

Secondly, under Article I:2 of G.A.T.S., services are supplied through four different modes, viz. cross-border (mode 1), consumption abroad (mode 2), commercial presence (mode 3), and movement of natural persons (mode 4). The supply of services from territory of one W.T.O. Member to the territory of another W.T.O. Member, without the presence of the service supplier in the territory of the service receiver, is the cross-border supply of service.[61] Therefore, the international data transfer is cross-border of service as the transfer occurs without the presence of the service provider in the destination country.

Thirdly, the W.T.O. Members also undertake specific commitments for M.A. and N.T. under Article XVI and Article XVII of the G.A.T.S. respectively, with the W.T.O. consistency of a measure determined based on the commitments taken. A Member, therefore, can rely on one of the three forms, by opting for "none" (no limitations to be

[53] *See* Group of Negotiations on Services, Uruguay Round, Scheduling of Initial Commitments in trade in Services, Explanatory Note, GATT Doc. MTN.GNS/W/164 ¶ 4 (Sept. 3, 1994).
[54] GATT Secretariat, *Services Sectoral Classification List*, GATT Doc. MTN.GNS/W/120 (July 10, 1991).
[55] U.N. Dep't of Int'l Econ. & Soc. Affairs, Provisional Central Product Classification, U.N. Doc. ST/ESA/STAT/SER.M/77 (2002).
[56] *See* Rolf H. Weber & Mira Burri, Classification of Services in the Digital Economy 93 (2013).
[57] *See* Ines Willemyns, *GATS Classification of Digital Services – Does 'The Cloud' Have a Silver Lining?,* 53 J. World Trade 59, 76 (2019).
[58] U.N. Dep't of Int'l Econ & Soc. Affairs, Provisional Central Product Classification, 140 Part II, U.N. Doc. ST/ESA/STAT/SER.M/77, U.N. Sales No. E.91.XVII.7 (1991), Part III, 223.
[59] G.A.T.S, Annex on Telecommunication Service.
[60] *See Data Transmission – Parallel v. Serial*, Quantil, https://www.quantil.com/content-delivery-insights/content-acceleration/data-transmission/.
[61] Panel Report, *Mexico – Measures Affecting Telecommunications Services*, ¶ 7.28, WTO Doc. WT/DS204/R (adopted June 1, 2004) [hereinafter Panel Report, Mexico – Telecommunications].

imposed),[62] "unbound" (any type of limitations can be imposed),[63] or bound (specific type of limitations are agreed upon to be imposed) in its Schedule of Commitments. In the case of G.D.P.R., the E.U. has not undertaken any commitments for the cross-border mode of supply in the telecommunication services by inscribing "none" under the concerned Schedule,[64] which means that it is now prohibited from imposing any restriction(s). Therefore, the requirements such as S.C.C.s[65] or B.C.R.s[66] or even the adequacy status becomes difficult to justify.

Article XVI:2 G.A.T.S., dealing with M.A., provides for a list of measures that should not be maintained in sectors where a country has undertaken full M.A. commitments.[67] Thus, the M.A. commitments which are not indicated in a Member's schedule of commitments, like the setting up of high privacy standards; or by providing any particular benefits on providing source code; or registration requirements for data collection[68], would end up being violative of Article XVI:2(a) and (c) of the G.A.T.S. .[69] It was in the *U.S. - Gambling*[70] dispute where the Appellate Body held that the remote supply of betting and gambling is to be categorised as *cross-border electronic delivery of services*, with the disputed measure of limiting the number of service suppliers held to be a quantitative restriction on cross border supply, thereby violative of the M.A. principle.[71] Also, considering the ratio in *Mexico - Telecom*, where the Panel held that the routing requirement for international telecommunication is inconsistent with Article XVI:2 (a), (b), and (c).[72] Therefore, since the E.U. has inscribed "none" in the "limitations on market access" column, it means that the G.D.P.R. cannot have an effect of zero quota

[62] Panel Report, *United States — Measures Affecting the Cross-Border Supply of Gambling and Betting Services*, ¶ 6.279, WTO Doc. WT/DS285/R (adopted Nov. 10, 2004) [hereinafter Panel Report, *US — Gambling*"].

[63] *See* MITSUO MATSUSHITA ET AL., THE WORLD TRADE ORGANIZATION LAW, PRACTICE, AND POLICY 598 (3rd ed. 2015).

[64] G.A.T.S, *supra* note 36.

[65] The Standard Contractual Clauses [hereinafter S.C.C.s] are the standard sets of contractual terms and conditions governing data transfer between the EU or E.E.A. and the non-E.U. entities which must be present in the contract between these two entities (if the non-E.U. countries has not obtained adequacy status). It ensures the compliance with the G.D.P.R. requirements by the non-E.U. entities at private individual level. *See Standard contractual clauses for data transfers between EU and non-EU countries (European Commission)*, https://ec.europa.eu/info/law/law-topic/data-protection/international-dimension-data-protection/standard-contractual-clauses-scc_en.

[66] Binding Corporate Rules are legally binding and enforceable internal rules and policies to govern international data transfers within the organization having its branches in the EU and outside the EU, where the country in which non-EU branch of the organization is located has not obtained adequacy status. *See Binding Corporate Rules - The General Data Protection Regulation*, PWC (2019), <https://www.pwc.com/m1/en/publications/documents/pwc-binding-corporate-rules-gdpr.pdf.

[67] Panel Report, *United States - Gambling*, *supra* note 62, § 6.298.

[68] Weber, *supra* note 5, at 25. Rolf H. Weber, *Regulatory Autonomy and Privacy Standards under the GATS*, 7 ASIAN J. WTO & INT'L HEALTH L & POL'Y 25 (2012).

[69] *Id.*

[70] Appellate Body Report, *United States Measures Affecting the Cross-Border Supply of Gambling and Betting Services*, WTO Doc.WT/DS285/AB/R (Apr. 7, 2005) [hereinafter Appellate Body Report, *US - Gambling*].

[71] Yakovleva & Irion, *supra* note 5, at 212.

[72] Panel Report, *Mexico - Telecommunications*, *supra* note 61, § 7.85.

or a complete prohibition. However, the requirements such as S.C.C.s, B.C.R.s, and other safety measures as required by the E.U. can be argued to be inconsistent with the E.U.'s obligations,[73] given that such high privacy standards may act as quantitative restrictions having the effect of limiting the supply of services – especially in the sectors like finance and banking, transport services (in terms of storing passenger data), global accounting firms, etc. In the recent 2020 decision of *Data Protection Commissioner v Facebook Ireland Limited, Maximillian Schrems (Schrems II)*, the Court validated the S.C.C.s but imposed an obligation on the data exporter to ensure the safeguard of data which being exported to a third country which must take into account the legal system of the importing country and the data access by the third country.[74] The non-fulfilment of aforementioned requirements would consequently lead to non-delivery of services, thereby naturally limiting the number of service suppliers, which can be successfully challenged under Articles XVI:2(a) and XVI:2(c) of G.A.T.S. .[75]

Meanwhile, the requirement of N.T. under G.A.T.S. is violated when a W.T.O. Member provides less favorable treatment to foreign services than the domestic services.[76] However, contrary to the G.A.T.T., N.T. is not a mandatory obligation, but a conditional one, with its violation based on the limitations existing in a Member's Schedule of Commitments.[77] Article XVII G.A.T.S. requires Members to accord "like" foreign services and service suppliers with either *de jure* or *de facto* "treatment no less favourable" than their domestic counterparts.[78] In determining whether the product is "like" or not, the Dispute Settlement Body [hereinafter D.S.B.] has adjudicated on the four parameters given by working parties on Border Tax Adjustment.[79] An interesting case study on the national treatment provision under G.A.T.S is *China-Publications and Audiovisual Products*.[80] The measure at dispute herein were the restrictions placed on the

[73] Blume, *supra* note 52, at 822-23; *see also* Daniel Crosby, Analysis of Data Localization Measures Under WTO Services Trade Rules and Commitments 7 (2016). (E15 Initiative Policy Brief, 2016).
[74] Case C-311/18 Data Protection Commissioner v. Facebook Ireland Ltd., para. 104-105 (2020).
[75] Weber, *supra* note 5, at 33.
[76] G.A.T.S., *supra* note 36.
[77] Panel Report, *China – Measures Affecting Trading Rights and Distribution Services for Certain Publications and Audiovisual Entertainment Products*, ¶7.950, WTO Doc. WT/DS363/R and Corr.1 (Jan.19, 2010) [hereinafter Panel Report, *China- Publications and Audiovisual Products*].
[78] *See* Aditya Mattoo, *Services Globalization in an Age of Insecurity: Rethinking Trade Cooperation* (World Bank Group Policy Research Working Paper 8579, 2018).
[79] Four parameters given in the working party report include – (i) whether the product is "similar", (ii) product's end-uses, (iii) consumers' tastes and habits, (iv) product's properties, nature and quality. *See* General Agreement on Tariffs and Trade, *Report by the working parties on Border Tax Adjustments,* L/3464, (adopted Dec.2, 1970) [hereinafter G.A.T.T.].
[80] Appellate Body Report, *China – Measures Affecting Trading Rights and Distribution Services for Certain Publications and Audiovisual Entertainment*, WTO Doc. WT/DS363/AB/R (Dec. 21, 2009) [hereinafter Appellate Body Report, *China- Publications and Audiovisual Products*]. After US-Gambling Appellate Body Report, China- Publications and Audiovisual Products Appellate Body Report is the only case that dealt with public morality under General Security Exception.

foreign-invested enterprises in China relating to the importation and distribution of *foreign* publications, electronic publications, and audiovisual products. The Chinese government imposed harsh and onerous content review requirements on foreign products in comparison to their like domestic products. The Appellate Body held that such Chinese measures are inconsistent with Article XVII as China undertook no commitments under the concerned sector, and thereby such application of content review requirement was held to be violative of Article XVII.[81]

Similarly, it can be contended that international data transfer rules under the G.D.P.R. may lead to a potential violation of N.T. principle as the non-E.U. entities have to comply with additional requirements such as adequacy status, B.C.R.s or S.C.C.s, and other special situations which may cause additional burdens, economic as well as administrative, on the foreign services. However, footnote 10 to Article XVII creates an exception by excluding any inherent competitive disadvantages which result from the foreign character of the relevant services or service suppliers as a ground for violation of the N.T.;[82] hence, the E.U. can justify such requirements under Footnote 10. Also, the level of data protection has a significant influence on the "likeness" analysis of the domestic and foreign services which may make them "unlike",[83] and the E.U. can easily rely on the argument that the services provided by the E.U. and foreign services are not alike.

2.2. EXCEPTIONS TO NON-DISCRIMINATION AND MARKET ACCESS COMMITMENT

The exceptions clause in the W.T.O. agreements exists to allow the Members to adopt such measures that are otherwise W.T.O. inconsistent with the objective of enabling the members to pursue their policy objectives that are legitimate and important.[84] The general exceptions clause under Article XIV of G.A.T.S. and Article XX of G.A.T.T. provide

[81] *See* Panel Report, *China- Publications and Audiovisual Products*, ¶ 7.1097, 7.1142, 7.1170 , WTO Doc. WT/DS363/R and Corr.1 (Jan.19, 2010).

[82] *See* G.A.T.S., Apr. 15, 1994, Art. XVII ,1869 U.N.T.S. 183.

[83] For instance, the data localization requirements in Europe for the business to business services has modified the competitive relationship between the services and service supplier because of the evolution of localization requirements into recurrent requirement of the enterprise customers. So, services coming from countries which have the adequate level of protection would receive favourable treatment; however, such services would not be considered "like" for the national treatment test. *See* Svetlana Yakovleva & Kristina Irion, *The Best of Both Worlds? Free Trade in Services and EU Law on Privacy and Data Protection,* 2 Eur. Data Prot. L. Rev. 191, 204 (2016). It must be noted that national treatment requirement would depend on claim brought under different sector having a connection with data transfer (such financial services, etc) and hence no specific analysis is given to any one service over another. Blume, *supra* note 52, at 807.

[84] *See* Appellate Body Report, *US – Gambling*, *supra* note 70, § 290-291.

different grounds for when the W.T.O. Members can get away from their agreed obligations and commitments. Under the G.A.T.S., these general exceptions can be invoked to: (a) protect public morals or maintain public orders, (b) protect human, animal or plant life or health, (c) secure compliance with laws or regulations which are not inconsistent with the provisions of G.A.T.S.,[85] (d) collection or imposition of direct taxes, and (e) differential treatment due to the agreement on avoidance of double taxation. Article XIV(c)(ii) G.A.T.S. is crucial here, specifically in the case of G.D.P.R., as this particular exception can be invoked if it is necessary to secure compliance with such laws or regulations that exist for "*the protection of the privacy of individuals in relation to the processing and dissemination of personal data and the protection of confidentiality of individual records and accounts.*"[86] Article XIV of G.A.T.S., much like Article XX of G.A.T.T., consists of several tests to check the legality of the measure. The authors believe that G.D.P.R. can be justified under the general exceptions only if it falls either under Article XIV(a) or Article XIV(c)(ii), but under these exceptions, the members are still required to prove the "necessity", which is based on "weighing and balancing" test involving the importance of the values being protected, the extent of the measure's contribution, and trade restrictiveness of the measure.[87] Apart from this, if any Member wants to claim a general exception, they would also have to fulfil the conditions in the Chapeau which requires that the *application of the measure must* not constitute "arbitrary" or "unjustifiable" discrimination where the same conditions prevail, and a "disguised restriction on trade in services". It ensures that the Members' right to avail exceptions is exercised in a reasonable manner and does not frustrate the rights accorded to other Members under the substantive rules of the G.A.T.S. .[88]

In the *U.S. – Gambling* dispute, the U.S. invoked the "public morals" exception under the Wire Act by claiming that the measures taken are necessary to prevent[89] and protect the nation's public morals and further maintain public order.[90] The pre-requisite for the justification of invoking a general exception clause is that it should be done strictly in a situation "*. . . where a genuine and sufficiently serious threat is posed to one of the fundamental interests of the society*".[91]

---

[85] *See* G.A.T.S., *supra* note 36.

[86] *Id.* Art. XIV(c)(ii); For a fully-fledged analysis of how this may occur, *see generally* Kristina Irion, Svetlana Yakovleva & Marija Bartl, Trade and Privacy: Complicated Bedfellows? How to Achieve Data Protection-Proof Free Trade Agreements 27-33 (2016). *see generally* Weber, *supra* note 5, at 32-34

[87] *See* Appellate Body Report, *US – Gambling*, *supra* note 70, § 306.

[88] *See also id.* § 338-369.

[89] *See* The measure at issue was "the total prohibition on the cross- border supply of gambling and betting services". *See* Appellate Body Report, *US – Gambling*, *supra* note 70, § 10-11.

[90] *See id.* § 316.

[91] Yves Poullet, *Transborder Data Flows and Extraterritoriality: The European Position*, 2 J. Int'l Com. L. & Tech. 141 (2007).

"Public order" points towards the need for safeguarding fundamental interests of the nation[92] which are highly subjective and dependent on many peculiar factors, giving Members a free hand to determine the appropriate level of protection for themselves.[93] The term "order" is read with footnote 5 of G.A.T.S., and it can also be termed as the preservation of the fundamental interests of society.[94] The U.S., in the gambling dispute, successfully established the prima facie case of "necessity" within the meaning of Article XIV(a) by showing that there were no other reasonably available measures.[95] Nevertheless, they failed the chapeau test because the measure did not apply equally on the domestic and the foreign service suppliers.[96] The authors believe that the restrictions under the G.D.P.R. can also rely on the same public morals and public order exception as these restrictions are currently being promoted as a means to enhance data security, i.e. protecting the privacy and security of personal information,[97] which is one of the fundamental interests of society.

Meanwhile, the exception under Article XIV(c)(ii) remains highly relevant for any measure seeking to secure protection of privacy of individuals under the G.A.T.S. .[98] It remains imperative that the three-fold test is fully satisfied before any Member relies on the said defence,[99] which involves (i) the identification of laws and regulations with which the measure is expected to secure compliance; (ii) such laws and regulations must not be inconsistent with the W.T.O. norms; and (iii) the measure is imposed to ensure compliance with those laws and regulations.[100] Further, a Member can impose any measure consistent with W.T.O. norms to achieve "zero-risk" level of protection.[101] Once the three-fold test is satisfied, the Member would still have to show "necessity" and pass other criteria under the Chapeau. One such essential requirement under the Chapeau is that of the "like condition" comparison. Since the third party requirements under the G.D.P.R. are based on the fact that the data protection regime, including the national security system, in the E.U. and the third country is not at the same level[102] which is a

[92] *See* Panel Report, *European Union and its member States – Certain Measures Relating to the Energy Sector*, ¶ 296, WTO Doc. WT/DS476/R/Add.1 (circulated Aug. 10, 2018) [adoption/appeal pending].
[93] *See generally* Appellate Body Report, *US – Gambling*, *supra* note 70, §6.461.
[94] *See* Appellate Body Report, *Korea – Measures Affecting Imports of Fresh, Chilled and Frozen Beef*, ¶176, WTO Doc. WT/DS161/AB/R, WT/DS169/AB/R (adopted Jan. 10, 2001) [hereinafter Appellate Body Report, *Korea – Beef*].
[95] *See e.g.*, Appellate Body Report, *US – Gambling*, *supra* note 70, § 326.
[96] *See e.g.*, *id.* § 372.
[97] *See also* Anupam Chander & Uyên P. Lê, *Data Nationalism*, 64 Emory L.J., 677, 718 (2015).
[98] *See also* G.A.T.S, *supra* note 36.
[99] *See* Panel Report, *United States — Measures Relating to Shrimp from Thailand*, ¶7.7174, WTO Doc. WT/DS343/R (adopted Feb. 29, 2008).
[100] *Id.*
[101] *See e.g.*, Appellate Body Report, *Korea-Beef*, *supra* note 94, §181.
[102] *See* Peter Swire, *Foreign Intelligence and Other Issues in the Initial Opinion in Schrems II*, Lawfare (Dec.23, 2019), https://www.lawfareblog.com/foreign-intelligence-and-other-issues-initial-opinion-schrems-ii.

vital consideration under the G.D.P.R. .[103] Therefore, it is established that the conditions prevailing in these two countries are not alike, and the arbitrariness and justifiability of the measures shall not be challenged. Therefore, there is a strong possibility that a complaint in the near future may be brought before the D.S.B. against the E.U., pertaining to the high regulatory standards imposed by the G.D.P.R., which go against the E.U.'s schedule of commitments under G.A.T.S. . Nevertheless, the E.U. can continue to rely on the justification of its measures under the general exceptions provided by Article XIV of G.A.T.S. .[104]

3. DIGITAL TRADE POST-TRANSITION PERIOD: EXPLORING THE VIABLE MODELS FOR THE U.K.

Once the U.K. becomes a "third country" from the context of G.D.P.R., that is where the things change. In order to obtain an adequate data protection level, the U.K. will have to, first, seek approval from the European Commission, like other third countries. Apart from opting for the adequacy ruling, there are other options available that the U.K. can rely upon to regulate its future digital trade with the E.U. . So far, it was seen that the U.K. was inclined towards opting for the adequacy status approach; however, recently, the U.K. has also planned to negotiate a comprehensive free trade agreement with the E.U. .

3.1 OPTION THAT THE U.K. COULD HAVE EXPLORED

3.1.1 U.K. AS AN E.E.A. MEMBER: EXEMPTION FROM DATA ADEQUACY REQUIREMENT, BUT A LOSS OF SOVEREIGNTY

A few scholars had attempted to explore the E.E.A. Agreement model for the post-Brexit U.K., not essentially considering it from a data transfer aspect. It was noticed that since E.U. laws are directly applicable to E.E.A. Members without specific incorporation by their members,[105] and G.D.P.R. already applies to all the E.E.A. Members,[106] if the U.K.

[103] *See generally Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions, A European strategy for data,* COM(2020) 66 final (Feb. 19, 2020) [hereinafter *European Strategy for Data*].

[104] *See* Federica Velli, *The Issue of Data Protection in EU Trade Commitments: Cross-Border Data Transfers in GATS and Bilateral Free Trade Agreements*, 4 Eur. Papers 881 (2019).

[105] *See* Consolidated Version of the Treaty on the Functioning of the European Union art. 15, May 9, 2008, 2008 O.J. (C 115) 47 [hereinafter T.F.E.U.].

[106] *See generally* Decision of the EEA Joint Committee No 154/2018 of 6 July 2018 amending Annex XI (Electronic communication, audiovisual services and information society) and Protocol 37 (containing the list provided for in Article 101) to the EEA Agreement, 2018 O.J. (L 183).

joins E.E.A., then the U.K. is not required to be treated like third countries in the post-transition period. In the context of Britain,[107] the most important issue is whether the U.K. would remain a member of the E.E.A. after its withdrawal from the E.U. .

The E.E.A. Agreement aims to build economic cooperation by promoting cross-border trade between the E.U. and the E.E.A. member States (i.e. Iceland, Liechtenstein and Norway), with equal competitive opportunities.[108] Although the E.E.A. members are not part of the E.U., they still get to enjoy cooperation with the E.U. member states in the field of economy, security and society.[109] Meanwhile, it is important to note that the automatic application of the E.U. legislation on the E.E.A. Members is bound to raise concerns by being equivalent to challenging a nation's sovereign powers as the E.E.A. members are not given the opportunity to participate in the discussions of all the E.U. legislations, but instead are required to incorporate it.[110]

One group of scholars opined that the U.K.'s withdrawal from the E.U. will imply its withdrawal from the E.E.A. too, because the E.E.A. Agreement applies to the territories to which the E.U.'s treaty apply.[111] The other group of scholars argue that the U.K.'s termination of the E.U. membership does not affect its membership at the E.E.A.[112] until and unless it gives a 12 months' notice under Article 127 of the E.E.A. Agreement[113] as the existence of the U.K. in the E.E.A. is based on its autonomous status, rather than based on its membership with the E.U. .[114] Nevertheless, exploring an E.E.A. member would have had many disadvantages for the U.K. in the areas other than the digital trade, some of them being the common reasons behind the Brexit vote.

Seeking data transfer through the E.E.A. mode would mean a transfer of power from the U.K. to a central E.U. agency, and direct application of E.U. law in the U.K. without participation in the enactment of the legislation(s),[115] which is an obvious threat to the

[107] *See also* Lord Owen, *The Legal Way to Exit the EU by 31st October is Through a Transition in the E.E.A.*, Brexit Central (Oct. 8, 2019), https://brexitcentral.com/the-legal-way-to-exit-the-eu-by-31st-october-is-through-a-transition-in-the-E.E.A./.

[108] *See* Agreement on the European Economic Area, art. 2(c), art. 10, Jan. 3, 1994, 1994 O.J. (L 001) 3 [hereinafter E.E.A. Agreement].

[109] *See Utenfor Og Innenfor*: *Norges Avtaler*. *Med EU* (*Outside and Inside:* Norway's *Agreements with the* EU) NOU 64-76 (Ministry of Foreign Affairs, 2012).

[110] *See* John Erik Fossum, *Representation Under Hegemony? On Norway's Relationship to the EU, in* The European Union's Non-Members: Independence Under Hegemony? 153, 155-56 (Erik O. Eriksen & John Erik Fossum eds., 2015).

[111] *See* Dóra Sif Tynes & Elisabeth Lian Haugsdal, *In, Out or In-Between? The U.K. as a Contracting Party to the Agreement on the European Economic Area*, 41 Eur. L.Rev. 753, 763 (2016).

[112] *See* Ulrich G. Schroeter & Heinrich Nemeczek, *The (Uncertain) Impact of Brexit on the United Kingdom's Membership in the European Economic Area*, 27 Eur. Bus. L. Rev. 921, 922 (2016).

[113] *See* David Allen Green, *Brexit: Can and Should the U.K. Remain in the EEA?*, Fin. Times (Aug. 2, 2017), https://www.ft.com/content/16b50be8-161c-38d3-83b8-14b04faa9580.

[114] *See* Schroeter & Nemeczek, *supra* note 112, at 921.

[115] *See* Desmond Dinan, Europe Recast: A History of European Union 19-25 (Red Globe Press 2014) (2004).

sovereignty of the U.K. .[116] Secondly, the applicability of the burdensome E.U. regulations has cost the UK approximately $880 million,[117] and similarly, the cost of joining the E.E.A. would also be high. Therefore, joining (remaining in) the E.E.A. would not be in the best interests of a post-Brexit Britain[118] merely because it may continue to benefit the U.K. in cases of data privacy due to G.D.P.R. .

3.2 MODELS THAT ARE BEING EXPLORED

3.2.1 U.K.-E.U. C.F.T.A.: A BILATERAL APPROACH TO (DIGITAL) TRADE

A draft U.K.-E.U. Comprehensive Free Trade Agreement [hereinafter U.K.-E.U. C.F.T.A.], that came out in May 2020, seems to be the U.K.'s new plan post the transition period. While parties are currently negotiating on the same; they continue to have major differences on matters concerning level playing field provisions, competition, and state aid, and the overall governance structure of the agreement.[119] This approach of the U.K. is a tad bit similar to the Switzerland's approach in establishing its relationship with the E.U. where they closely integrated, and continue to cooperate on bilateral levels on a sectoral basis,[120] having concluded multiple agreements.[121] However, U.K. C.F.T.A.'s approach is essentially for seeking greater autonomy for the U.K.,[122] so that the U.K. will retain its sovereign right; and "whatever happens", the U.K. will not negotiate on any matters in which it loses its control over its laws and politics.[123] In furtherance of the same, the U.K. has also maintained its position on not being bound by the decision(s) of

[116] *See generally* Ralph C. Bryant, *Brexit: Make Hard Choices but Don't Confuse Sovereignty with Autonomy*, Brookings (Dec. 21, 2018), https://www.brookings.edu/blog/up-front/2018/12/21/brexit-make-hard-choices-but-dont-confuse-sovereignty-with-autonomy/.

[117] *See* Timothy B. Lee, *Brexit: the 7 most important arguments for Britain to leave the EU*, Vox (Jun. 25, 2016), https://www.vox.com/2016/6/22/11992106/brexit-arguments.

[118] *See* MJ Pérez Crespo, *After Brexit... The Best of Both Worlds? Rebutting the Norwegian and Swiss Models as Long-Term Options for the U.K.* 36(1) Yearbook of European Law 94, 108 (2017).

[119] *What does the U.K.'s draft EU FTA text tell us about the negotiations?*, Wiredgov (May 21, 2020),https://www.wired-gov.net/wg/news.nsf/articles/What+does+the+UKs+draft+EU+FTA+text+tell+us+about+the+negotiations+21052020112500.

[120] *Id.* at 11.

[121] *See e.g.*, Bilateral Agreement I EU- Swis, Jun. 21, 1999, 2002 O.J. (L 114). The areas covered by the Bilateral Agreement I were (i) free movement of persons; (ii) overland transport; (iii) air transport; (iv) public procurement markets; (v) participation in EU research program; (vi) agriculture; (vii) technical barriers to trade; Bilateral Agreement II EU- Swis, Oct. 26, 2004, 2004 O.J. (L 114).The areas covered by this Bilateral Agreement II were (i) taxation of savings, (ii) fight against fraud; (iii) Schengen/Dublin; (iv) processed agricultural products; (v) statistics; (vi) pension; (vii) environment; (viii) audio-visual industry; (ix) education and occupational training.

[122] *See* Marius Vahl & Nina Grolimund, Integration Without Membership: Switzerland's Bilateral Agreements with the European Union 3 (2006).

[123] *See* HM Government, *The Future Relationship with the EU-The UK's Approach to Negotiations*, CP211, ¶ 5 (Feb. 2020) [hereinafter Negotiations Document].

the European Court of Justice [hereinafter E.C.J.].[124] Taking into consideration the bilateral nature of the agreement, the U.K. remains to be at the same bargaining position as that of the E.U. .[125]

This model may pose certain structural challenges like (i) negotiating such a big agreement with multiple arrangements on different subjects will be a lengthy and complex process, but would invariably allow the U.K. to retain its sovereignty; (ii) the U.K. would be less bothered by implementation of all the E.U. laws, and would only have to comply with those laws which would be agreed upon in the bilateral agreement between them; and (iii) the U.K. would not be bound by an interpretation of the E.C.J. in cases of dispute between its bilateral agreement with the E.U., as it is in the case of Switzerland.

In context of the data transfer, the U.K. has declared that data protection will not be a part of U.K.-E.U. C.F.T.A. as the U.K. considers it separate from a more comprehensive future relationship,[126] yet it is crucial to note that the U.K.-E.U. C.F.T.A. does provide a separate chapter on digital trade.[127] Also, the draft U.K.-E.U. C.F.T.A. specifically talks about emerging technology which is inextricably related to the G.D.P.R. as the G.D.P.R. has been "designed to cover new technologies".[128]

The importance of digital trade and adequacy decision has been noted by the European Commission, especially in the context of Brexit, by stating that Brexit is "... an enabling factor for trade, including digital trade, and an essential prerequisite for a close and ambitious cooperation in the area of law enforcement and security".[129] Noting the importance of data protection in digital trade,[130] which has also been considered by the E.U., the authors feel that data adequacy should be a part of the U.K.-E.U. C.F.T.A., as the exclusion of data privacy laws would have implications on areas such as banking and financial services, e-commerce, etc. which are part of the U.K.-E.U. C.F.T.A..[131] However, Europe's data strategy clearly stipulates that the *only* way for data protection,

[124] *Id.* ¶ 6.

[125] *See* René Schwok & Cenni Najy, UK Returning to EFTA: Divorce at 40 and Going Back to Mom and Dad? ¶ 40 (2012).

[126] *See* Negotiations Document, *supra* note 123, ¶ 60.

[127] *See generally* Draft UK-EU Comprehensive Free Trade Agreement, March 18, 2020, UKTF (2020) 14 [hereinafter C.F.T.A.].

[128] *See* Draft UK-EU Comprehensive Free Trade Agreement, March 18, 2020, UKTF (2020) 14 [hereinafter C.F.T.A.]; *Communication from the Commission to the European Parliament and the Council, Data protection as a pillar of citizens' empowerment and the EU's approach to the digital transition - two years of application of the General Data Protection Regulation*, COM (2020) 264 final (Jun. 24, 2020) [hereinafter E.U.'s Approach].

[129] E.U.'s Approach, *supra* note 128, at 11.

[130] *See generally* Aaditya Mattoo & Joshua P. Meltzer, *Resolving the conflict between Privacy and Digital Trade*, Vox (May 23, 2018), https://voxeu.org/article/resolving-conflict-between-privacy-and-digital-trade.

[131] *See* For instance, Article 18.13 specifically talks about personal information protection and hence exclusion of adequacy status issue from the EU-UK FTA would cause additional burden to the UK.

irrespective of its importance in digital trade, is adequacy decision or other existing mechanisms in the G.D.P.R., rather than subjecting data protection under the free trade agreement,[132] and the process for obtaining the adequacy decision has been initiated by the U.K.,[133] however the E.U. has casted doubts in this regard.[134] It is crucial to note that while adequacy decision are yet to be finalised, data transfer between the U.K. and the E.U. is currently taking place under the Trade and Cooperation Agreement signed on 24 December 2020, which has provided a six-month bridging period to allow the continued flow of data.

3.2.2 FULL ADEQUACY FINDING UNDER THE G.D.P.R.: A HARD NUT TO CRACK

The transboundary transfer of data from one W.T.O. Member to another is analogous to the free flow of trade in (electronic) services. In situations involving the transfer of data from the E.U. to a third country, it is essential to consider whether that third country is a "secured" country (i.e. the country that has been afforded adequacy status) or an "unsecured" country under the G.D.P.R. .[135] Article 45 of the G.D.P.R.[136] provides that secured countries are the ones where the national laws provide an essentially equivalent level of protection to the data as that in the E.U., and have in that way successfully achieved adequacy decision. The E.C.J. gave a landmark decision in the *Schrems v Data Protection Commissioner* [hereinafter Schrems I] case concerning Article 25 of the E.U. Data Protection Directive, which is broader than Article 45 of the G.D.P.R., wherein the Court held that there must be an evaluation of the third country's data protection laws and its application.[137] In other words, *Schrems I* case advocated the essential equivalency test. In addition to these requirements, Article 45 should also evaluate whether third country's laws on national security, human rights, etc. are similar to the E.U. laws,[138] and while determining such adequacy of the data protection law, the European Commission must take into account the existence and effective functioning of one or more independent

[132] *See* European Strategy for Data, *supra* note 103.
[133] *See, e.g. What does Adequacy mean?*, ICO https://ico.org.uk/for-organisations/data-protection-at-the-end-of-the-transition-period/what-does-adequacy-mean.
[134] *See* Letter from Andrea Jelinek, Chair of the European Data Protection Board (EDPB) to Member States, (Jun. 15, 2020), https://edpb.europa.eu/sites/edpb/files/files/file1/edpb_letter_out_2020-0054-uk-usagreement.pdf.
[135] *See G.D.P.R., Third Countries*, Intersoft Consulting, https://gdpr-info.eu/issues/third-countries.
[136] *See* G.D.P.R., *supra* note 13, Art. 45. Article 45 of the G.D.P.R. imposes an obligation on the member states to ensure that the data can be transferred to a country outside the EU only if that country ensures an adequate level of protection.
[137] *See* Case C-362/14, Max Schrems v. Data Protection Commissioner, ECLI:EU:C:2015:650, ¶ 75 (Oct. 6, 2015).
[138] *See, e.g.*, Prashant Mali, GDPR Articles with Commentary & EU Case Laws 109 (2019).

supervisory authorities, the rule of law, relevant legislation, respect for human rights and fundamental freedoms, and the international commitments the third country or international organisation has entered into.[139]

Meanwhile, unsecured countries are considered as the ones which have not achieved an 'adequacy decision', and in the absence of that adequacy decision, the international transfer of data can take place through B.C.R.s,[140] S.C.C.s,[141] or as per the procedures under Article 46 or Article 49 of the G.D.P.R. . The data transfer that happens through S.C.C. between private parties can also ensure data protection if assurances are made to comply with the G.D.P.R.'s approved Code of Conduct.[142] These model Code of Conduct must be approved[143] and monitored[144] by a supervisory authority under the G.D.P.R.; and it must be noted that these corporate rules or standard data protection clauses are available only for arrangements among the parties without any direct intervention by the State as a regulatory authority, while an adequacy status is given to a State without the involvement of private parties. The comparison between the B.C.R.s and S.C.C.s suggests that B.C.R.s might be overall more effective than the S.C.C.s as the former can be tailored as per the requirements of a corporation which will differ with the kind of business operations being carried out. B.C.R.s can be comparatively costly as they require an approval from a competent data protection authority [hereinafter D.P.A.], and then the decision of the D.P.A. is sent to the European Data Protection Board [hereinafter E.D.P.B.], and it is only after the opinion of E.D.P.B., the D.P.A. approves the concerned B.C.R.s. However, in the case of a S.C.C., *firstly*, the S.CC.. is approved by the Commission which can be used by the parties without any approval from any authority - making it less costly than the B.C.R.s; however, in case of an ad-hoc S.C.C., there is a similar approval mechanism as in the case of B.C.R.s.[145] Also, the scope of B.C.R.s is limited to only intra-corporate transfer as opposed to the S.C.C.s, which have a wider scope. It is also observed that there exist possibilities when compliance with "adequacy decision" is not needed in certain circumstances, viz. (i) if the data subjects have consented to it; (ii) the transfer pertains to a legal or a contractual claim; or (iii) in the public interest.[146]

[139] *See* G.D.P.R., *supra* note 13, Art. 45.

[140] Binding Corporate Rules are the rules dealing with international data transfer within intra-organizations located in different countries. *See generally International personal data transfers: binding corporate rules (BCRs) under the GDPR*, I-Scoop, https://www.i-scoop.eu/gdpr/binding-corporate-rules-bcrs-gdpr.

[141] *See generally* G.D.P.R., *supra note* 13, Art. 46.

[142] *See id.* Art. 40; *see also GDPR and approved codes of conduct – demonstrating compliance*, I-Scoop, https://www.i-scoop.eu/gdpr/gdpr-codes-conduct.

[143] G.D.P.R., *supra* note 13, Art. 55.

[144] *Id.* Art. 41.

[145] *See, e.g.*, Patrick Van Eecke et al., *EU: Binding Corporate Rules are Generating Greater Interest*, DLA Piper (Oct. 16, 2019), https://blogs.dlapiper.com/privacymatters/eu-binding-corporate-rules-are-generating-greater-interest.

[146] G.D.P.R., *supra* note 13, Art. 49.

The authors note that the option of moving forward with an "adequacy decision" approach seems to be the easiest amongst all the others discussed above,[147] and is being pursued by the U.K. to retain its sovereignty. It has also been argued that seeking an "adequacy decision" in the case of the U.K. is merely a formal requirement as U.K. was already a member of the E.U.,[148] and they would be better off with a fully-fledged comprehensive E.U.-U.K. data agreement.[149] The advantage of this bilateral data *only* arrangement approach would be that it cannot be revoked by the European Commission or invalidated by the E.C.J. as would be in the case of an "adequacy decision",[150] and it makes this bilateral data only arrangement more beneficial since we understand that a sudden revocation or invalidation of an adequacy decision would freeze the E.U. – U.K. digital trade, eventually hampering the U.K.'s economy.

Furthermore, the U.K. Information Commission Office [hereinafter I.C.O.][151] has already implemented the G.D.P.R., with the promulgation of the Data Protection Act 2018.[152] Nevertheless, despite the compatibility that exists between the U.K. Data Protection Act and E.U.'s G.D.P.R., there exists a doubt on specific grounds as to whether the UK will be able to achieve its adequacy decision without any problems. So, in *Secretary of State for the Home Department v Watson*, the E.C.J. held that the collection of data and its retention by the security services in all cases without having any exception, even for the collection and/or retention of data for combating crime,[153] violates the fundamental right to privacy of a person.[154] A possible hindrance in accomplishing the essential equivalence test by the U.K. exists concerning their "Investigatory Powers Act, 2016". The High Court and the European Court of Human Rights have held that the powers granted to the U.K.'s security forces and intelligence services to intercept, hold, and examine any data would end up violating the privacy rights of individuals, and hence put the likelihood of getting an adequacy decision by the E.U. into question.[155]

---

[147] *See* Warwick Ashford, *UK surveillance laws a potential 'sticking point' post-Brexit*, COMPUTERWEEKLY.COM (Apr. 19, 2018), https://www.computerweekly.com/news/252439434/UK-surveillance-laws-a-potential-sticking-point-post-Brexit.

[148] *See* Department of Exiting the EU, *Framework for the UK-EU Partnership Data Protection* (2018), 16.

[149] *See* HM Government, *Technical Note: Benefits of a new data protection agreement* (2018), 1; *See also* Ashford, *supra* note 147.

[150] *See* OLIVER PATEL & NATHAN LEA, EU-UK DATA FLOWS, BREXIT AND NO-DEAL: ADEQUACY OR DISARRAY? 9 (2019). *see also* G.D.P.R., supra note 13, Recital 107, Art. 45(5).

[151] The UK's independent authority set up to uphold information rights in the public interest, promoting openness by public bodies and data privacy for individuals. *See* ICO. https://ico.org.uk/.

[152] *See generally Guide to the General Data Protection Regulation (GDPR)*, ICO, https://ico.org.uk/for-organisations/guide-to-data-protection/guide-to-the-general-data-protection-regulation-gdpr

[153] *See Secretary of the State for the Home Department v. Watson: Case Analysis*, GLOBALFREEDOMOFEXPRESSION.COLUMBIA.EDU, https://globalfreedomofexpression.columbia.edu/cases/secretary-state-home-department-v-watson.

[154] *See* C-698/15, Secretary of State for the Home Department v. Watson, 2016 ECLI: EU:C:2016:970, ¶ 103 (Dec.21, 2016).

[155] William R.M. Long & Francesca Blythe, *Data Protection and Post-Brexit Issues*, INT'L J. DATA PROT. OFFICER, PRIV. OFFICER & PRIV. COUNS., no. 2, 2018, at 8.

Similarly, the "Five Eyes Alliance", an intelligence alliance consisting of the U.S., Canada, Australia, New Zealand and the U.K., dealing with cooperation on signal intelligence sharing, proposed that the law enforcement agencies should have access to encrypted data to make them stronger.[156] The concerns of the European Commission herein also remain the same, i.e. this exercise might lead to an unchecked transboundary transfer of personal data of the E.U. citizens under the garb of national security.[157] Therefore, such issues will act as a major hindrance in the achievement of an adequacy status, even though the laws are entirely consistent with G.D.P.R. .

CONCLUSION

While there are several provisions in G.A.T.S which are helpful in the regulation of data through various mediums, it is well documented that none of those provisions is conducive for protection of the same data, even when the transfer of such data is happening through electronic services.[158] The rationale behind the non-existence of such standards can be attributed to the fact that the protection of data deals with issues directly linked to privacy, and the right to privacy is more of a fundamental and constitutional right raising questions revolving around the sovereignty of a state. Therefore, it is necessary to look for solutions that can act as a bridge between the need for regulation of data under the W.T.O.-law and reasonable application of privacy laws for the protection of this data.[159]

Weber suggested that one of the most viable justifications for a W.T.O. member's need for applying stringent data protection laws can be read in reference to the general exception under Article XIV (lit. c) of the G.A.T.S.[160] Nonetheless, the level of the protection of data can always be increased in bilateral or regional trade agreements by having specific chapters looking into the concerns of members for the regulation and protection of their data.[161] However, as far as protection of data under G.D.P.R. is concerned, it has been argued by several scholars that if the G.D.P.R. is challenged before a W.T.O. panel, the E.U. would have to heavily rely on the exceptions under G.A.T.S,

[156] *See "Five Eyes" security alliance calls for access to encrypted material*, Reuters (Jul. 30, 2019), https://www.reuters.com/article/us-security-fiveeyes-britain-idUSKCN1UP199.
[157] *See* David Meyer, *Post-Brexit U.K.'s Surveillance Practices Could Spell Big Problems for Business,* Fortune (Feb. 7, 2020), https://fortune.com/2020/02/07/post-brexit-uk-data-protection-adequacy/.
[158] *See* Weber, *supra* note 5, at 25.
[159] *Id.*
[160] *Id.*
[161] *Id.*

creating a difficult situation for them considering the fact that only in one case the W.T.O. D.S.B. has allowed the innovation of such exception.[162]

Now, after the transition period, the Commission has not given its adequacy decision as of now, and there are possibilities that the data protection regimes of the U.K. and E.U. would remain consistent;[163] nonetheless, the authors are of the opinion that all the legal entities must start preparing themselves with an alternative method, having included such B.C.R.s or S.C.C.s, which may help them deal with the international transfer of data; otherwise, the entire economic activity surrounding the banking, finance, and e-commerce sectors will be halted. Furthermore, post-transition period, the authority of the Information Commission Office as the D.P.A. will cease under the G.D.P.R. . Hence, business organizations need to approach the E.U. Data Protection Authority for any kind of approval, which will put more burden on them.

Lastly, the authors also are of the opinion that the option of the U.K.-E.U. C.F.T.A. with proper incorporation of data privacy arrangements under it would be the best option for the U.K. moving forward, since this approach allows the U.K. to retain its autonomy and sovereignty, whilst also managing to get an adequate level of protection in the eyes of G.D.P.R. for areas/sectors where the lack of an adequacy decision, immediately after the end of transition period, might have significant implications. Therefore, it seems fit to say that whatever path the U.K. may choose for its relationship with the E.U. pertaining to the concerns surrounding the international transfer of data, it must start finalising the guidelines around the same, as otherwise the economic activity of the business organisations will be endangered due to the hindrances in the transfer of data.

[162] *See* Appellate Body Report, *European Communities – Measures Affecting Asbestos and Asbestos-Containing Products*, ¶3243, WTO Doc. WT/DS135/AB/R (adopted 5 April 2001).
[163] *See generally* Wimmer & Jones, *supra* note 12, at 1561.

VOLUME 6 ISSUE 1 ARTICLES & ESSAYS

UNIVERSITY OF BOLOGNA LAW REVIEW

https://doi.org/10.6092/issn.2531-6133/12881

Received: 26 Sep. 2020 | Accepted: 30 Nov. 2020 | Published: 10 May 2021

Challenging the Impact of FIOST Clauses on Cargo Interests

AHMET GELGEC

Ahmet Gelgec is a Lecturer in Maritime and Commercial Law at Istanbul Medeniyet University (Turkey).

@ ahmet.gelgec@medeniyet.edu.tr

0000-0002-6075-7211

ABSTRACT

Loss of, or damage to goods is a frequent occurrence in the shipping industry, which may often occur as a result of improper cargo-handling operations during loading, discharging or even stowing. This highly concerns cargo interests, as they will seek to reimburse their loss from their carriers under bills of lading. Often, the bill of lading may well contain terms of a charterparty by way of incorporation that allow the carrier to contract out their cargo-related operations. Once this is the case, the cargo interest is unjustly left without a remedy for loss of, or damage to his goods vis-à-vis the carrier under English law. This paper, instead of challenging the correctness of the law firmly established concerning the transfer of these obligations via Free In and Out Stowed and Trimmed (FIOST) clauses, rather, aims to propose ideas to tackle the impact arising out of the status quo under English law. Finally, it offers some plausible suggestions for cargo interests to surmount this undesired outcome.

KEYWORDS

Fiost clauses; The Hague Rules; Cargo Interests; Shipowners; Charterers

TABLE OF CONTENTS

INTRODUCTION

The International Convention for the Unification of Certain Rules of Law relating to Bills of Lading, and Protocol of Signature [hereinafter Hague Rules] or the Hague-Visby Rules[1] as amended by the Protocol to Amend the International Convention for the Unification of Certain Rules of Law Relating to Bills of Lading signed at Brussels in 1968 apply either by way of paramount clause agreed in or compulsorily, by force of law to the contracts contained or evidenced in bills of lading. The Rules prescribe the rights and liabilities arising out of the desired voyage starting from loading to discharge operations, including the core obligations of the carrier, like providing a seaworthy vessel[2] and other responsibilities caring for the cargo such as loading, handling, discharging and stowage operations.[3] Hence, when failing to deliver one of these responsibilities, the carrier may become subject to liabilities against cargo interests under the bill of lading.

However, sometimes the shipowner may not be willing to undertake some of these obligations that would normally fall upon him as defined under the Rules. Accordingly, as a result of freedom of contract, when drafting charter party terms, by virtue of free-in and free-out clauses, these responsibilities may have been shifted to charterers who often are not the bill of lading holders. Sometimes even a term may have been inserted into the bill of lading transferring these responsibilities to cargo interests, like shippers or receivers. And it is not uncommon in practice that some germane terms of the charterparty may have as well been incorporated into the bill of lading by general words of incorporation. When this is the case, under English law, as it will be shown below, cargo interests' claims *vis-à-vis* carriers arising out of breach of these obligations appear to be falling short of success, as they are allowed to contract out them. It is, however, submitted that this undermines the balance that is struck out by the Rules between cargo interests and carriers, making the former weaker and the latter even stronger, as the former are left without a remedy, whereas the faulty party escapes liability.

It is worth noting that by no means, does this paper seek to challenge the correctness of the law firmly established concerning the transfer of these obligations via Free In and Out Stowed and Trimmed [hereinafter FIOST] clauses, but rather aims to

[1] The Hague Rules of 1924 are not in force in the UK. The Hague-Visby Rules as amended by the Brussels Protocol 1968 are in force in the UK by the Carriage of Goods by Sea Act 1971. Unless otherwise stated, the arguments are to cover both the Hague and the Hague-Visby Rules, as the relevant provisions on the subject of this paper bear no difference under both the Hague and the Hague-Visby Rules.

[2] Art.III(1) of the Rules prescribes an obligation that requires the carrier to exercise due diligence before or at the beginning of the voyage to make the vessel seaworthy to undertake the intended voyage. "Seaworthiness" covers all aspects that put the vessel in a condition to perform properly the contractual voyage. For detailed explanations on seaworthiness, see *The Eurasian Dream* [2002] EWHC 118; [2002] 1 Lloyd's Rep 719, 735.

[3] Art.III(2).

propose ideas to tackle the undesired outcome arising out of the *status quo*. To do so, first, it shall set out in brief, the purpose of FIOST clauses and its impact between carriers and cargo interests under the bill of lading contract. Secondly, and more importantly, it shall seek to surmount the undesired outcome for cargo interests by offering plausible suggestions.

1. FIOST CLAUSES AND THEIR INCORPORATION INTO BILLS OF LADING

Cargo-related obligations such as loading, stowing and discharging, in the absence of an express provision, fall upon the shipowner rather than the charterer in common law. Though, this is hardly the case in practice, in most cases, parties are not silent on this matter in their contracts. Under charter parties, the shipowner and the charterer are free to draft or negotiate any term in allocating their respective responsibilities among themselves, such as these cargo-related operations for the manner in which they are to be carried out.

Free In and Out Stowed [hereinafter FIOS] or FIOST clauses or other variations are often drafted in charterparties.[4] Depending on the wording used, these clauses might have different impact on the rights and responsibilities of the parties. In an unqualified form of FIOS or FIOST clause, "free" only indicates that it is free of cost.[5] This is to say, such a clause is not capable of transferring the risk of these operations, and therefore it does not discharge the shipowner from responsibility against cargo interests.[6] Although, an effectively drafted clause may transfer the risk of these operations, making the shipowner relieved of responsibility.[7] Whereas the latter type enables the shipowner to contract out these obligations and hence the responsibility, the former type does not offer shipowners

[4] FIO (free in and out), FIOS (free in and out stowed), FIOST (free in and out stowed trimmed), FILO (free in liner out), FISLO (free in stowed liner out), etc. On FIOS or FIOST clauses in general see, Richard Aikens et al., Bills of Lading 338-9 (2nd ed. 2016) ;Bernard Eder et al., Scrutton on Charterparties and Bills of Lading 344-5 (23rd ed. 2015) ; 3h Paul Todd, Principles of the Carriage of Goods by Sea 261-2, 319-22 (1st ed. 2016) ;Julian Cooke et al., Voyage Charters 14.36-14.45 (4th ed. 2014) ; Simon Baughen, Shipping Law 106-9 (7th ed 2019). For more detailed analysis of the nature and scope of FIOST clauses see, Ilian Djadjev, The Obligation of the Carrier regarding the Cargo 101-47 (2017).

[5] *Jindal Iron and Steel Co Ltd v. Islamic Solidarity Shipping Co Jordan Inc (The Jordan II)* [2003] [2003] EWCA Civ 144; 2 Lloyd's Rep 87, 103. See also, *The Mohave Maiden* [2015] EWHC 1747 (Comm).

[6] For an example such a clause, see clause 8 of the NYPE, "...and Charterers are to load, stow, and trim the cargo at their expense under the supervision of the Captain..." *See* also *Sea Master Shipping Inc. v. Arab Bank (Switzerland) Ltd* [2020] EWHC 2030 (Comm).

[7] *See*, clause 5 of the GENCON charter party, "the cargo shall be brought into the holds, loaded, stowed and/or trimmed, tallied, lashed and/or secured by the Charterers, free of any risk, liability and expense whatsoever to the Owners".

a way out from responsibilities against cargo interests. It thus appears that there is no invariable general outcome from these clauses and accordingly cargo interests' right of recourse against carriers is highly dependent upon the wording of each clause.

Although FIOST clauses are sometimes inserted into bills of lading, they might often be incorporated from a charterparty. Use of general words of incorporation clause in the bill of lading is not uncommon, such as "all terms and conditions, liberties and exceptions of the charterparty . . . ".[8] The general rule is that as long as clauses are germane to the subject matter of the bill of lading (ie., carriage or delivery of the goods or the payment of freight and demurrage), those germane terms of the charterparty would be treated as incorporated into the bill of lading.[9] As for incorporation of FIOST clauses, in *the Eems Solar*, general words of incorporation were held to be successfully incorporated into a FIOST clause of a charterparty into the bill of lading.[10] Sometimes such incorporated clauses may require a degree of manipulation of the words so as to fit exactly the bill of lading. However, it must be said that the English courts have shown no reluctance in manipulating such incorporated clauses that are considered germane to fit them into bills of lading.[11] In supporting this argument, in *the Eems Solar,* manipulation was found to be unnecessary for a FIOST clause – which was assumed to be germane - relieving the shipowner of the cargo-handling obligations, where the charterers had them imposed on.[12] Though, this does not necessarily mean that an incorporated FIOST clause always sits well with the other terms of the bill of lading, which will be discussed below.

2. CONTENTION BETWEEN FIOST CLAUSES AND THE HAGUE RULES

As stated above, FIOST clauses can be successfully incorporated into bills of lading by the use of general words. If the incorporated clause did not only transfer of cost but also the risk of cargo-related operations to the shipper, the receiver or even to the charterer, a

[8] For examples of general words of incorporation clauses, see, *Skips A/S Nordheim v. Syrian Petroleum Co and Petrofina SA (The Varenna)* [1984] Q.B. 599; *Federal Bulk Carriers Inc v. C Itoh & Co Ltd (The Federal Bulker)* [1989] 1 Lloyd's Rep. 103.

[9] *TW Thomas & Co Ltd v. Portsea SS Co Ltd (The Portsmouth)* [1912] AC 1; *Yuzhny Zavod Metall Profil LLC v. Eems Beheerder BV (The Eems Solar)* [2013] 2 Lloyd's Rep 487.

[10] *Yuzhny Zavod Metall Profil LLC v. Eems Beheerder BV (The Eems Solar)* [2013] 2 Lloyd's Rep. 487, [94-95]. For a detailed commentary on this case see, Paul Todd, *Hague Rules and Stowage*, Lloyd's Mar.& Com. L Q. 139 (2014).

[11] Though where manipulation was not possible for a demurrage clause, see *Miramar Maritime Corp v. Holborn Oil Trading (The Miramar)* [1984] A.C. 676.

[12] *The Eems Solar* [2013] 2 Lloyd's Rep 487, [94-95].

tension would be deemed to arise between that clause and art.III(2) of the Rules in principle, as it imposes the obligations to load, stow and discharge onto the carrier[13] under the bill of lading. Therefore, a question arises at this point: would an incorporated FIOST clause lowering the carrier's cargo-handling responsibilities prevail over art.III(2)?

The Hague or the Hague-Visby Rules were created to strike a balance between the weak party (the cargo interest) and the party that would enjoy upper bargaining power (the carrier). In order to find a balance between them, the Rules therefore lay out standard liabilities of the carrier *vis-à-vis* the cargo interest. So as to protect cargo interests as the weaker party, thus art.III(8) of the Rules reads:

> Any clause, covenant, or agreement in a contract of carriage relieving the carrier or the ship from liability for loss or damage to, or in connection with, goods arising from negligence, fault, or failure in the duties and obligations provided in this article or lessening such liability otherwise than as provided in these Rules, shall be null and void and of no effect. A benefit of insurance in favour of the carrier or similar clause shall be deemed to be a clause relieving the carrier from liability.

The provision goes to the root of the underlying reason that caused the creation of the Rules; where the carrier seeks to lower his liability provided in the Rules by other contractual terms drafted or incorporated, these terms would be rendered void and of no effect hereunder. As for FIOST clauses, although it might seem plausible at first sight that they should be struck out by art.III(8), once the carrier seeks to relieve himself of responsibilities under art.III(2), such as for loss or damage to goods arising from improper stowage, the approach taken by the English courts may prove this wrong.

Devlin J in *Pyrene v Scindia* held that art.III(2) does not define the scope of the carriage contract, but the terms upon which the service is to be performed.[14] This is to say, art.III(2) by no means is to be construed to override freedom of contract that the parties may enjoy under the bill of lading, when reallocating the cargo-related operations. The proposition of Devlin J was later affirmed by the House of Lords in *Renton v Palmyra*.[15] Thus, the terms incorporated shifting these responsibilities to a third party are construed as the terms merely defining the scope of the contract, and therefore they appear to be not falling within the ambit of art.III(8). This was the case both in *the Jordan*

[13] It is assumed that the shipowner is the contracting carrier under the bill of lading (shipowner's bill), which is signed, by the master or other person as agent for the shipowner. If the contracting carrier is the charterer who is also the performing party of the cargo-related operations via FIOST clauses, the cargo interest will then be able to sue the charterer/contracting carrier for the breach of art.III(2) under the bill of lading. On identification of the carrier under the bill of lading see, Eder, *supra* note 4, at 115-118.

[14] *Pyrene Co Ltd v. Scindia Steam Navigation Co Ltd* [1954] 2 Q.B. 402, 418.

[15] *GH Renton & Co Ltd v. Palmyra Trading Corp of Panama (The Caspiana)* [1957] AC 149, 170, 173, 174 and 176.

II[16] and *the Eems Solar*.[17] In the former, the charterparty contained a FIOST clause and another clause reading "Shipper/Charterers/Receivers to put the cargo on board, trim and discharge cargo free of expense to the vessel . . . " and the bill of lading incorporated the terms of this charterparty. Following the discharge, the cargo was found to be damaged. The issue revolved around whether the shipowner was liable for damage occurred during loading, stowing or discharge. Whilst the charterers brought a claim against the shipowner under the charterparty incorporating the Hague Rules, the shipper and the consignee sued under the bills of lading. Having followed the *ratio decidendi* of *Renton v Palmyra* and *Pyrene v Scindia*, in the House of Lords held that the terms transferring cargo-related responsibilities such as loading, stowing or discharging from shipowners to shippers, consignees, charterers were not caught by art. III (8).[18]

In *the Eems Solar*, the shipowners were sued under the bill of lading, which incorporated terms of the charterparty including a FIOST clause transferring the cargo-related responsibilities from the shipowner to the charterer by the cargo owners for damage caused to a cargo of steel sheets by adverse weather. The cargo owners submitted that the shipowners were in breach of art.III(1) and (2), as the cargo was improperly stowed to protect it from damage during voyage. They further argued that the incorporated FIOST clause would be rendered null and void under art.III(8), as it would discharge the shipowner from responsibility under the Rules. First, it was found that the damage to the cargo was caused by improper stowage and not unseaworthiness. Though, having followed *Renton v Palymyra* and *the Jordan II*, the court eventually held that the shipowner was entitled to contract out the cargo-related operations, and equally the shipowner was relieved of responsibilities arising from these operations. It is therefore safe to assume that a FIOST clause may not be invalidated by art.III(8) under English law, accordingly it is considered consistent with the relevant provisions of the Rules.

3. THE APPROACH TAKEN BY OTHER JURISDICTIONS

Under English law, as the authorities cited above illustrate, cargo interests appear to be destined arguing without success that FIOST clauses are caught by art.III(8), as shipowners are entitled to freely negotiate their contract including reallocation of

[16] *The Jordan II* [2005] 1 W.L.R. 1363; [2005] 1 Lloyd's Rep. 57, [14].
[17] [2013] 2 Lloyd's Rep 487, [68], [100].
[18] *The Jordan II* [2005] 1 W.L.R. 1363; [2005] 1 Lloyd's Rep 57, [12].

cargo-related operations under bills of lading. Though the author opines that it is necessary to re-strike the balance between those parties, as the standard protection provided by the Rules for cargo interests to this extent are undermined by English case law. Scrutinising other jurisdictions might provide some assistance to see if there is any useful solution for this purpose.

To start with the Commonwealth countries, the approach taken by the English courts seems to have been followed by Australia,[19] New Zealand,[20] India[21] and Pakistan.[22] Italian law can also be said to show similarities to the English approach.[23] On the other hand, in South Africa, it was held that cargo-handling operations are non-delegable duties of the carrier and accordingly, cargo interests are not denied redress against carriers in that respect.[24] As for French law, the view adopted is that FIOST clauses are only capable of transferring the risk of cost for cargo-handling operations such as loading, trimming, discharging, stowing etc., meaning that the carrier will not be entitled to contract out the risk of these operations, and accordingly may not be relieved of liability arising from damage done to goods resulting from these operations.[25]

In respect of Dutch law, it is safe to contend that FIOST clauses are capable of transferring the risk and responsibilities of these operations along with the cost.[26] On the other hand, there is no predominant view adopted by the courts in the United States.

Whereas some circuits evidently favoured the view that unless the carrier supervises or takes no part in these operations, he might contract out of responsibility for these duties by FIOST clauses,[27] the US Court of Appeals for the Second and Fifth Circuits held that cargo-related operations were inalienable, and as a consequence a FIOST clause was rendered null and void under art.III(8) of the Rules.[28]

---

[19] *See Shipping Corporation of India v Gamlen Chemical Co A/Asia Pty Ltd* [1980] 147 CLR. 142 (Austl.); *Hunter Grain Pty Ltd v Hyundai Merchant Marine Co Ltd* [1993] 117 ALR 507 (Austl.); *Nikolay Malakhov Shipping Co Ltd v SEAS Sapfor Ltd* [1998] 44 NSWLR 371 (Austl.).

[20] *See BJ Ball New Zealand v. Federal Steam Navigation* [1950] NZLR 954 (N.Z.); *International Ore & Fertilizer Corporation v. East Coast Fertilizer Co Ltd* [1987] 1 NZLR 9 (N.Z.); *Dairy Containers v. The Ship (The Tasman Discoverer)* [2003] 3 NZLR 353 (N.Z.).

[21] *See The New India Assurance Co Ltd v. M/S Splosna Plovba* [1986] AIR Ker 176 (India).

[22] *See East and West Steamship Co v. Hossain Brothers* (1968) 20 PLD SC 15 (Pak.).

[23] *The Saudi Prince (No 2)* [1988] 1 Lloyd's Rep 1.

[24] *The Sea Joy* [1998] 1 SA 487 (S. Afr.).

[25] William Tetley, Marine Cargo Claims 1297-300 (Yvon Blais, 4th ed. 2008)

[26] *De Atlantic Coast* [1991] Nr 15 34. For a detailed analysis of the Dutch approach on the matter see, Djadjev, *supra* note 4, at 107-110.

[27] *See Sumitomo Corp of America v. M/V Sie Kim* (1985) 632 FSupp 824, (United States District Court); *Atlas Assurance Co v. Harper Robinson Shipping Co* (1975) 508 F2d 1381 (9th Circuit); *Sigri Carbon Corp v. Lykes Bros Steamship Co* (1987) 655 FSupp 1435, (United States District Court).

[28] *See Associated Metals Minerals* v. *MV Arktis Sky* (1992) 978 F2d 47 (2nd Circuit) and *Tubacex* v. *MV Risan* (1995) 45 F3d 951 (5th Circuit).

As for the propositions adopted by South African and French and, to some extent, by the US law, as much as they might seem favourable, they do not suggest any mechanism in reinstating the balance between cargo interests and shipowners under the Rules in this sense, as carriers are not allowed to contract out of responsibility at all in the first place.

As regards the other jurisdictions favouring the English approach where carriers may be relieved of liability pursuant to a FIOST clause, the author is not aware of any reported case that has sought to alter the position of cargo interests. That is to say, the laws of these jurisdictions appear to fall short of providing assistance in altering the impact of these clauses.

4. QUEST FOR RE-STRIKING THE BALANCE

By a FIOST clause inserted or incorporated into bills of lading, as is shown in the above sections, it would not be wrong to note that carriers are able to surpass any claim against cargo interests concerning loss arising during cargo-handling operations, which they have taken no part in. It is thought that this bears an undesired outcome that needs revision. Assume that a receiver took delivery of a cargo, which was damaged as a result of improper stowage. Also suppose that a shipper delivered his goods to the shipowner/carrier for shipment, which was also damaged due to bad stowage. Also suppose that in both cases, there was a FIOST clause of a charterparty incorporated into bills of lading, transferring responsibilities to a third party. First, in both scenarios, both the receiver and the shipper could bring a claim against their carriers/shipowners for their respective losses, as it is their carriers that they both had direct contractual proximity with under the bills of lading.

As for the first scenario, the shipowner could relieve himself of responsibility in the light of a FIOST clause incorporated, as the loss in question would result from a third party's, namely the shipper's or the charterer's, actions or negligence. Second, the receiver's recourse *vis-à-vis* the shipper or the charterer would be destined for failure too, as his contractual partner was the shipowner/carrier under the bill of lading rather than the shipper or the charterer. The very same can be said for the second scenario for the shipper against the charterer. In the first scenario, if the shipper or the charterer is a free on board [hereinafter FOB] or a cost, insurance and freight [hereinafter CIF] seller –unless there is a string sale-, the cargo interest as a buyer, indeed might have a direct contractual proximity with either of these parties under the sale contract, and

depending on its terms, he could be able to claim his loss under the sale contract, but not under the carriage contract. This may not be always the case though, as cargo interests often can be said to have been left without a remedy for the matter in question under English law. The *status quo* proves that loss arising from these operations does not fall upon the party that undertakes to perform these obligations, as he is allowed to walk away without liability.[29] It is hence opined that the end result is highly unsatisfactory for cargo interests, and they should therefore have a means of recourse directly or indirectly against the faulty party.

4.1. INTERVENING ACT OF THE SHIPOWNER

The very first question at this point needs answering is whether the shipowner/carrier would be subject to liability, in case he intervened or took some role in the stowage, where these operations were transferred to the charterer via a FIOST clause. It is trite law that regardless of whether it is the charterer who is to perform cargo-handling operations, the shipowner *"is practically bound to play some part in"*[30] in these operations and the master, to some extent, is under a duty to exercise supervision and control over the charterer's performance of these services.[31] If this is the case, then it is plausible to argue that the charterer's responsibility should be limited to the extent of the shipowner's contribution to the loss arising from cargo-related operations. Put simply, if the shipowner intervened with these operations, which contributed to it being improper, he should be liable to a corresponding degree against the bill of lading holder. This argument was embraced in *the Eems Solar* by the learned judge and he subsequently opined that the shipowner should not be relieved of liability for improper stowage, had the stowage plan made by the master and the mate –which failed to provide for the necessary precaution in the case- been followed by the stevedores and their employers. Though he went on to note that there was no evidence showing that the stevedores relied on or paid any attention to this plan in the case. On this ground, it was thus held that the entire responsibility for stowage rested with the charterers via their stevedores and accordingly the damage resulting from the failure to stow the cargo properly could not lead to hold the shipowner responsible.

In *the Eems Solar*, the cargo interests' attempt led to a failure. It was nonetheless only on the ground that there was no evidence that the shipowner intervened with the

[29] For similar criticisms *see* Todd, *supra* note 10. Simon Baughen, *Tripartite Contracts and the Missing Link*, 2004 Lloyd's Mar.& Com. L Q. 129 . Simon Baughen, *Defining the Limits of the Carrier's Responsibilities*, 2005 Lloyd's Mar.& Com. L Q. 153

[30] *Pyrene v. Scindia* [1954] 2 Q.B. 402, 418.

[31] *See generally Court Line Ltd v. Canadian Transport Co Ltd* [1940] AC 934, 944.

stowage or that the stevedores relied on their plan. However, it would not be wrong to argue that *the Eems Solar* left the door ajar for cargo interests' favour; as long as they can prove that the shipowner's intervention has contributed to the loss arising from these operations, shipowners can be held liable for a corresponding degree under the bills of lading. That being said, it might still prove to be difficult to show the intervening actions of the shipowner and their contribution to damage.

4.2. ACTION IN TORT

As shown above, it may not be often easy to prove the shipowner's intervention in these operations, as was the case in *the Eems Solar*. They therefore may need alternative means to fend off this undesired result. As for a claim in tort, the claimant must prove that he had the property in the goods when the damage occurred.[32] The vast majority of commercial vessels carry the goods sold under sale contracts on *shipment terms* –FOB and CIF sales– title to the goods will often pass after shipment such as on delivery or during transit. It is unusual for these sales that property in the goods pass prior to shipment. Therefore, the receiver would have a difficulty in proving that he had proprietary interest in the goods during loading, stowing or trimming operations causing damage to the goods, whereas, the shipper might have a better chance to claim in tort against the charterer performing these operations. On the other hand, for a loss arising during discharge, unless it is undertaken by the receiver, perhaps he might have an arguable claim against the faulty party provided that he establishes that he had the title to the goods at the time damage occurred. However, as *the Starsin*[33] proves, it may not always be encouraged to claim in tort actions, as it could be really challenging and difficult to show that they had the property in the goods at the time the goods were damaged.[34]

Even if shown, they may not be able to recover purely economical losses, which are not loss consequential upon physical damage, which might fall short of the actual damage borne by cargo interests.[35] In order to provide a right of recourse for cargo interests it is therefore important to scrutinise other potential legal devices, which will be examined below.

[32] *See Candlewood Navigation Corporation Ltd v. Mitsui OSK Lines Ltd* [1986] AC 1; *Leigh & Sillivan Ltd v. Aliakmon Shipping Co Ltd (The Aliakmon)* [1986] A.C. 785; [1986] 2 Lloyd's Rep. 1.

[33] *Homburg Houtimport BV v. Agrosin Private Ltd (The Starsin)* [2003] UKHL 12; [2004] 1 AC 715; [2003] 1 Lloyd's Rep. 571.

[34] *Id.*

[35] *See also Candlewood Navigation Corporation Ltd v. Mitsui OSK Lines Ltd* [1986] AC 1; *Leigh & Sillivan Ltd v. Aliakmon Shipping Co Ltd (The Aliakmon)* [1986] AC 785; [1986] 2 Lloyd's Rep. 1. For the application and approval of this principle see, *Losinjska Plovidba v. Transco Overseas Ltd (The Orjula)* [1995] 2 Lloyd's Rep. 395 and *Virgo Steamship Co SA v Skaarup Shipping Corp (The Kapetan Georgis)* [1988] 1 Lloyd's Rep. 352. For a detailed analysis on *the Orjula* see also, Andrew Tettenborn, *Dangerous Cargo: Tort Liability and Environmental Responsibility,* 1996 Lloyd's Mar. & Com. L. Q. 6, 8.

4.3. IMPLIED CONTRACT AND AGENCY PRINCIPLE

Since a claim in tort is challenging and the argument for the shipowner's intervention is not waterproof, a question arises immediately at this point: Is there an alternative means of solution favouring the cargo interest to reverse this undesired outcome? The short answer may not be affirmative, as the English courts have not issued a decision in this context favouring the cargo interest. Nevertheless the Court of Appeal in *the Coral*[36], suggested two propositions to address this issue based on the decision in *Pyrene v Scindia*; first, that the charterer could be deemed as having a contract with the shipper through the agency of the shipowner, and second that an implied contract might come into existence between the charterer and the shipper, once the goods are delivered by the shipper and they accepted by the charterer to perform the operations of loading and stowing.[37] Though the court did not go beyond noting these suggestions, nor did it explain how they would be applicable to the matter in question. The courts so far have not developed any solution in this context. These propositions, therefore, need to be explored and examined to see whether any of them could be a remedy for cargo interests.

4.3.1. AGENCY PRINCIPLE

To start with the agency principle, it would be safe to say that it is no longer considered as good law in the light of the decision of the House of Lords in *the Midland Silicones*[38] and the decision of the Court of Appeal in *the Kapetan Markos*[39] where the agency principle allowing a third party to participate in a contract was laid to rest.[40] It is, therefore, highly doubtful that the agency principle would prove a fruitful solution for cargo interests on the matter.

4.3.2. IMPLICATION OF A CONTRACT

As it is evident from the rulings in Midland Silicones and the Kapetan Markos, the first proposition made in the Coral seems not feasible, as a third party is unable to participate in a contract that was not concluded by him. However, both *Midland Silicones* and the *Kapetan Markos* went on to explain *Pyrene v Scindia* on the grounds of the doctrine of implied contract. It is trite law that the implication of a contract is a matter of fact to be

[36] *Balli Trading Ltd v. Afalona Shipping Co Ltd (The Coral)* [1993] 1 Lloyd's Rep. 1, 7.
[37] *Id.*
[38] *See, e.g.* [1962] AC 446, 471.
[39] *See generally The Kapetan Markos (No 2)* [1987] 2 Lloyd's Rep. 321, 331.
[40] *See also* Only parties to the contract can sue or to be sued thereunder; *Adler v. Dickinson (The Himalaya)* [1955] 1 Q.B. 158, 161, 162. See also, Aikens, *supra* note 4, at 276; Eder, *supra* note 4, at 69.

determined depending on the circumstances of each case.[41] The Court of Appeal in *the Coral* found the implied contract established in *Pyrene v Scindia* between the FOB seller and the shipowner could be analogous to the relationship between the charterer and the shipper. Devlin J in *Pyrene v Scindia* explained the implied contract "by delivering the goods alongside the seller impliedly invited the shipowner to load them, and the shipowner by lifting the goods impliedly accepted that invitation. The implied contract so created . . ." .[42] By way of analogy, it is submitted that a contract can be implied on the ground that the shipper by delivering the goods, could be regarded as inviting the charterer to load and stow them - as these operations are transferred to him by a FIOST clause- and the charterer could also be deemed to be accepting this invitation.

Under English law, a contract can be created as a result of parties' conduct, which could be deemed as offer and acceptance.[43] As was the case in *Pyrene v Scindia*, it is arguable that the shipper's offer of the goods to the charterer and the charterer's acceptance of them for loading, in return may be considered as offer and acceptance. Though in order to infer a contract, alongside offer and acceptance, contractual intention is also a necessary tool.[44] Considering the issue in question, indeed these parties can be said to be acting under another contract and it is not always easy to establish a contractual intention for an implied contract in the light of ordinary contractual principles. Though in *Pyrene v Scindia*, a contract was successfully implied from the conduct of the parties and the implied contract in *Pyrene v Scindia* was embraced by the House of Lords in *Midland Silicones* notwithstanding the lack of contractual intention.[45] In supporting this, case law appears to be proving that a contract could be created, although it was not possible to support its formation in the light of standard offer and acceptance analysis,[46] and that the existence of the implied contract was accepted, even regardless of the actual intention of the parties.[47] Also many learned scholars, like Treitel,[48] Debattista,[49] Lorenzon,[50] as well as the editors of *Chitty on*

[41] *The Elli 2* [1985] 1 Lloyd's Rep. 107, 111.
[42] [1954] 2 Q.B. 402, 426.
[43] *See Hart v. Mills* (1846) 15 LJ Ex 200; *Steven v. Bromley & Son* [1919] 2 K.B. 722; *see Greenmast Shipping Co SA v. Jean Lion et Cie SA (The Saronikos)* [1986] 2 Lloyd's Rep. 277; *see also Confetti Records v. Warner Music UK Ltd* [2003] EWHC 1274 [2003] EMLR 35.
[44] *See e.g., The Aramis* [1989] 1 Lloyd's Rep. 213, 224.
[45] [1962] AC 446, 471.
[46] *See, The Satanita* [1895] P 248; *Clarke v. Dunraven* [1897] AC 59; *The Eurymedon* [1975] AC 154. *See* also *Butler Machine Tool Co Ltd v. Ex-Cell-O Corp (England) Ltd* [1979] 1 W.L.R. 401.
[47] *See Cremer v. General Carriers SA (The Dona Mari)* [1974] 1 W.L.R. 341.
[48] *See generally* G.H. Treitel, *Bills of Lading and Implied Contracts,* 1989 Lloyd's Mar. & Com. L. Q. 162, 168-72.
[49] *See* Charles Debattista, Bills of Lading in Export Trade 10 (3rd ed. 2008).
[50] *See* Filippo Lorenzon et al., Sassoon on C.I.F. and F.O.B. Contracts 10-52 (6th ed. 2016).

Contracts[51] and *Carver on Bills of Lading,*[52] consider that a *Pyrene* type implied contract is still alive.

The courts, when explaining the underlying ratio for the implication of a contract, have taken a pragmatic approach in order to satisfy commercial expectations.[53] Regardless of the actual intention of the parties, contracts have so far been implied by the courts, when it was necessary to do so[54] and when business efficacy[55] required it. Thus, when a relationship between two parties is not entirely efficacious without the implication of a contract between them, the courts have shown readiness to imply a contract, when other requirements are satisfied.

FIOST clauses arguably have a detrimental impact on the business efficacy, which requires that a party should not be substantially deprived of the benefit that the parties intended him to receive, under the contract.[56] When shippers or consignees are not the charterers, the only document in their hand that provides contractual nexus with the carrier is the bill of lading and they hardly have a say on drafting its terms including the terms that are incorporated from a charterparty in relation to the obligations to load, discharge, stow etc. On the other hand, third party bill of lading holders are usually buyers under CIF and often FOB contracts[57] or sometimes banks who hold those bills as pledge or security in letter of credit transactions. These parties, as consignees, do not conclude the carriage contract contained in the bills of lading, never negotiate or draft its terms or never see the bill of lading until it is transferred, but only purchase the carriage contract which includes the terms that are incorporated from a charterparty.

Therefore, they purchase a carriage contract under which they may not be able to recover their losses from the shipowner as a result of a clause permitting the shipowner to reallocate the obligations of art.III(2). Loss or damage to the cargo is not uncommon in practice, and art.III(2) is one of the most frequently resorted provisions of the Rules by

[51] *See e.g.*, Hugh Beale, Chitty on Contracts 190 (32d ed. 2017).
[52] Treitel, *supra* note 48, at 314.
[53] *See The Satanita* [1895] P 248; *Clarke v. Dunraven* [1897] AC 59; *The Eurymedon* [1975] AC 154. *See* also *Butler Machine Tool Co Ltd v. Ex-Cell-O Corp (England) Ltd* [1979] 1 W.L.R. 401.
[54] *See The Aramis* [1989] 1 Lloyd's Rep. 213, 224.
[55] *Id.*
[56] *See Hongkong Fir Shipping Co Ltd v. Kawasaki Kisen Kaisha* [1962] 2 Q.B. 26, 70.
[57] Particularly under the second and third type of FOB contracts, where the buyer is not considered the shipper of the goods; *see Pyrene & Co. v. Scindia Navigation Co* [1954] 2 Q.B. 402; *Wimble v. Rosenberg & Sons* [1913] 3 KB 743; *The El Amria and The El Minia* [1982] 2 Lloyd's Rep 28; *Concordia Trading BV v. Richco International Ltd* [1991] 1 Lloyd's Rep 475; *Scottish & Newcastle Int Ltd v. Othon Ghalanos Ltd* [2008] UKHL 11; [2008] 1 Lloyd's Rep 462. For more detailed analysis on FOB contracts *see e.g.*, Guenter Treitel, *C.I.F. Contracts, in* Benjamin's Sale of Goods, ch. 19 (10th ed. 2017); Lorenzon, *supra* note 50, at 247; Debattista, *supra* note 49, at 8; Michael Bridge, The International Sale of Goods 89 (4th ed. 2017); Paul Todd, Cases and Materials on International Trade Law 94 (1st ed. 2003); Carole Murray et. al., Schmitthoff: The Law and Practice of International Trade 21 (12th ed. 2012); Andrea Lista, International Commercial Sales: The Sale of Goods on Shipment Terms 23 (2018).

cargo interests. Under these circumstances, cargo interests can be said to have been deprived of its business effect, as the protection provided by art.III(8) is also undermined.

As explained above, cargo interests appear to be left without a right of recourse unjustly, as a result of a FIOST clause incorporated into the bill of lading shifting the responsibility of the cargo-related operations to the charterer. It is submitted that it is entirely unsatisfactory and commercially inconvenient, since the charterer as the source of defect is allowed to escape liability and cargo interests are substantially deprived of the benefit that they are intended to receive under the contract. The author by no means proposes challenging the law that the shipowner is entitled to contract out these responsibilities. The problem is however concentrated on the loophole created as a result of FIOST clauses that the party who transferred these duties enjoys lack of liability, while he should also face consequences of these responsibilities. On this ground, the relationship between the charterer and the cargo interest is not entirely efficacious without the existence of an implied contract. Therefore, based on the principles of commercial necessity and business efficacy, the author promotes that a contract should be implied as analogous to *Pyrene* and used as a legal device in order to establish a right of recourse for cargo interests.

4.3.2.1. TERMS OF THE IMPLIED CONTRACT

Existence of an implied contract between these parties does not suffice to make the charterer responsible though. The question, therefore, remains as to what are the terms of this contract so as to see whether the charterer could be subject to liability. In general terms, it is trite law that although it is implied as a separate contract, it is accepted that it comes into existence on the bill of lading terms.[58] In *Pyrene v Scindia,* Devlin J noted "[t]he implied contract so created must incorporate the shipowner's usual terms . . . "[59] He went on to find that "they enter into upon . . . which they know or expect the bill of lading to contain" and the *Hague Rules* were considered "usual in the trade".[60] Eventually, Devlin J held that the implied contract was assumed to incorporate the Rules and accordingly the carrier was entitled to trigger the package limitation provisions *vis-à-vis* the seller. As for the issue in question, assuming that a contract is implied between the shipper and the charterer, a question therefore arises immediately at this point: would

[58] *See Brandt v. Liverpool, Brazil & River Plate Steam Navigation Co* [1924] 1 K.B. 575; *See also Pyrene v. Scindia* [1954] 2 Q.B. 402; *The Elli 2* [1985] 1 Lloyd's Rep. 107.
[59] [1954] 2 Q.B. 402, 426.
[60] *Id.* Provided that the Hague or the Hague Visby Rules are incorporated by a paramount clause in the bill of lading.

the shipper be able to trigger art.III(2) which reads ". . . the carrier shall properly and carefully load, handle, stow, carry, keep, care for, and discharge the goods carried"?

The Rules were found applicable in *Pyrene v Scindia* between the shipowner and the FOB seller who was not the shipper in the bill of lading, despite the fact that art.I(a) prescribes " "Carrier" and includes the owner or the charterer who enters into a contract of carriage with a shipper". It was evident in *Pyrene v Scindia* that the carrier in fact entered into a contract with the buyer who was the real shipper, not with the seller. It is not easy to accommodate the terms of the implied contract between its parties. As was the case in *Pyrene v Scindia*, the courts however showed some readiness to read its terms with more flexible approach and equally applied some verbal manipulation to give proper effect to the Rules thereunder. On this ground with some verbal manipulation, it would not be entirely implausible to argue that art.III(2) would be triggered by the shipper against the charterer under the implied contract, particularly with the support of art.I(a) in which it reads; " "Carrier" includes [. . .] the charterer[. . .]". Indeed, the charterer is not the real carrier under the bill of lading in respect of the matter in question. Although the seller in *Pyrene v Scindia* was not the real shipper under the bill of lading either, the *Hague Rules* applied between them regardless.

Some additional support -perhaps surprisingly- could be found in FIOST clauses. As stated above, it is common ground that the implied contract is on the bill of lading terms. A question arises at this point as to whether a FIOST clause that is incorporated into the bill of lading from a charterparty transferring the cargo-related responsibilities to a third party, such as the charterer, would be incorporated into the implied contract. In *The Elli 2*, there was a bill of lading incorporating charterparty terms including demurrage clauses. In the Court of Appeal, it was held that the contract implied was on the bill of lading terms which also included the incorporated clauses from the charterparty, accordingly, the shipowner was found entitled to demurrage under the implied contract.[61] On the basis of *The Elli 2,* provided that there is an incorporation clause in the bill, there is no reason why a FIOST clause of a charterparty should not be incorporated into the implied contract. If this proposition is correct, then a clause that deprives the cargo interest of having a recourse under the bill of lading contract might arguably give rise to a right to recovery under the implied contract against the charterer who is the source of defect. Assuming that a bill of lading incorporated a widely used FIOST clause, such as clause 5 of the Gencon 1994 Charterparty which reads:

[61] *See The Elli 2* [1985] 1 Lloyd's Rep. 107, 112 and 116.

> 5. Loading/Discharging (a) Costs/Risks
>
> The cargo shall be brought into the holds, loaded, stowed and/or trimmed, tallied, lashed and/or secured by the Charterers, free of any risk, liability and expense whatsoever to the Owners. The Charterer shall provide and lay all dunnaged material as required for the stowage and protection of the cargo onboard, the Owners allowing the use of all dunnaged available on board.

Once successfully incorporated into the implied contract on the basis of *The Elli 2,* without applying art.III(2), perhaps it might be argued that the shipper would put such a FIOST clause in use in his favour against the charterer, as it makes it clear that the risk of these operations and responsibilities fall on the "charterer". On the other hand, as for the receiver, it is submitted that it is doubtful whether the implication of a contract would provide any assistance. By virtue of the Carriage of Goods by Sea Act 1992, contractual rights and liabilities under the bill of lading[62] are transferred to receivers.

However, it does only transfer rights and liabilities of the original contract contained in the bill of lading. It is not plausible to argue that the Act also transfers rights and liabilities under a separate implied contract to a receiver. There is no reported case supporting this inference either. One argument would be through suggesting a similar proposition that a contract could be implied between the receiver and the charterer[63] by way of analogy with *Brandt v Liverpool.*

For a couple of reasons, it is however unlikely that a contract can be implied between the receiver and the charterer. First, it is often the receiver that undertakes to perform discharge, not the charterer. Once this is the case, if the fault during discharge lies with the receiver causing damage to his own goods, the charterer will not be subject to any liability. Second, even if the entire cargo-handling operations are performed by the charterer, it is doubtful whether a *Brandt v Liverpool* type implied contract would arise between the receiver and the charterer, particularly on the basis that its application is restricted by some authorities.[64] In *Brandt v Liverpool*, a contract was successfully implied between the carrier and the receiver on delivery of the goods against production of the bills of lading. Although, to imply such a contract there should also be a financial consideration moving from the receiver such as payment of freight or demurrage etc. Even if these operations are to be carried out by the charterer, pursuant to a FIOST clause, the receiver will be technically considered as demanding his goods from the

[62] As well as under seaway bills and delivery orders by virtue of Section 1 (b) and (c).
[63] Or the shipper depending on the wording of FIOST clause.
[64] *See The Owners of Cargo Iately Laden on Board the Ship Aramis v. Aramis Maritime Corp (The Aramis)* [1989] 1 Lloyd's Rep. 213; *See, e.g., Mitsui & Co Ltd v. Novorossiysk Shipping Co (The Gudermes)* [1993] 1 Lloyd's Rep. 311. *Seealso The Aliakmon* [1986] 2 W.L.R. 902.

carrier/the shipowner to whom the receiver will present the bill of lading. One may argue against this, as there is some authority indicating that a contract can be implied, even if no document is tendered in exchange for the discharge of the goods.[65] Even in case of this, it would be still doubtful that there could be any payment for freight or any other charges made to the charterer by the receiver. Therefore, unlike the shipper, the implication of a contract would unlikely be of any assistance to the receiver in creating a right of recourse *vis-à-vis* the charterer. One practical solution could be that the shipper might sue the charterer on behalf of the receiver under the implied contract that would come into existence between him and the charterer on the grounds advocated above.

4.4. ANALOGY WITH TRANSHIPMENT CASES

It is argued above that a *Pyrene* type implied contract could be a useful tool in providing a right of recourse for shippers against charterers. To fend off the detrimental impact of FIOST clauses on cargo interests, developing alternative legal devices might be necessary, given that the suggestions put forward above have yet to be tested before the courts and they may fall short in providing a right of recourse against the faulty party. Therefore, an alternative suggestion is submitted by way of analogy with the cases where transhipment occurred.

When a cargo is shipped under a through bills of lading for an intercontinental route, due to practical reasons, sometimes the contractual carrier may not undertake to perform the entire voyage but transship the cargo on a vessel that is to complete the remaining sea leg.[66] When this is the case, the contractual carrier sub-contracts with a carrier that is to perform the actual carriage. Though, this does not necessarily mean that the contractual carrier would be able to disclaim any liability after transhipment, for the remaining part of the voyage, which is to be carried out by the actual carrier, as there is a *dicta* in the *Holland Colombo* suggesting otherwise.[67] According to *dicta* of Privy Council in the *Holland Colombo*, a clause relieving the contractual carrier of responsibilities arising from loading, stowage, discharge after transhipment is of doubtful validity. Therefore, a bill of lading holder will be entitled to bring an action against the contractual carrier for the damage done to his cargo after transhipment

[65] *Cf. The Elli 2* [1985] 1 Lloyd's Rep. 107.

[66] The holder of the bill of lading will be able to have action *vis-à-vis* the contractual carrier for the entire voyage, as the bill will provide a continuous documentary cover; *see Hansson v. Hamel & Horley* [1922] AC 36, 46 in which the House of Lords followed *Landauer & Co v. Craven and Speeding Bros* [1912] 2 K.B. 94. *See also* Benjamin, *supra* note 57, at 19-027; Debattista, *supra* note 49, at 7.14.

[67] *Cf. Holland Colombo Trading Society Ltd v. Alawdeen* [1954] 2 Lloyd's Rep. 45, 53-54.

under a through bills of lading. The contractual carrier, however, will be able to claim indemnity from the actual carrier.

This might be of some assistance to the matter in question by way of analogy. Suppose that a FIOST clause of a charterparty allowing the shipowner to contract out his cargo-related obligations was incorporated into the bill of lading. He would be able to disclaim any liability arising from these operations, as he successfully transferred them to the charterer. Also, suppose that a contractual carrier sub-contracted for transhipment with an actual carrier who would perform the remaining leg and the cargo was damaged after transhipment due to improper stowage, which was undertaken by the actual carrier. In the latter example, in light of *dicta* in the *Holland Colombo,* the contractual carrier would remain liable for the improper stowage performed by his agent, the actual carrier after transhipment, even if there was a clause under the bill of lading relieving the contractual carrier from any liability concerning cargo-related operations, such as bad stowage. If this is the correct statement of the law, it is difficult to see that the carrier in the former example would enjoy contracting out these operations through a FIOST clause, and as a consequence would be relieved of liability, whereas the carrier in the second example would not, regardless of the existence of a similar clause disclaiming liabilities. Perhaps one plausible explanation would be that the contractual carriers are entitled to indemnity from their agents, actual carriers, for liabilities arising after transhipment, such as the ones arising from improper stowage. As for the matter in question, it is submitted that by way of analogy with the *dicta* in the *Holland Colombo*, once sued by the cargo interest, the shipowner could be allowed to claim indemnity from the charterer, if the charterer was deemed as agent of the shipowner in performing these operations, as is the case in transhipment cases. This would hardly do any harm to the shipowner's freedom of contract concerning the scope of the services that he would want to undertake, and on the contrary, it would be lawfully and commercially more convenient; first, because cargo interests would not be unjustly left without a remedy; second the risk of these operations would ultimately stay with the party that undertakes to perform them, namely the charterer.[68]

[68] Depending on the wording used in the FIOST clause, it could be the shipper as well.

5. CONCLUDING REMARKS

The burden of this paper has been to suggest solutions in order to revise the undesired outcome of FIOST clauses in the bill of lading over cargo interests under English law, not to challenge the law established that the parties are free to negotiate or draft their own terms under their contract. Having said that, such clauses, as argued throughout the paper, might create a loophole in favour of the party that undertakes to carry out these services, and as a consequence, allow them to escape liability. The outcome evidently has a detrimental impact over the business efficacy of the bill of lading, which requires that a party should not be deprived of substantially the benefit that the parties intended him to receive thereunder.

In an attempt to alter the impact of FIOST clauses, several propositions have been discussed above. Among these propositions, it is eventually submitted by the author that the two of them, -the implication of a contract and the analogy with transshipment- cases appear to be the most plausible legal devices that could be successfully put in use before the courts in order to establish a right of recourse for cargo interests. Indeed, these devices have yet to be tested before the English courts.

Nevertheless, it is thought that whoever is undertaking to carry out these obligations should also bear the consequences. Therefore, it is strongly opined by the author that the courts should show readiness to embrace these two devices to alter the displeasing result borne by cargo interests, as these two advocated devices may produce commercially and lawfully more convenient outcomes without confronting the freedom to contract out these responsibilities via FIOST clauses.

VOLUME 6 ISSUE 1 ARTICLES & ESSAYS

UNIVERSITY OF BOLOGNA LAW REVIEW

https://doi.org/10.6092/issn.2531-6133/13299

Received: 28 Aug. 2020 | Accepted: 14 Jun. 2021 | Published: 19 Jul. 2021

Constituting Over Constitutions

CARTER DILLARD

Carter Dillard, Master of Laws at New York University School of Law (United States of America) with a thesis reformulating the right to have children (Jeremy Waldron); former member of the Steering Committee of the Population Ethics and Policy Research Project, Future of Humanity Institute, and former visiting scholar, Uehiro Center, Oxford University (United Kingdom); policy advisor, Fair Start Movement; he can claim a string of extensive academic and public policy work on the fields of family planning and animal protection.

@ cjd328@nyu.edu

0000-0002-3774-9643

ABSTRACT

In philosophy, legal theory and law, the *Grundnorm*, or basic norm, is often assumed to be the constitution, or that which overrides other norms. That is incorrect. This paper argues that the *Grundnorm* should be the norm which regulates human procreation. This norm must proceed from the theoretical absence of human power, or a zero baseline. This essay attempts to correct the *Grundnorm* fallacy with what will be called the Zero-Baseline Model. The correction reorients our human rights regimes and family planning systems, in ways that lead to an inevitable list of specific policy reforms that largely invert current family planning models and policies in use at the United Nations, European Union, the United States, and elsewhere. Those reforms can all be described in a simple narrative of reorienting family planning laws and policies from what would-be parents desire, subjectively, towards what all future children need, objectively. And as the evidence shows, those reforms prove highly effective and much more efficient in promoting child welfare, reducing economic and other inequalities, mitigating the climate and other ecological crises, protecting non-humans, and building democracy, than their alternatives.

KEYWORDS

Constitutional Law; Public Policy; Grundnorm; Zero Baseline Model; Family Planning

TABLE OF CONTENTS

INTRODUCTION

This paper will invert what is commonly seen as an act of self-determination by parents, procreation, to reveal instead its true nature as an act of other-determination, on a massive scale more determinative of the future child's life and the communities future children will comprise than of the lives of the parents. This will lead to a revision of Hans Kelsen's *Grundnorm* theory, and in this revision, I will argue that polities, or ideally what will be legalities, form dynamically from the inside out via the norm that determines procreation, rather than I will argue that legalities form dynamically the norms which determines procreation from inside-out, rather than through traditional outward-facing borders defined by traditional Constitutions and defended by violence. We are constituting, not constituted, in the past. Or if put another way, the creation norm always precedes and overrides other norms which are wrongly perceived to be more fundamental, including rules of rec ognition and other seemingly constitutive norms. The practical revisions in law and policy to correct what we perceive to be fundamental will entail a special form of liberation. This revision raises key questions in the formation of any conception of liberal governance: If governance should create morally valuable options while treating all equally, must family planning systems be designed to give all kids a fair start in life, eliminating the arbitrary impacts of nurture that procreative autonomy would impose on future generations? Does that process include an ecological baseline, such that the future generations can self-determine free of all forms of oppression – including climate change – imposed by prior generations? Can states intervene in family planning decisions, as a matter of enforcing a peremptory norm and a child's human right to a fair start in life, to assure this? This revision is offered to inform the process, and to reverse a common misperception about what actions are personal versus interpersonal in nature in the context of procreation. That reversal could raise an impossibility theorem: That there is no way to make sense of liberation, empowerment, obligation, the rule of law, or legitimacy without accounting for the creation of persons, oriented from the zero-baseline discussed below, and determining how objective values in that process will inevitably enable or disable subjective choice.

In philosophy and law, the most basic norm, or *Grundnorm*, is usually assumed to be the superior norm guiding human behavior from which all inferior norms are derived.[1] That is incorrect. The fallacy is based on a series of mistakes, described in more

[1] *See, e.g.*, Stanley L. Paulson, *The Neo-Kantian Dimension of Kelsen's Pure Theory of Law*, 12 Oxford J. Legal Stud. 311 (1992); Brian H. Bix, *Kelsen, Hart, and Legal Normativity*, 34 Revija Za Ustavno Teorijo in Filozofijo Prava [Revus J. Const. Theory & Phil. Law], 25 (2018) (Spain); Hans Kelsen, Reine Rechtslehre: Einleitung in die rechtswissenschaftliche Problematik [Pure Legal Doctrine: Introduction to The Legal Problem] (F.

detail below.[2] Instead, the most basic norm is that which ought to account for the creation of humans, or the creation norm. This is the norm that frames—in the realm of human rights—population ethics, going beyond quantities of persons and maximizing economic growth to deal with the unavoidable qualities, relative positioning of persons, relation of persons to their ecologies, etc., as we constitute legalities.[3] It is lexically primary human right and correlative duty, the first and constantly recurring duty we face, as we reconstitute ourselves through children. If population ethics is about counting people, then the form referred to as constituting herein is about making people count – which is perhaps the core commitment of liberalism.

The mistake of not seeing the *Grundnorm* as creation norm is impeding family-planning policy reforms that are ten to twenty times more effective than their alternatives at promoting child welfare, reducing economic and other inequalities, mitigating the climate and other ecological crises, protecting non-humans, and building democracy.[4]

This inquiry began during the development of the deontological framework for population ethics and policies, as part of the Population Ethics and Policy Research Project, within the Future of Humanity Institute at Oxford University in 2014. The inquiry and the claims in this essay can be thought of in these simple terms: The fundamental idea of liberalism and the advent of human rights and democracy all rested on the necessary condition of people capable of self-rule, people who differed sharply from one another at the time, those born and raised under the preceding ideologies. Where did these new people come from? This essay will argue that, for the most part, they have yet to appear on earth.[5]

Deuticke ed., 1934); Miriam E. Oatman, *General Theory of Law and State by Hans Kelsen*, 40 Am. Pol. Sci. Rev. 131 (1946).

[2] *See infra* Part I.

[3] *See* "How Important Is Population Ethics?", *Practical Ethics* (blog), http://blog.practicalethics.ox.ac.uk/2014/10/how-important-is-population-ethics/.

[4] Matthew Hamity et al., *A Human Rights Approach to Planning Families*, 49 Soc. Change 469 (2019); Editorial, Population Growth and Climate Change, 337 BMJ (2008), https://www.bmj.com/content/337/bmj.39575.691343.80. M. Greenstone, A Looney, J Patashnik, and M Yu, Thirteen economic facts about social mobility and the role of education, Brookings Institution website, available at: https://www.brookings.edu/research/thirteen-economic-facts-about-social-mobility-and-the-role-of-ed; John Guillebaud, *There Are Not Enough Resources to Support the World's Population*, Ockham's Razor with Robyn Williams website, available at: https://www.abc.net.au/radionational/programs/ockhamsrazor/there-are-not-enough-resources-to-support-the-worlds-population/5511900 published June 10, 2014; Smaller families mean better lives for all. Population Matters website. Available at: https://populationmatters.org/smaller-families. Accessed February 12, 2020; Paul A. Murtaugh & Michael G. Schlax, *Reproduction and the Carbon Legacies of Individuals*, 19 Glob. Env't Change 14 (2009); "Having Kids White Paper", Having Kids website, https://havingkids.org/faqs/white-paper-draft/.

[5] The inquiry could also be stated this way: What is the state's interest in procreation? The answer disposes of the artifice of the state and becomes simply this: Being free and equal people.

This essay will describe the creation norm as a simple scale (0, 1, 2, . . .) in which: (1) all subsequent things being equal, self-determination should fall and being determined by others should rise, in linear form (given the political equality of persons), as people join the relevant polity (or "legality", more accurately), which exists against a baseline of non-polity (or point zero). (2) Self-determination is maximized through unique, temporalized, and interlocking values as well as specific thresholds for the quantity, qualities, relative positioning of persons to one another and their ecologies, and formation of new legalities, described in detail below. These values and thresholds orient around zero as a balance point (for example, the baseline deviation of which is causing harms like climate change and induced a shortening of pregnancies[6]). These values and thresholds represent the objective frame that makes subjectivity possible and are an ideal balance between group security and personal independence. (3) Persons are positioned to convert the influence of others into norms through consent, i.e., we can imagine the positioning of the antecedent and decentralized "we" group of free and equal persons,[7] the reference to which expressly or implicitly initiates and justifies modern Constitutions as in "we the people". This means that, consistent with point (1) above, legalities become increasingly illegitimate[8] because self-determination and personal sovereignty should give way to being determined by others, and the consent to others' power—and hence inclusive normativity—become increasingly impossible. (4) Concentration of power in extant persons and institutions is transferred to future generations, to make them a sufficient antecedent group, or "we the people" in quantity, quality, and relative positioning as they enter the legality.[9]

The norm maximizes consent by balancing the positive freedoms to morally valuable options in life that come with children's development and increased roles in self-governance, with the negative freedom of smaller and decentralized legalities and the restoration of nature.[10] It re-orients our current systems of socially and ecologically unsustainable human rights which give would-be parents the unfettered right to

[6] *See, e.g.*, "Many pregnancies are shorter as climate change causes more 90-degree days", *UCLA Newsroom* (blog), https://newsroom.ucla.edu/releases/shorter-pregnancies-climate-change.

[7] This differs from Rawls and others' original positions approach. See Richard Dworkin, *The Original Position*, 40 Univ. Chi. L. Rev. 500 (1973).

[8] For a more traditional take on illegitimacy see Carl Schmitt, Legality and Legitimacy (Jeffrey Seitzer trans., 2004); *see also*, "Legal Theory Lexicon: Legitimacy", *Legal Theory Blog*, at https://lsolum.typepad.com/legaltheory/2018/05/legal-theory-lexicon-legitimacy.html.

[9] *See* Gary S. Becker, A Treatise On The Family (2nd ed. 1991) ("[T]he interaction between quantity and quality explains why the education of children, for instance, depends closely on the number of children".); Susan J. G. Alexander, *A Fairer Hand: Why Courts Must Recognize the Value of a Child's Companionship*, 8 T.M. Cooley L. Rev. 273, 302 (1991). ("[W]hen people have fewer children, the value of each individual child increases enormously. Parents are willing to spend more on each child, e.g., by investing more in each child's training".).

[10] *See* Hamity et al., *supra* note 4.

proliferate, and would give future children have no right to guaranteed levels of welfare at birth, and would give humans no right to nature.[11] The norm pushes in the opposite direction of recent and blatant pro-natal moves by many governments to maintain population-driven economic growth despite the fact that it exponentially increases climate and other ecological impacts.[12] Thus, this norm is a necessary condition for any conception of human freedom, which makes the conception (4) dimensional, reforming former and relatively static conceptions, like Isaiah Berlin's.[13] Freedom, in this sense, could be better referred to as our natural sovereignty; it values people, makes people count or matter, and extends liberalism beyond the state–subject paradigm to protect us from one another at an irreducibly fundamental level.

This essay refers to the creation of norm—which tries to constantly reconstitute the ideal antecedent "we"—as the "zero-baseline model", or "constituting", as in "constituting an inclusive legality", as opposed to the relatively static, top-down, and enforced concept of a modern nation-state's constitution. Legitimacy derives from people, and people derive from the creation norm—that is the value and nature of constituting, as opposed to some historical document disconnected from the people to whom it refers. No country is constituted; if we take human rights and democracy seriously, there are groups of people who are either constituting or de-constituting relatively legitimate or illegitimate societies, or relatively consensual social organizations that are more or less legalities.[14] *Constituting*, in this sense, means moving beyond being a people for whom a cessation of violence and settling of property rights in a top-down state structure suffices as freedom and toward being a people physically constituted into consensual and normative communities of free and equal people.

Through the lens of the zero-baseline model, things like the climate crisis and its impact on self-determination might take on a new significance, as we are forcibly subjected to the threatening power and influence of others. Things like recent pro-natal legislation in the United States, which will continue to exacerbate the crisis, also take on new meaning.

The creation norm has never existed in positive law or been practiced. Had it been, the world would look nothing like it does today. This essay will argue that the

[11] *See* Carter Dillard, *Comprehensive Animal Rights*, Having Kids website, https://havingkids.org/comprehensive-animal-rights/.

[12] *See* Dylan Matthews, "Mitt Romney and Michael Bennet just unveiled a basic income plan for kids", *Vox* (blog), https://www.vox.com/future-perfect/2019/12/16/21024222/mitt-romney-michael-bennet-basic-income-kids-child-allowance; and Jeff McMahon, "The World Economy Is a Pyramid Scheme", Forbes, April 5, 2019, at https://www.forbes.com/sites/jeffmcmahon/2019/04/05/the-world-economy-is-a-pyramid-scheme-steven-chu-says/.

[13] *See* Isaiah Berlin, Liberty (2nd ed. 2002).

[14] For a take on legality moving in this direction see Scott J. Shapiro, Legality (2011).

creation norm is that which constitutes legalities, or legitimate social organizations, and if what this essay argues is true, the world remains in pre-constitutional states of illegitimacy. Establishing the norm as the foundation from which future generations are created, legitimating all forms of social organization would require radically rethinking population ethics, law, and family-planning systems from their current trajectories.

Establishing the norm would require moving from a subjective norm that protects what parents want, to an objective "child-centric" norm that prioritizes what children need (because having a child is more of an other-determining than a self-determining act), using the specific values and thresholds referred to above, for their creation. The shift would require recognizing the creation norm as a peremptory norm in international law and embracing it as a socially and ecologically regenerative—and hence central organizing principle—for new policies that could prove ten to twenty times more effective at mitigating the climate crisis, restoring our ecology, improving child welfare, reducing economic equality and building democracy than current policies.[15] The transition to this norm essentially entails us becoming sufficiently other-regarding, or child-centric, to reconstitute our polities into legalities, by changing the way we "make" people.

What does the transition look like in practice, and how feasible is it? In short, the transition involves a significant and worldwide transfer of resources—treated as fundamental decentralization of public and private power in the form of substantial future child entitlements that will merge the economic margins of rich and poor (or more accurately powerful and powerless)—in order to incentivize changes in the way we plan our families and have children. Given the current distribution of resources, the malleability of family-planning decisions, current fertility rates, shifting gender dynamics, the threat of the climate crisis, and the way the zero-baseline model unifies the many generations of human rights, liberalism, communitarianism, etc., this shift is feasible.

This essay will lay out the zero-baseline model and the assumptions it is based on and will begin to spell out a system and discourse for the model's implementation in law and social praxis, including its potential role in specific campaigns developing international law and in pending federal climate litigation and proposed federal and state child welfare legislation in the United States. While the model is discussed in the familiar context of nation-states, it can be applied to the dynamics of any grouping of persons and forms of power (e.g. exemplification) that influence upon and within those groups.

[15] *See* Hamity at al., *supra* note 4.

If there is one thing this essay should convey, and especially at the discussion of "pregnant pronouns" and the lesser power asymmetry below, it is that the way most of us experience and react to power is myopic, temporally and otherwise, and that we are capable of instead more broadly comprehending the dynamic social ought of constituting legitimate societies. Attacking the pregnant pronouns in any normative claim, with questions about how the claimant is accounting for values like wellbeing, equity, nature, and democracy in the creation of the persons to whom the pronouns refer, seriously problematizes the claims.

BACKGROUND AND THE LIMITS OF HUMAN COGNITION

As discussed above, creating a comprehensive liberal normative system, like a human-rights-based democracy,[16] would first require an antecedent norm that accounts for the creation of the system's members, the relations between them, and the degrees of power or influence (as opposed to specific forms of power or influence, like violence)[17] they would exert upon one another. And, as discussed, this norm does not exist for many reasons and may have never so-far existed.[18] These reasons include a series of mistakes in the development of liberalism in addition to the cognitive challenges discussed below.

This essay will not discuss these mistakes in detail, which are covered in a companion article.[19] In short, liberal theory and practice (1) conceived power too narrowly, focusing on forms like violence, and ignoring subtleties like the conditions of child development and anthropogenic impacts on our ecology; (2) never seriously accounted for the birth of persons into the social contract and the way they inevitably exclude others from their sovereign role in that contract; (3) never seriously accounted the need to ensure that all persons are capable of being emancipated, or actually capable of participating in a social contract; (4) made the order error of conceiving freedom in

[16] *See* Mary Ann Glendon, A World Made New: Eleanor Roosevelt and the Universal Declaration of Human Rights (2002).

[17] "Power" or "influence" here is conceived of simply as a degree of *any* effect, consequence, influence, result, etc., the cause of which is a human. This includes the sort of subtle effects that collectively caused the climate crisis. That crisis threatens political structures, as well as our species, in ways never seen before, and it is proof that more-narrow conceptions of power—as authority, violence, control, etc.—should now be seen as dangerously myopic and as tied into the psychological limits of the way humans perceive. Power is about the capacity of each person to influence, and, therefore, it is about each person's presence. Alexander Barker, St. Benets, University of Oxford, "What is Power?" (unpublished manuscript, Word file).

[18] Sarah Conly, One Child: Do We Have a Right to More? (2016); Carter Dillard, *Antecedent Law: The Law of People-Making*, 79 Miss. L. J. 873 (2010).

[19] Carter Dillard, *Illegitimacy*, Willamette L. Rev., Forthcoming Spring 2020.

terms of what people can do without first accounting for who people are and first relate, in terms of their quantities, qualities, relative positioning, etc.; and (5) tried to seed an ideology of human liberation from a completely antithetical position: that of parents' property rights in their future children.[20] Perhaps the biggest mistake in all of this was relying on social processes, e.g. courts' interpretations of traditional constitutions, to subvert the *Grundnorm.* That is an order error; there is no social source or process that precedes the norm itself. Understanding that means taking rights seriously.

The absence of a theorized creation norm may account for the key conundrum in liberalism today: The average person has little influence over the forms of human power that influence him or her, like climate change and other environmental crises, global markets, risk of nuclear war, etc. A person most often lives under compulsory norms and social forms that do not reflect his or her deliberate input and will. How is that possible in an era defined by the supposed rise of human rights and democracy, which, by any interpretation, was meant to empower individuals and make them self-determining?

Could the absence of such a creation norm be accounted for by limits on human cognition, our temporal myopia, our tendency to notice and focus on obvious forms of influence, like violence, over subtle ones like greenhouse-gas emissions?

Accounting for political obligation and autonomy by temporally accounting for all the people with potential to influence our lives, in a world of close to 8 billion people, is not something our human brains can easily perceive.[21] The dynamics of autonomy and self- versus other-determination in this sense, or the temporalized presence and interaction of the actual people in the social contract are not things we normally comprehend. Human brains readily focus more on the symbols of that power and contract, our elected representatives, the words of an ancient and property-based constitution,[22] and violence and the threat of violence from law enforcement than focusing on the myriad ways we influence and define—and thereby limit—the lives of persons who do not yet exist.[23] We are more apt to see freedom in a myopic and downstream way (as exemplified by the notion of procreative autonomy), as if freedom

[20] *Id.*

[21] Darrell A. Worthy, A. Ross Otto and W. Todd Maddox, *Working-Memory Load and Temporal Myopia in Dynamic Decision-Making*, 38 J. Experimental Psych.: Learning, Memory, and Cognition 1640 (2012); Bruno Goncalves, Nicola Perra, and Alessandro Vespignani, *Validation of Dunbar's Number in Twitter Conversations*, arXiv (2011): arXiv:1105.5170v2. The birth and development of humans—*qua* the addition of new members of a liberal normative system contingent on particular ecologies—simply happens too slowly and amorphously for humans to readily perceive and account for.

[22] Carter Dillard, *Future Children as Property*, 17 Duke J. Gender L. & Pol'y 47 (2010).

[23] Philippe Julien, Jacques Lacan's Return to Freud: the Real, the Symbolic, and the Imaginary (1994); *Id.*; Dillard, *supra*, note 18.

were some sphere that magically separates us in time and space from the rest of the universe.[24]

This temporal error,[25] i.e., the difficulty of even conceiving of future humans, goes both ways: humans fail to account for how behaviors today determine the future as well as how social position or class is determined by past events. And that temporal error leads to a spatial one, a myopia, exemplified by H.L.A. Hart's famous statement of the basis for choice-rights theory (and a basis for many subsequent free-market theories and policies).[26] Hart's theory of choice-based rights begins with a minimal and subjectivist moral account of rights:

> [I]f there are any moral rights at all, it follows that there is at least one natural right, the equal right of all men to be free. The moral justification [for the right] does not arise from the character of the particular action to the performance of which the claimant has a right; what justifies the claim is simply—there being no special relation between him and those who are threatening to interfere to justify that interference—that this is a particular exemplification of the equal right to be free.[27]

Hart's account is myopic. How did the actors in this scene get to where they are? What is the context, historically, that accounts for their interaction? What form of limiting social organization has surrounded them and surrounds them now?[28] What accounts for any difference in the levels of influence they have over each other in that instance? This myopia has affected modern political discourse in what Nagle and Murphy called the "myth of ownership", or the perception that property rights and social positioning are naturally occurring phenomena,[29] as is an expectation that the state exists to protect them (through violence). Is there any reason to institutionalize this myopic notion of freedom as a value?[30]

The zero-baseline model tries to account for these errors in the way humans perceive and process information about their environment and that interfere with

[24] Joseph Raz, The Morality of Freedom (1986).
[25] Worthy et al., *supra* note 21.
[26] *See* Walter E. Block, *Private Property Rights, Economic Freedom, and Professor Coase: A Critique of Friedman, Mccloskey, Medema, and Zorn*, 26 Harv. J. L. & Pub. Pol'y 923,951 (2003).
[27] H. L. A. Hart, *Are There Any Natural Rights?*, 64 Phil. Rev. 175,191 (1955).
[28] Planned Parenthood v. Casey, 505 U.S. 833 (1992); Washington v. Glucksberg, 521 U.S. 702 (1997).
[29] Liam Murphy & Thomas Nagel, The Myth of Ownership: Taxes and Justice (2002).
[30] *See* Philip Pettit, *The Instability of Freedom as Noninterference: The Case of Isaiah Berlin*, 121 Ethics 693,716 (2011). To Raz and Pettit's critique, the zero-baseline model simply adds this: any sense of personal autonomy is an illusion that ignores the norm or norms which created the person in question and surrounds him or her, and his or her future, with other persons. The antecedent and external autonomy defined by the people with whom we share the world encompasses any subsequent and internal autonomy, or "personal sphere", we may misperceive as the exclusive form of autonomy.

achieving liberalism.[31] The model requires and uses novel concepts, like the temporalizing verb-form conception of "constituting", as opposed to the noun-form conception of a "constitution", which will be unfamiliar to the reader, but are explained below, are necessary to account for the model of intergenerational consent and self-determination advocated for here. The model involves other concepts, like a spatial-temporal border that seeks to account for all ways humans might influence one another, as well as a duality, or pincer, described below, in which a non-human-centric ecological ethic is paired with a child-centric family-planning model.[32]

This missing antecedent norm, or creation norm (*Grundnorm*[33]), or the "zero-baseline model", could hold the key to solving the conundrum and ensuring the special form of political evolution—the dynamic balance between community and autonomy—that liberalism always promised. The zero-baseline model contributes to the fields of population ethics, biopolitics, and legal theory by offering a new and comprehensive deontological framework, baseline, or point of orientation for evaluation,[34] discourse, praxis for family-planning and population policy reform that would be ten to twenty times more effective[35] at mitigating climate change, eliminating child poverty, improving economic inequality, and reaching other sustainable development goals than downstream approaches (i.e., norms that address behaviors after the creation of persons).[36] The model will be presented as a necessary, but not sufficient, condition of any reasoned conception of political obligation and self-determination/autonomy, and hence liberty and freedom.[37]

[31] *See* "What Is the Zero Baseline Model", *available at* https://fairstartmovement.org/what-is-zero-baseline-modeling-or-constituting-in-politics-and-law/.

[32] *See* Dillard, *supra*, note 11.

[33] Contrast traditional but relatively downstream conceptions of the *Grundnorm*.*See* Hans Kelsen, Pure Theory of Law (Max Knight trans., 1967); Juan Carlos Riofrío, *Kelsen, The New Inverted Pyramid and the Classics of Constitutional Law*, 7 Russ. L. J. 87,118 (2019), https://doi.org/10.17589/2309-8678-2019-7-1-87-118; Kelsen's *Grundnorm* was subsequent. What norm accounts for the presence of the people in Kelsen's systems? What norm should that be? Would a *Grundnorm* be normative without accounting for and aligning the fundamental values described below?

[34] *See* Legal Theory Lexicon 066: Baselines, https://lsolum.typepad.com/legal_theory_lexicon/2008/05/legal-theory–2.html.

[35] *Ibid*., note 2.

[36] *See* The Fair Start Model of Family Planning, Having Kids website, accessed August 15, 2019, https://havingkids.org/wp-content/uploads/2019/05/FairStartModel2019-1.pdf.

[37] *See* Alex Cole, *Topics in Political Theory: Is Freedom from Power Possible?*, Discourses on Liberty (blog), February 3, 2012, http://discoursesonliberty.blogspot.com/2012/02/is-freedom-from-power-possible.html.

THE ZERO-BASELINE MODEL

The zero-baseline model works by revising the consent-as-justification theory[38] of political obligation and autonomy by temporalizing it, de-abstracting its spatial component (accounting for actual people and their influence on the world), and expanding its perspective by three hundred and sixty degrees to account for all human influence.[39] It does so by first treating the relevant borders of human influence as starting at the pre-existence of future generations,[40] and ending with the edges of extant spheres of human influence.[41]

The perspective this model requires envisions, a four-dimensional social contract—or *Grundnorm*—requiring the consent or constructive consent of all extant and future persons[42] and thus, their integration into human-rights-based democracy. The model accounts for the mutual consent between extant and future persons, through those persons' entry into, relationship to, norms of, and right to exit the human-rights-based democracy.[43] In this model, humans come together in social contract, constituting society from a state of nonpolity, thereby converting all actual and potential human influence into law through the consensual integration of extant persons. The model allows for the measuring and thus, the minimizing of the inevitable and linear loss of political autonomy and self-determinability that comes from adding new members to the human-rights-based democracy. It maximizes members' consent to the influence that results from the existence of the others.[44] The model is antecedent, or a *Grundnorm*, because it accounts for the capacity to influence simply by existing.

Conceptualizing the model begins with a series of assumptions derived from the conceptions of consent traditionally used in consent-as-justification theories.[45] These assumptions spell out the necessary conditions for the existence of consensual

[38] *See* Harry Beran, *In Defense of the Consent Theory of Political Obligation and Authority*, 87 Ethics 260 (1977).
[39] *See supra*, note 17, for definition of "influence".
[40] *See* Matthias Doepke, *Gary Becker on the Quantity and Quality of Children,* 81 J. Demographic Econ. 59 (2015). *See e.g.* Eric A. Hanushek, *The Trade-off Between Child Quantity and Quality,* 100 J. Pol. Econ. 84 (1992).
[41] *See* Joel Feinberg, *The Rights of Animals and Future Generations, in* Philosophy and Environmental Crisis 43 (William T. Blackstone ed., 1974). Carter Dillard, *The Primary Right,* 29 Pace Env't. L. Rev. 860 (2012). *See also*, Alan Weisman, The World Without Us (2007).
[42] Derek Parfit, *Future People, the Non-Identity Problem, and Person-Affecting Principles,* 45 Phil. & Pub. Aff. 118 (2017).
[43] Dillard, *supra* note 41.
[44] The model could be represented in this way: 0 is nonpolity, 1 is absolute political autonomy or the individual sovereignty, and >1 begins the social contract with varying levels of political obligation and autonomy, depending in part on the application of the *Grundnorm*, relative to 0. With the zero-baseline model, systems remain pre-constitutional because they cannot account for the paradox of maximizing autonomy through the integration and obligation that happens between 1 and >1, for example, at the level of the townhall meeting hypothetical discussed below. The model thus acts as a measuring stick for the antecedent conditions of political obligation and autonomy.
[45] *See* Richard Dagger & David Lefkowitz, *Political Obligation, in* Stanford Encyclopedia of Philosophy (2021).

identities,[46] including the ways in which persons in a group themselves within society (i.e., how they form, or "constitute"[47]). (The assumptions do not spell out the conditions for what those identities subsequently *do*; they are assumptions simply about the entities' existence.) The model relies on the following assumptions:

a. any theory of political obligation and autonomy must account for all forms of human influence, including the absence of human influence, or non-polity, as a point of orientation.[48]
b. because we *are* before we *do*, the norm that creates us is lexically primary.[49]
c. members of a human-rights-based democracy must be of certain constitutive quality,[50] i.e., capable of consenting,[51] reasonable,[52] emancipated,[53] able to communicate with each other at a constitutive level,[54] etc.

[46] In case it's not readily apparent, "consenting identities" refers to those people capable of consenting, i.e., the actually emancipated adults, the adults who are at or above the constitutive quality threshold required to be met for membership into the society in the model.

[47] *See* " Reframing Population and Family Planning to Focus on Justice", *available at* https://fairstartmovement.org/reframing-population-and-family-planning-to-focus-on-justice/.

[48] *See* Carter J. Dillard, *Rethinking the Procreative Right*, 10 Yale Hum. Rts. & Dev. L. J. 1, 10 (2007). The interplay between population growth and nonpolity is what drove the need for government, according to Locke. Or, as may be the case, the government was created by self-interested elites, with its creation leading to population growth:

> The imperative of collecting people, settling them close to the core of power, holding them there, and having them produce a surplus in excess of their own needs animates much of early statecraft. Where there was no preexisting settled population that could serve as the nucleus of state formation, a population had to be assembled for the purpose.

From James C. Scott, Against the Grain: A Deep History of the Earliest States (2017).

[49] *See* John Rawls, Justice as Fairness: A Restatement (Erin L. Kelly ed., 2001).

[50] Some scholars have taken the perspective of current members and their willingness to admit and enter into contract with others, in some case with those who may be disabled or of another species. *See, e.g.*, Martha C. Nussbaum, Frontiers of Justice: Disability, Nationality, Species Membership (2006). But what if the prospective members are not disabled or of another species but will simply lack some sort of minimum civic quality?

[51] Contrast this with the constructive consent of the new parties themselves (i.e., offspring). *See, e.g.*, John Lawrence Hill, *What Does It Mean to Be a "Parent"? The Claims of Biology as the Basis for Parental Rights*, 66 N.Y.U. L. Rev. 353, 384 (1991). ("[T]he procreative right arguably is contingent upon the constructive consent of the resulting child".).

[52] *See* Marilyn Friedman, *John Rawls and the Political Coercion of Unreasonable People*, *in* The Idea of a Political Liberalism: Essays on Rawls 16 (Victoria Davion & Clark Wolf eds., 1999); *see e.g.*, Carter Dillard, *Empathy with Animals: A Litmus Test for Legal Personhood?* 19 Animal L. Rev. (2012). Could certain meta-ethical positions, like subjectivism, be explained as simply lacking a constitutive perspective, and requisite qualities, like empathy? *See also*, Maxine Eichner, *Who Should Control Children's Education?: Parents, Children, and the State*, 75 U. Cin. L. Rev. 1339, 1361 (2007) (describing why a democracy requires critical thinking in its citizens), at 1341; see also John Finnis, Natural Law and Natural Rights 260, 276 (2d ed. 2011); *see e.g.*, Joseph Raz, Practical Reason and Norms 144 (1975). Referring to the need for people of "sufficiently high quality" in democracies (quoting Joseph A. Schumpeter, Capitalism, Socialism and Democracy 290 (3d ed. 1950).).

[53] This term refers to a child we would accept into society who becomes mature enough to be considered an adult and who is free of the wardship of his or her parents. *See* David Archard, Children: Rights and Childhood (1993). Archard's discusses that varying levels of peoples' quality may be controversial, but that controversy disappears when parents are forced to explain why they put certain efforts (or any efforts) into rearing their children.

[54] For example, the United States Constitution says, "we the people", and there is a common assumption that the whole document should represent a discourse all Americans can have. And this assumption includes in

d. the constitutive quality requirement must be a part of the model that creates a human-right-based democracy because it is preferable to exclude merely possible people rather than extant people.[55]
e. each new member of a human-rights-based democracy is an equal individual sovereign who excludes (via things like majority rule) other members from the sovereignty equally,[56] and must account for and consent to each new member's capacity to influence, as well as their actual influence over themselves.[57]
f. because each member of a human-rights-based democracy retains ultimate political authority, the first would-be members must (in what might be called the "first election") constructively consent to become a new member of the democracy, consent to the addition of new members to the democracy, and must be free to reconstitute new human-rights-based democracies relative to the non-polity baseline.
g. in an ideal human-rights-based democracy, the consensual addition of new members would be the only thing that reduces political autonomy and creates political obligation.[58]

These assumptions could be summarized more simply: there is no law without obligation[59]; no obligation without consent[60]; no consent without accounting for the quantity, quality, and relative positions of the relevant entities; and no such accounting

"all Americans" those Americans who are illiterate or who cannot read English; the Constitution is still a means of common communication.

[55] *See* Dillard, *supra* note 22. *See also* Santosky v. Kramer, 455 U.S. 745, 789 (Rhenquist J., dissenting):

> It requires no citation of authority to assert that children who are abused in their youth generally face extraordinary problems developing into responsible, productive citizens. The same can be said of children who, though not physically or emotionally abused, are passed from one foster home to another with no constancy of love, trust, or discipline.

Consider John Stuart Mill's take on the circumstances one is born into:

> Capacity for other nobler feelings is in most natures a very tender plant, easily killed, not only by hostile influences, but by mere want of sustenance; and in the majority of young persons it speedily dies away if the occupations to which their position in life has devoted them, and the society into which it has thrown them, are not favorable to keeping that higher capacity in exercise.

[56] Each member's role in a democracy is, *cetus paribus*, inverse to the size of that democracy. *See e.g.*, Jean-Jacques Rousseau, On the Social Contract 50-51 (Donald A. Cress trans., Hackett Publishing Company 1988) (1762) at 50-51. Note that equality and egalitarianism, in this formulation, can simply be explained as comprehensively accounting for power. Anything other than equality would require an explanation of the power differential.

[57] This assumption, and the constitutive quality assumption, may have been why Mill seemed to think that procreation violated the harm principle in many cases. *See e.g.*, (Elizabeth Rapaport ed., Hackett John Stuart Mill, On Liberty Publishing Company 1978) (1859).

[58] Among the ways we commonly otherwise reduce it: delegation of authority to representatives, the use of sanctions or incentives to ensure compliance with norms, tolerating power not accounted for by its conversion into law, etc.

[59] *See generally* H.L.A. Hart, The Concept of Law (3d ed., Oxford University Press 2012) (1961).

[60] Contrast this with others' take on the nature of law. *See* Frederick Schauer, The Force of Law (2015).

without nonpolity, or the relative absence of human influence. Thus, nonpolity (to be contrasted with states of nature and original conditions in prior political-obligation models)[61] is the zero-baseline from which quantity, quality, relative positioning, and the resulting loss of human autonomy must be measured; any other baseline would not be comprehensive. Just as we use total darkness as the implicit baseline to determine how bright a room is, we must use the total absence of human influence (nonpolity) to measure the influence resulting from a person's existence or actions.

Given these assumptions, the *Grundnorm* (the constituting norm) is comprised of four interdependent requirements: to minimize the loss of political autonomy in a human-rights-based democracy, and hence maximize consent, integration, and justified political obligation, would require, as a matter of lexical primacy, that (1) each new entrant be of a minimum constitutive quality or capable of constituting with others, (2) that there be a maximum number of members constituting the human-rights-based democracy, (3) that they enter relative to each other so that they exclude equally,[62] and (4) that they are capable, given their quality, quantity, and relativity, of reconstituting their legalities relative to some level of nonpolity.

The existence of these four variables is inevitable in the act of procreating. The *Grundnorm,* thus, necessitates backdating the standards, e.g., requiring conditions of entry that ensure constitutive quality and emancipation,[63] rather than ignoring entry and subsequently excluding persons as insufficient in some way.

This model would alter common approaches in fields such as population ethics[64] that currently focus only on quantity and utility rather than accounting for the antecedent factors at play when we account for the constituting norm.[65]

The constituting norm changes our perspective. Instead of a quantitative concept of population, the focus turns to the dynamic "we" that enables constitutions and the many values which that thick concept[66] implies. Instead of asking what role a given population or quantity of persons plays in a particular outcome,[67] e.g., the creation of greenhouse gas emissions, we can begin to evaluate how the four variables have and

[61] *See* David Boucher & Paul Kelly, The Social Contract from Hobbes to Rawls (1994).
[62] For example, each person gets equal time to speak in the townhall hypothetical, *infra.*
[63] Locke's infeasible solution was to prevent unreasonable people from ever leaving the wardship of their parents. *See e.g.*, Dillard, *supra* note 48, at 40.
[64] What is being espoused here is to be distinguished from the field of philosophical inquiry currently known as "population ethics", which is dominated by utilitarians and largely focuses on how population and welfare relate. Population ethics will partly determine what the substantive content of post-constituting law should be, but this article raises legal and political questions beyond population ethics.
[65] *See generally* Hilary Greaves, *Climate Change and Optimum Population,* 102 The Monist 42 (2019).
[66] *See* "Thick concept", Wikipedia website, https://en.wikipedia.org/wiki/Thick_concept.
[67] *See* Paul R. Ehrlich & Anne H. Ehrlich, Population, Resources and the Faith-Based Economy: The Situation in 2016, Biophysical Econ. & Res. Quality, Aug. 2016, at 1, 3.

will subject persons, who in liberal societies should be self-determining, to existential threats created by others, or in other words, having their lives and futures so determined by others.

Note also how equality or fairness as a value, or being what Rawls called "free and equal people" [68] is accounted for simply as a matter of power, or accounting for all human influence relative to nonpolity or a zero-baseline. Any power differential that would precede their coming together, or constituting, would mean we had not accounted for all the power people have had over each other. When we do so from a zero-baseline, we coincide with equality. Hence the phrase "free and equal" might actually be redundant.

This model uses eight ultimate and interdependent values (each of which can be pegged to existing positive law standards)[69] to set the minimum and maximum thresholds for the quality, quantity, relativity,[70] and nonpolity requirements. The model is put into action, as discussed below in Part II, by using the values to restructure the fundamental — and universal — human right to procreate as it currently appears in ethics, law, policy, and praxis and gearing all family planning around the restructured right.[71] The model replaces the current value at the core of the structure of the right, which is "procreative autonomy", a contradiction in terms that beautifully exemplifies human temporal myopia,[72] and that seems to have been designed to alleviate collective obligations to future children.[73] How can the act of creating another person be autonomous and self-determining, as opposed to inter-relational and other-determining?[74] The mode replaces procreative autonomy with the eight values in the *Grundnorm*.[75]

[68] Claus Dierksmeier, "John Rawls on the Rights of Future Generations", *in* Handbook of Intergenerational Justice (2006).

[69] Backdating existing positive law standards can convince skeptics to accept the zero-baseline model, but these standards should not supplant our ideals when using the model to develop policies.

[70] That is, the relative position of the people created, for example, how they relate to each other in terms of, say, access to healthcare?

[71] *See* Dillard, *supra* note 48.

[72] *See* Martin Held, *Sustainable Development from a Temporal Perspective*, 10 Time & Soc'y 351 (2001).

[73] *See* Dillard, *supra* note 22. Parental procreative autonomy may also be a way in which parents' guilt about having brought a being into existence who was bound to suffer and to die, experiencing harms that would not have befallen the being had he or she not come into existence, an existence that was solely created by an action of his or her parents. Some see this creation of life as unethical regardless of the child's status, health, or place in the world. *See* David Benatar, Better Never to Have Been: The Harm of Coming into Existence (2006).

[74] *See* Dillard, *supra* note 48.

[75] U.N. Secretary General U Thant famously stated that "the Universal Declaration of Human Rights describes the family as the natural and fundamental unit of society. It follows that any choice and decision with regard to the size of the family must inevitably rest with the family itself, and cannot be made by anyone else". *See also* Proclamation of Teheran, Final Act of the International Conference on Human Rights, A/CONF.32/41 at 3 (April 22 to May 13, 1968) ("Parents have a basic human right to determine freely and responsibly the number and spacing of their children".). Thant's statement makes no sense. The choice and decision should

Is there an analogy for this move? Speech norms (which are based on self-expression rather than autonomy) are a much better model for the values that might underlie a creation norm. Speech norms are also nuanced, with clear limiting duties, to ensure human-rights-based democracies function. Why wouldn't the creation norm that preceded speech norms be as well?

The values that comprise the *Grundnorm*, which try to maximize obligation and autonomy through entry, role, and exit are the following: Treating procreation as the improved continuation of the parent's life ("improved continuity", or I.C.),[76] a minimum level of welfare at entry, the nexus or zero-baseline point at which the values are first determined, fairly defined by the welfare[77] of other entrants (> M.W.(F)), a qualitative threshold of a minimum quality of equitable individual sovereignty defined by the standard for emancipation, or reason (> E); a quantitative threshold of a maximum number of individual sovereigns defined by a minimum of individual sovereign roles in their democracy (< R);[78] and finally, a quantitative threshold of a maximum number of individual sovereigns defined by nature (< N), the nexus or zero-baseline point or balance point at which the values are last determined.

The zero-baseline model can be illustrated and supports intuitions triggered by examining the pronouns, which might be called "pregnant pronouns",[79] in any abstract normative claim. The claimant should be asked to de-abstract their usage, which will be relatively contextual and atemporal to specify the people to whom the pronouns refer, taking into account (1) the dynamic nature of those pronouns (temporalizing them to account for various relevant slices of time, using the preexistence and nonpolity borders discussed above), (2) the quantity, quality, and relative positions of the people in those pronouns, and (3) the relation of those people to their physical environment or relative states of nonpolity. Requiring temporalizing and specification should (1) require the

follow a norm that is consonant with the declaration and as such should reconstitute human-rights-based democracies, a third form of social organization, beyond families and the state, that Thant may not have been able to envision.

[76] *See* Tehran Proclamation; *See* Carter Dillard, *Valuing Having Children*, 12 J. L. & Fam. Stud. 151 (2010) (arguing that rather than procreative autonomy, the objective value at the base of the right to procreate is the values of continuing one's life, but making it better). Improving continuity, as a value, structures the relationship between constituting, eugenics, and what might be called "euphenics" in a neologism. In short, the three overlap.

[77] Despite the political framing of the model, this value is highly material because humans use objects to orient their relations with one another. *See* Julien, *supra* note 23; *see also* Dillard, *supra* note 22. The unique position of homeless persons is a poignant example of this. We identify, categorize, and treat them differently based on the specific form of property they lack.

[78] The model, especially through the value of one's individual sovereign role, implies a fundamental human right to be heard and to matter and a correlative duty to listen. Ideally our dominant norms, or laws, are normative because they reflect our input and, as such, our consent.

[79] *See, e.g.*, https://havingkids.org/what-are-pregnant-pronouns-the-key-to-a-better-future-and-consensual-and-legitimate-governance/.

claimant to move toward the zero-baseline model in the specification or the contextualizing of claims (contravening commitments to things like procreative autonomy[80] or representative democracy) and (2) insisting the claimant prioritize applying the model before making his or her initial claim as a necessary condition and a more effective means of achieving his or her outcomes.

After defining the pregnant pronouns, we can begin to sharpen the impact of specification and measure the loss of political autonomy by examining the pregnant pronouns for what might be called the "lesser-power asymmetry". In the ascertainable set of persons to whom the pregnant pronoun is referred, can we identify persons that the claimant, after careful consideration of commitments might be making in their claim or elsewhere, be required to exclude from the set of persons to which they are referring? Through the discourse and by examining the claimant's own behavior relative to the eight values in how she or he constitutes and reconstitutes her or his various social groups, are there people the claimant would not reasonably trust with a particular power, duty, privilege, etc., upon which their claim relies and that is implied in it?[81] Moreover, is that particular power, duty, or privilege one that is lesser, included in the set of powers, duties, privileges, etc., in those granted by membership in the human-rights-based democracy? If we do not trust our fellow citizens with custody of our children, why do we elect them in the first election of citizenry,[82] i.e., trust them in the democratic rule-making process?

In other words, by generalizing about the actual persons in our normative claims, we allow ourselves to ignore the actual implications of their quantity, qualities, relative positioning, and positioning relative to their ecology. Pushed to actually specify, we will find the pregnant pronouns are often the key variable in our claims and are not who we need them to be. Moreover, when we look at the asymmetry of how we would exclude people from lesser-included powers that those excluded people should have as members of whatever democracies we share, we find disturbing things about our own place in those democracies.

How can we account for that asymmetry? One answer could be that the human-rights-based democracy either excludes all members equally from their sovereign role, in a variety of ways (crowding out voices and the value of each vote, the use of representatives, the use of procedural limitations, etc.) or does so in ways that exclude some more than others. In other words, we tolerate sharing our democracies with people whom we don't really trust, because neither they, nor we, play a role in the

[80] *See* Sarah Conly, *supra* note 18.
[81] *See* Wesley N. Hohfeld, *Some Fundamental Legal Conceptions as Applied in Judicial Reasoning*, 23 YALE L. J. 16 (1913).
[82] *See* Dillard, *supra*, note 19.

democracy, or because we believe we can exclude those persons from any role while maintaining our own. The presence of those persons who we would exclude from the lesser rights but not greater rights drives home the point: we disempower or are disempowered by others.[83]

The pregnant pronouns and lesser-power asymmetry, as well as concepts like constituting, integrating, and constitutive identities help overcome the cognitive dissonances and other limitations on how we perceive and process information by pushing us toward what might be called a "liberal" or "constituting" perspective. That perspective simply requires accounting for the people who influence and thus have power over your life. But unlike Rawls' model, wherein a particular perspective or sense of justice was a precursor of being reasonable,[84] this constituting perspective comprehensively accounts for how those people, and hence their forms of organization and normativity, came to be and are coming to be. The lesser-power asymmetry forces us away from seeing future children as economic persons,[85] whose creation is geared around filling shopping malls, toward whom it would be necessary for the functional town hall that precedes and regulates the shopping mall. Contrary to Rawls, from this perspective, future generations are not some amorphous entity, they are instead a vulnerable class of people to be protected by the sort of concrete policies and praxis described below.

Five additional aspects of the model are worth noting.

First, the values and dynamics addressed by the model apply to any social grouping and system of norms. We should resist the tendency to think in the framework of existing nation states, though that framework will be useful for the praxis and policies that will implement the model.

Second, the zero-baseline model, or constituting and integrating *Grundnorm*, requires the explicit decentralizing of power—in its many forms—into future generations. For instance, the California Legislature could enact statutes that require persons with an amount of wealth above a certain threshold to pay the State, in the form of a tax, money that will then be used for "baby bonds", in which low-income babies receive a set amount at birth and more each year. Such "baby bonds" would transfer power (in the form of wealth) to those born without power. Under this scheme, existing

[83] Obviously, this is a way of demonstrating the differences in the world today between pre-constitutional systems and an ideal human-rights-based democracy in which the consensual addition of new members would be the only thing that reduces political autonomy.

[84] *See* Friedman, *supra* note 52, in Autonomy, Gender, Politics, ch. 8 (2003).

[85] *See generally* Having Kids, letter, https://havingkids.org/wp-content/uploads/2019/07/Wajahat-Ali-letter-final.pdf.

power is decentralized into future generations and allows the entry (via birth) of a baby in a low-resource family to meet the consent requirement for all relevant entities.[86]

Third, although the model may look minimal, it is merely initial. It is limited to the distinct work of accounting for constitutive identities, rather than accounting for what they subsequently do.

Fourth, zero-baseline modeling and the internationally coordinated intensive family-planning interventions it requires was not feasible until recently and certainly was not at the mid-millennial advent and early development of liberalism. Humans simply did not have the international law,[87] reproductive technology,[88] knowledge of child development, rise of gender equity, etc., necessary for this modeling. But there are many pressing consequential reasons (beyond the sufficient reason of legitimating our human-rights-based democracies with a *Grundnorm*), including climate change, rising inequality, the need to meet sustainable-development goals, rising automation of labor, the rapid development of gene-editing technology—for why we should implement this model today.

Last, from the constitutive perspective, hierarchy, in any form, including representative democracy or celebritism,[89] may be accounted for as the pre-constitutional and de-constituting normative push in directions away from the zero-baseline model and its values. Social hierarchy is the institution of the chaos and cacophony that invites top-down solutions and orientation around objects rather than interpersonal relations between free and equal subjects.[90] The hierarchists will use de-constitution to make orienting from the liberal or constitutional perspective infeasible, while also propagating Hart's illusion of autonomy, because it supports their

[86] *See* "The Right to Have Children", *Having Kids* (blog), available at https://havingkids.org/wp-content/uploads/2018/06/The-Right-to-Have-Children-What-is-the-best-account-and-how-do-we-implement-it-in-law-and-practice.pdf.

[87] Consider how foreign threats have justified things like pro-natalism and other degradations of democracy. https://www.independent.org/pdf/tir/tir_20_02_03_howden.pdf. Pre-constitutionalism abroad, as an existential threat, justifies pre-constitutionalism at home. But what if we could apply the model universally, targeting pre-constitutionalism everywhere, in ways the world could not in the early days of liberalism?

[88] *See generally* https://www.nature.com/articles/d41586-019-01906-z. Could we realistically control things like human gene editing in a top-down fashion, or does the history of things like the enforcement of abortion prohibitions show the need for a truly integrated norm?

[89] "Celebritism" refers to the political commitment of celebrities. Lena Partzsch, *The Power of Celebrities in Global Politics,* 6 Celebrity Stud. 178 (2015).

[90] *See* Julien, *supra* note 23.

hierarchy, or what some call a population pyramid Ponzi scheme.[91] As a result, subjects become lost in a sea of people.[92]

The lived experience of political obligation, autonomy and a reconciliation of seemingly conflicting fundamental modes of positive communitarian freedom and negative liberal freedom could come in the dynamism of the values that compose the zero-baseline model. The model is essentially emancipatory, aligning the positive freedom to and the negative freedom from with the emancipation that comes from a spatial-temporal model that maximizes universal consent.

To illustrate, we can apply the thresholds described above to a simple hypothetical of the scenario that must precede all liberal collective action (including the decision to minimize the collective and separate off into a free-market or comparable system): A town hall meeting in which members of the town must agree on a plan.

In this scenario, each additional member that joins the town hall meeting excludes other members, e.g., through time at the podium, of the possibility that the plan will reflect the will of another. That inevitable exclusion is reduced or exacerbated contingent upon the constitutive qualities of each member and whether they are reasonable, e.g., respecting limits on their time at the podium, calling for an unreasonable plan, etc.[93] Whether the town hall meeting actually involves free and equal people engaged in collectively making the rules under which they will live, depends on the relative positioning of the members to each other and the power differentials that existed before they entered the meeting. It also depends on the relation of the members to the natural ecology surrounding the process (e.g., are members preoccupied with some existential ecological crisis?). As the town grows, a negative feedback loop (a self-exacerbating collective-action problem) develops, with each member having less of a role and less incentive to have a role as the learned helplessness of being lost in the crowd sets in.[94] The speed with which that loop develops depends on the quantity, qualities, relative positioning of the members to each

[91] *See* Jeff McMahon, *The World Economy Is A Pyramid Scheme, Steven Chu Says*, Forbes, April 5, 2019, available at https://www.forbes.com/sites/jeffmcmahon/2019/04/05/the-world-economy-is-a-pyramid-scheme-steven-chu-says. *See also*, Santosky v. Kramer, 455 U.S. 745, 790 (1982) (Rehnquist, J., dissenting) ("Few could doubt that the most valuable resource of a self-governing society is its population of children".); Brown v. Bd. of Educ., 347 U.S. 483, 493 (1954) (referring to education as "perhaps the most important function of state and local governments"); Prince v. Massachusetts, 321 U.S. 158, 168 (1944) ("[D]emocratic society rests, for its continuance, upon the healthy, well rounded growth of young people into full maturity as citizens, with all that implies".).

[92] *See* https://havingkids.org/wp-content/uploads/2019/07/Wajahat-Ali-letter-final.pdf.

[93] The Supreme Court has addressed "constituting" in a sense in Fiallo v. Bell, 430 U.S. 787, 792 (1977) ("[O]ver no conceivable subject is the legislative power of Congress more complete than it is over the admission of aliens".) (quoting Oceanic Navigation Co. v. Stranahan, 214 U.S. 320, 339 (1909)).

[94] *See* Ilya Somin, Democracy and Political Ignorance: Why Smaller Government Is Smarter (Stanford University Press, 2016).

other and their ecology of the members, i.e., whether the group is relatively constituting or de-constituting before they ever even begin the process of collective rulemaking. All of this becomes more complex as we account for intergenerationality.

Minimizing the loss of social self-determination in the planning process requires that each new entrant be of a minimum constitutive quality. It requires there be a maximum number of entrants. It requires they be arranged relative to each other and their ecology in particular ways.

The model leads to a specific confluence of values. For example, the autonomy would-be parents gain through self-development and meeting a standard of readiness to parent, aligns with objective, constitutive and naturalistic reasons for having children, the diminishing interest in having additional children,[95] and the quality–quantity tradeoff of a smaller family. These things align with the autonomy of a highly developed child and the morally valuable options and equality of opportunities in life they will have, which further aligns with the autonomy created by gender equity, cooperativeness within groups like our town hall scenario above, a level of inherent security in groups that avoids the need for top-down limitations on autonomy, as well as fluidity among groups.

The resulting smaller populations and high levels of development align with one's having a meaningful voice in public affairs and the rules under which one must live. This participatory agency aligns with the autonomy created by fulfilling one's need for meaningful group membership, which in turn aligns with reduced consumption and property-oriented around nonpolity, rather than group belonging or power and ranking over others. One's consumption and property orient around nonpolity instead of one using those things to relate to other people. And that re-orientation supports the decentralization of concentrations of power into future generations, discussed below.

And all of these align with a healthy and safe environment and the freedom—or autonomy—from others that is only possible through interaction with the non-human world. This in turn aligns with the autonomy the animal liberation movement advocates for.[96] Note that this alignment spans many types of freedom, from communitarian to libertarian, that might be seen as conflicting, were we not temporalizing our perspective.

The model is unique, particularly in one important regard. The thresholds above can be set using existing legal standards we claim to already adhere to, like basic parental fitness, the Children's Rights Convention, biodiversity and wilderness restoration targets, education benchmarks, federalism and representative ratios, redistribution

[95] *See* Dillard, *supra* note 76.
[96] *See* Dillard, *supra* note 11.

policies that attempt to create equality of opportunity, etc. The model will show that, by our own standards, we are not who we should be or who we claim to be.[97]

Also, because the schema above represents the standard for political obligation and autonomy by maximizing consent, it also represents the *Grundnorm* for truly integrating (by their own true will) members into a legality. Hence, it serves as a human-rights-based criterion for evaluating the legitimacy of legal systems.

In summary, how can human-rights-based democracies solve the democratic dilemma of making good choices while being integrative and inclusive?[98] How do we ensure freedom, or the balance between community and autonomy, inherent in concepts like unified independence and ordered liberty?[99] We must understand that we cannot include persons in any type of grouping without excluding persons, in some way, and as such, constitute ourselves in quality, quantity, relative arrangement, and, relative to our ecology, to maximize consent and obligation and thereby minimize the loss of political autonomy. Incontrast, we ensure that we never meet the political ideals of liberalism by ignoring procreation, seeing birth positioning as a matter of random fortune, and pretending the people with whom we share the world just fell from the sky (or were delivered by a stork).

A comprehensively normative system[100] requires that we create normative people. When people say "there ought to be a law", there should be an antecedent norm that makes the statement meaningful. At its base, democracy involves the capacity to influence, which in turn begins with the existence of the person. If we take democracy (or individual sovereignty, from the appropriate and more compelling individualized perspective) seriously, we must focus on the creation of individuals and thereby unravel the age-old legal conundrums of balancing community and autonomy, unified independence, and ordered liberty.[101]

[97] *See* Rawls, *supra* note 49. A rational person could not honestly conclude that their being wealthy is a good reason to accept a system that only favors the wealthy.

[98] *See* Robert A. Dahl, *A Democratic Dilemma: System Effectiveness Versus Citizen Participation*, 109 Pol. Sci. Q. 23 (1994). https://www.jstor.org/stable/2151659?seq=1#page_scan_tab_contents, accessed August 15, 2019.

[99] *See e.g.*, Planned Parenthood v. Casey, 505 U.S. 833 (1992); *see also* Washington *v.* Glucksberg, 521 U.S. 702 (1997).

[100] *See* Ronald Dworkin, Law's Empire (1986).

[101] *See* Washington v. Glucksberg, 521 U.S. 702 (1997).

THE PRAXIS OF CONSTITUTING

Modifying political obligation and autonomy models is complex. For that reason, and because the new modeling requires the cooperation of prospective parents across the globe, this article provides a useful heuristic throughout: Remodeling family planning means changing from a parent-focused and subjective family planning model to a child-focused and objective "Fair Start" family planning model,[102] as well as specific actions to ensure that change, messaged through an operational non-profit that has begun the discourse, Having Kids.[103] A simple test for the change in perspective is to ask whether we can seed resource transfers through universal basic income laws that provide resource-poor children with money that has been collected in the form of taxes paid by the wealthy, as cooperative entitlements owned by future children, or whether those resource transfers are seen as ethically suspect incentives aimed at parents.

In essence, our reliance on future generations as economic inputs and non-humans and their habitats as resources, keeps us in a state of pre-constitutionality. We can invert the population pyramid to change this and target those with the most resources in order to fund the transition.[104]

The Fair Start model (which is the public branding for the zero-baseline model) requires (1) parental readiness before procreating; (2) a transfer of resources to would-be parents designed to decentralize influence *into* future generations, using the heuristic of "giving each child a fair start in life"; and (3) a norm, for all, of smaller families who can invest more in each child.[105] By focusing on the objective needs of future children, parents and communities are forced to abandon the myth of procreative autonomy. They must replace it with objective reasons for having children, which must align with other aspects of the model to maximize consent and thereby maximize political obligation and autonomy. Practically speaking, Fair Start modeling involves (1) reforming the interpretation and practice of the fundamental right to have children as it exists in the Universal Declaration of Human Rights and implementing conventions[106] and (2) intensive family-planning incentives (which are actually entitlements held by the

[102] The model, and the ability to operate a child-first normative framework, would be a useful litmus test for whether a person is acting with the minimum level of empathy and helping behavior, or constitutive quality, necessary to constitute.

[103] https://havingkids.org/. The use of a possessive verb with the creation norm is inconsistent, but the organization adopts common parlance for pragmatic reasons.

[104] *See* https://havingkids.org/protect-future-ivanka-trump-marco-rubio/; *See also* and https://www.vox.com/future-perfect/2019/12/16/21024222/mitt-romney-michael-bennet-basic-income-kids-child-allowance.

[105] *See* Doepke, *supra* note 40.

[106] *See* Dillard, *supra* note 48.

future child),[107] administered by the U.N.F.P.A. and U.N.I.C.E.F. using the Children's Rights Convention as one initial baseline.[108]

The zero-baseline model accounts for the loss of autonomy and the liberating confluence of smaller families that develop each child, ensuring children play a meaningful role in their democracies, and lower population variants that match the natural ecologies in which democracy was born and that best promote human flourishing. Sovereignty requires that democracies be small, dynamically forming and reforming, and populated by highly developed people. This is consistent with the healthy, safe, and biodiverse ecologies in which they exist. In other words, the values above align under the zero-baseline model as iterations of human autonomy or freedom and enable a dynamic of group formation and reformation and of balancing values while minimizing the loss of autonomy. This view of democracy, focused on the people that actually compose it, looks nothing like modern, static, and bloated nation states.

There are reasons to consider the zero-baseline model as the dominant peremptory norm, including the need to preserve the model simply as a hedge against the harm being done to the values today (e.g., the destruction of the non-human world), as well as the efficacy of the model in furthering agreed-upon outcomes like the Paris Accord and the Sustainable Development Goals.[109] The zero-baseline model proves to be ten to twenty times more effective at mitigating climate change, improving child welfare and equity, building democracy, reaching other development goals—and avoiding unacceptable ecological and social risks—than any downstream efforts.[110]

But one reason that the model becomes the dominant peremptory norm, overriding all conflicting interests, including property rights claims against the demand for resource transfers outlined above, stands out in particular. Because we are before we do, the creation norm is lexically prior, and therefore overrides conflicting norms, including public and private property rights that might be asserted against the transfer of resources described here. There can be no justification for the inequity and centralization of power in some extant persons, relative to the creation-norm based claims of incoming children, and Murphy provides a useful analogy: They challenge the

[107] *See* The Fair Start Model of Family Planning, Having Kids website, at https://havingkids.org/wp-content/uploads/2019/05/FairStartModel2019-1.pdf (last visited Aug. 15, 2019).

[108] *See* Ashley Berke, "Dear UN: We Can Do Better for Kids", *Having Kids* (blog), July 31, 2019, 2019, at https://havingkids.org/uncandobetter/ (last visited Aug. 15, 2019).

[109] *See, e.g.*,"Paris Agreement", *Wikipedia*, https://en.m.wikipedia.org/wiki/Paris_Agreement (last visited Aug. 15, 2019); *see also* https://sustainabledevelopment.un.org/?menu=1300; https://havingkids.org/fair-start-research/.

[110] *See* Hamity at al., *supra* note 4; The Fair Start Model of Family Planning, Having Kids website at https://havingkids.org/wp-content/uploads/2019/05/FairStartModel2019-1.pdf (last visited Aug. 15, 2019).

assumption that we own property in a meaningful way that can be separated from taxes and other public obligations we owed relative to that property.[111] But have they in turn assumed we are properly constituted in any meaningful way and can assess public and political obligations without first accounting for the borders of human power, without a way for electing one another, without a way to actually emancipate citizens? Is there an antecedent myth of a human-rights-based democracy that legitimates downstream norms without a truly constituting and legitimating *Grundnorm*?

The overriding nature of the zero-baseline model means that any person may further the *Grundnorm* by any means effective, regardless of any contravening norms. Constituting is analogous to self-defense against the aggressive and non-consensual invasion of our world by others. This means securing guaranteed minimum incomes for children (or "baby bonds") pegged to family-planning reform (including through things like zero-interest family-planning loans),[112] enabling legislation for no-procreation probation and parole orders for abusive and neglectful parents, expediting approval of long-acting male contraceptives for market, incorporating a "smaller family" policy into UN sustainable-development goals, and undertaking a variety of other measures.[113]

A key question remains as to how to choose which concentrations of power, both public and private,[114] to prioritize the decentralization required for universal basic income measures for poorer children. The applicable types of concentrations of power include any resources currently used to incentivize large or unprepared families, and the resources of individuals and organizations that pose acute threats to nonpolity through things like the exacerbation of climate change. Those concentrations of power would be targeted for downstream welfarist redistribution, through taxes and other measures. Why these resources? These resources should be targeted because they pose the greatest threats to the legitimate obligation and autonomy the model seeks to promote.

Our framing of the praxis and constitutive discourse employs three different narratives. Fair start family planning and the zero-baseline model ensure the following:

(1) The simultaneous liberation of future children and non-humans (the most vulnerable and numerous moral entities we can conceive) from property hood and constant colonization at the spatial-temporal border discussed above. The praxis of the model targets the hegemony of a small group of living persons over an untold number of vulnerable moral entities.

[111] *See* Murphy & Nagel, *supra* note 29.

[112] See kiva.org as a possible vector. *See also*, Byrd Pinkerton, Jillian Weinberger, and Amy Drozdowska, "Free college tuition helps, but it's not a silver bullet", *Vox* (blog), at https://www.vox.com/2020/2/12/20997880/free-college-tuition-kalamazoo-promise-the-impact.

[113] *See generally* https://havingkids.org/.

[114] This distinction conflicts with the model but is used here for the sake of ease.

(2) The unification of child welfare, environmental and animal protection, and human rights and democracy movements around this new modeling, given the model's comprehensive and superior efficacy.[115]

(3) The manifestation of a real (spatial-temporal) social contract that can be used as a model agreement between would-be parents and concentrations of power that satisfies disparate and competing political perspectives and excludes unreasonable or pre-constitutional people.

Given climate change, the need for an immediate, conscious, and accelerated move toward United Nations low-variant world-population projections, sustainable-development goals, and the meeting of the obligations of the Children's Rights Convention,[116] all facilitated by the specific wealth transfers and new family-planning incentives (which are actually entitlements held by future children) that will be most effective at minimizing human suffering in the future.[117]

Part of the praxis involves two actual social justice initiatives: The first is a lawsuit recently filed against the United States for how its actions and inactions regarding climate change have violated the constitutional right to privacy in wilderness (nonpolity), by degrading conditions in that wilderness,[118] and the second is a legislative initiative that would authorize family courts to impose "no procreation" orders against persons convicted of felony child mistreatment.[119] Together, these concrete actions begin the tasks of redefining political autonomy and obligation by illuminating the bookends of political obligation and autonomy: the preexistence of future generations and nonpolity. These bookends apply the zero-baseline model to initiate the constitutive discourse through a simple question:

How can freedom deprive children of minimum levels of welfare at entry, e.g., the right of parents to create additional children who will be kept and raised in the horrific conditions of state custody because the parents previously abused or neglected their first child, but not the right to simply be left alone—free from the life-threatening impacts of climate change—in the solitude of wilderness? Such a state of affairs is not freedom; it is pitting extant and future people against each other for the benefit of elites and others

[115] "The Fair Start Model of Family Planning", Having Kids, https://havingkids.org/wp-content/uploads/2019/05/FairStartModel2019-1.pdf (last visited Aug. 15, 2019).
[116] *See id.*
[117] *See* Hamity at al., *supra* note 4.
[118] *See* Karen Savage, *Judge Dismisses "Right to Wilderness" Climate Suit Against U.S. Government*, Climate Liability News, August 1, 2019, https://www.climateliabilitynews.org/2019/08/01/right-to-wilderness-climate-lawsuit/ (last visited Aug. 15, 2019).
[119] *See* Ashley Berke, "Ask Florida Senator Lauren Book to Prevent Child Abuse With SB90", *Having Kids* blog, March 14, 2019, https://havingkids.org/florida-senator-lauren-book-prevent-child-abuse-sb-90/ (last visited Aug. 15, 2019).

above them in the social hierarchy. It is using people to disempower one another and ignoring the simple truth that we need to create free people before they even begin to socially self-determine and regulate their own affairs.

This discourse requires people to de-abstract their normative claims and account for the quantity, quality, relativity, and ecological positioning of all of the people upon which their claims rely and the values at stake and dynamics at play. This would in turn push people toward accepting constituting (which is essentially a thick conception of a constitution) as a fundamental normative framework. Constitutive discourse, the discourse around the values and process of the model, assesses the relative illegitimacy of systems of social organization, combining the subjective exercise of all forms of human influence with the objectivity of the model and its irrefutable basic values, starting from and ending with the touchstone of nonpolity.

Quite simply, we can only reduce the age-old tensions between the individual and the collective and between freedom and security and thereby further human autonomy by creating sufficiently capable people (in terms of quantity, constitutive qualities, relative positioning, and ecological positioning) who may reconstitute their legalities at will.[120] Families cannot create that capability in isolation.[121] We create capable people by overcoming the cognitive dissonances and misperceptions we carry regarding the nature of human autonomy and decentralizing concentrations of human power (governmental, corporate, oligarchical, prospective parental, etc.) into future generations through zero-baseline modeling and its layperson translation, Fair Start family planning. Because, as a preremptory norm, the model overrides all conflicting norms, all may engage in its praxis by any means effective. And because of the nature of the model, those who oppose it can be classified, treated as pre-constitutional, and as a threat to the autonomy of the model would further.

A simple analogy: a river with future children on one side and would-be parents on the other. The parents are pulling the kids across, acting in isolation from one another. As such, some kids cross in turbulent parts of the river and are washed downstream, or come across harmed for life, while others wade through ankle-deep parts with all the resources they need. As the mass pulling-across happens, the ecology of the river is trampled and destroyed. The process has resulted in our side of the river looking like chaos, inequity, and ecological collapse. Our challenge is to build a bridge in the middle of the river that

[120] How close a particular polity is to being able to subdivide and reconstitute into legitimate legalities could be assessed, initially, by examining how the lesser power asymmetry applies there.

[121] *See* "Marshmallow Test" Redux: New Research Reveals Children Show Better Self-Control When They Depend on Each Other", Association for Psychological Science website, https://www.psychologicalscience.org/news/releases/marshmallow-test-redux-new-research-reveals-children-show-better-self-control-when-they-depend-on-each-other.html.

can bring the children over in a way that emancipates them and ensures they are agents that can consent to the power of others. Is that simple ask—to work together to give all kids a fair start in life—something around which we can build revolutionary action?[122]

CONCLUSION

The zero-baseline model derives from irrefutable truths about how freedom requires limiting the total amount of human power (or "influls") present in the world, inherent limits on human cognition and processes for determining rules, minimum levels of human development to determine those rules, etc. The model provides a liberal or constitutive perspective and tool for a specific spatial-temporal cognition to help overcome the psychological limitations discussed above, those that hide the inevitable first behaviors and dynamics that promote or degrade human autonomy. This cognition goes well beyond Rawls' prerequisite "sense of justice" and looks to actually exclude pre-constitutional people. This modeling could prove the key to solving many of today's social and ecological crises through a concrete praxis that is emancipatory, effective, and designed to appeal across the political spectrum.

Again, the only way to be free from human power is to consent to it, via norms. That requires accounting for certain first and inevitable dynamics in the process and reflecting that accounting in a first norm, or *Grundnorm*, that should constitute us.[123] There is no person, no realm of human experience, including freedom, outside of the realm of a creation norm. The zero-baseline model is a candidate for this first norm and, hence, a necessary condition of political obligation and autonomy. If we want free people, we have to secure the resources necessary to do so. They will not fall from the sky.

[122] *See generally* Richard A. Posner, *The Problematics of Moral and Legal Theory*, 111 Harv. L. Rev. 1637 (1997).

[123] This interpretation of legal systems, which follow the *Grundnorm*, would comprise a fourth version of the conception of "law as a seamless web". see Lawrence Solum, "Legal Theory Lexicon: The Law Is a Seamless Web", *Legal Theory Blog* (blog), October 1, 2006, https://lsolum.typepad.com/legaltheory/2006/10/legal_theory_le.html.

VOLUME 6 ISSUE 1 ARTICLES & ESSAYS

UNIVERSITY OF BOLOGNA LAW REVIEW

https://doi.org/10.6092/issn.2531-6133/13410

Received: 31 Aug. 2020 | Accepted: 17 Nov. 2020 | Published: 4 Aug. 2021

Is a Requirement to Wear a Mask Economically Valid During COVID-19?

STEVE G. PARSONS

Steve Parsons is a partially-retired consulting economist and former adjunct faculty member at Washington University in St. Louis. (United States). Thanks are due to Philip Liang for research assistance. Thanks, in particular, are due to Carly Gibbons for research assistance and useful comments.

@ Steve@ParsonsEcon.com
0000-0002-3674-5050

ABSTRACT

Two of the most important categories of government intervention in response to COVID-19 are business closures and mask mandates. The scientific literature supports the efficacy of mask-wearing to reduce the transmission of respiratory viruses (including COVID-19). However, the efficacy is greater in stopping outbound transmission (meaning that my mask protects you) than inbound transmission (meaning that my mask protects me). Evidence suggests that the full benefits to society of wearing masks are far greater than the full costs to society of wearing masks. The author argues that mask-wearing is far more cost effective than business closures in controlling the spread of COVID-19. Moreover, the author argues that highly infectious diseases have an externality dimension. The person infected with COVID-19 makes a decision regarding whether to wear a mask based on their own perceived costs and benefits of mask-wearing, but that decision has consequences for those they come in contact with: the infected person's decision not to wear a mask imposes costs on others that are external to the infected person's decision process not to do so. The author further argues that some possible methods by which to deal with such an external cost (individual negotiations, a tax on spreading COVID-19, or as subsidy for wearing masks) are impractical. This makes a mask-wearing government mandate economically valid.

KEYWORDS

Covid-19; Externalities; Mask; Business Closure; Stay-At-Home Order

JEL CODES

I28; I12

TABLE OF CONTENTS

INTRODUCTION

In December of 2019, an outbreak of coronavirus [hereinafter COVID-19] was detected in mainland China. The outbreak was caused by a new virus, technically known as the "severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2)." On February 12th, the World Health Organization named the disease caused by the novel coronavirus "Coronavirus Disease 2019" (COVID-19).[1] As of the date of this writing, COVID-19 has killed over 1.2 million people worldwide with nearly 240,000 of those deaths occurring in the United States.[2] Consider the two primary categories of government intervention in the United States in response to COVID-19: 1) business closures and stay-at-home orders; and 2) mandatory mask-wearing. By April 6, 2020, forty-three States had issued stay-at-home orders.[3] Moreover, virtually every State had some business closures, such as in the case of gyms, sporting venues, bars and indoor dining. In contrast, by April 6, 2020 (that same point in time), only seven States had mandated masks in public.[4] However, this has changed over time. By June 3, 2020 "[a]ll 50 US states [had] loosened restrictions put in place earlier in the pandemic, allowing some businesses to reopen."[5] Moreover, as of July 17, 2020, twenty-eight States had mandatory mask orders[6] and the

[1] World Health Organization, *Naming the Coronavirus Disease (COVID-19) and the Virus That Causes It*, WHO (last accessed Nov. 8, 2020), https://www.who.int/emergencies/diseases/novel-coronavirus-2019/technical-guidance/naming-the-coronavirus-disease-(covid-2019)-and-the-virus-that-causes-it.

[2] The New York Times, *COVID in the U.S.: Latest Map and Case Count*, NYT (Nov. 8, 2020), https://www.nytimes.com/interactive/2020/us/coronavirus-us-cases.html.

[3] *See* Jason Silverstein, *43 States Now Have Stay-at-Home Orders for Coronavirus. These Are The 7 That Don't*, CBS News (Apr. 6, 2020), https://www.cbsnews.com/news/stay-at-home-orders-states/.

[4] *See* Scottie Andrew & James Frio, *These Are the States That Require You to Wear a Face Mask in Public*, CNN (Apr. 20, 2020), https://www.cnn.com/2020/04/20/us/states-that-require-masks-trnd/index.html.

[5] Holly Secon, *An Interactive Map of the US Cities and States Still under Lockdown – and Those That are Reopening*, Business Insider (June 3, 2020), https://www.businessinsider.com/us-map-stay-at-home-orders-lockdowns-2020-3.

[6] *See* Arielle Mitropoulos, *28 States, Washington D.C., and Puerto Rico Have Issued Mask Mandates to Prevent Spread of COVID-19*, ABC News (July 17, 2020),https://abcnews.go.com/US/28-states-washington-dc-puerto-rico-issued-mask/story?id=71842266; Grace Hauck, *What States Require Face Masks in Public? Alabama, Arkansas, Colorado Join Growing List of States Where It's Mandatory*, USA Today (July 3, 2020), https://www.usatoday.com/story/news/health/2020/07/03/covid-face-masks-states-require-public/5371503002/.

nine largest brick and mortar retailers (including Walmart, Home Depot, and Lowes) now require shoppers to wear a mask.[7]

In this paper, I consider the possible economic rationales for governments to intervene in markets or personal affairs. I do this by introducing the economic concept of an externality (where one person's actions affect other people) first in general, then later, specifically, in the context of a mask mandate by the government. I briefly examine the scientific research on the effectiveness of masks, both with respect to air-borne infectious diseases in general and the evidence with respect to COVID-19.

1. THE ECONOMICS OF EXTERNALITIES

Like most economists, I prefer limited government interference in markets, business, and personal affairs. Economics has a short list of potentially-valid rationales for government intervention.[8] One of these potentially-valid rationales is an "externality," which is usually in the form of an external cost (sometimes called an external diseconomy). An external cost "occurs when an action by a firm or individual results in uncompensated costs or harm to others."[9] Perhaps the most commonly discussed external cost is pollution. For example, a steel company makes decisions regarding what technology to employ and how much steel to produce based upon the private costs it incurs vis-à-vis the private benefits it receives (revenues from the sale of the steel). What is not included in (or, in other words, is external to) the steel company's decision are the effects of sulfur dioxide, for example, that the steel plant emits, which can trigger respiratory illnesses for people near the plant.

[7] *See* Antonia Noori Farzan et al., *Top Nine Brick-and-Mortar Retailers Now Require Coronavirus Masks in U.S. Stores*, Washington Post (July 17, 2020), https://www.washingtonpost.com/nation/2020/07/17/coronavirus-live-updates-us/.

[8] Francis M. Bator, *The Anatomy of Market Failure*, 72 Q. J. Econ. 351, 357 (1958). (describing only three categories of market failure: monopoly, public goods, and externalities.) Moreover, one can argue that public goods, and even monopolies, are a form of externality. Carl J. Dahlman, *The Problem of Externality*, 22 J. L. & Econ 141, 142 (1979). If one accepts this perspective, externality is the central cause of market failure and the prime potential rationale for government intervention.

[9] W. Bruce Allen et al., Managerial Economics: Theory, Applications, and Cases 707 (8th ed. 2021). Often, discussions of external costs focus on a price for a market transaction (such as the price of steel when pollution occurs). However, a market transaction is not required for an externality as the quoted definition allows. *See also, e.g.*, James M. Buchanan & Wm. Craig Stubblebine, *Externality*, 29 Economica 371 (1962). The authors do not focus on market transactions but rather on "activities." Traditionally, the entities would be described as "economic agents," which are most commonly consumers and producers. *See* Gary Becker, Economic Theory 84-7 (1971). Hal R. Varian, Intermediate Microeconomics: A Modern Approach 569–89 (5th ed. 1999). Here, however, I will largely focus on externality effects of decisions outside of a typical market transaction.

These additional "social costs" are external to the steel company's decisions involving which technology to employ and how much steel to produce. The full welfare efficient or socially efficient results occur where the full social marginal costs (including the external costs) equal the full social marginal benefits (including any external benefits).[10]

It is noteworthy that the simple existence of an externality is insufficient to justify government intervention. Under certain circumstances, societal members can resolve (or reduce the impact of) the externality themselves by negotiating a result that is socially efficient. The key to such a result are well-established property rights and low transaction costs.[11] For example, if a cattle rancher and a wheat farmer exist on the same island, transaction costs are likely low enough that a solution to the problem of cattle wandering into the wheat field can be resolved by negotiation between the parties.[12]

Free markets (markets that exist without government interference) may try to reduce these externality problems via: a) voluntary clubs, consortia, or unions to jointly negotiate;[13] b) charities or major philanthropists; c) the collection and dissemination of information;[14] or d) lawsuits (including class action lawsuits).[15]

[10] *See* virtually any textbook on welfare economics. For the earliest complete treatment on the topic, *see generally* Arthur Cecil Pigou, THE ECONOMICS OF WELFARE (1920). For an intuitive treatment, *see* Tejvan Pettinger, *Social Efficiency*, Economics Help (Sept. 17, 2019), https://www.economicshelp.org/blog/2393/economics/social-efficiency. This concept is similar to "pareto efficiency," named after the Italian economist, Alfred Pareto. To be pareto optimal, there can be no change by which one agent can be better off without making other agents worse off.

[11] Nobel Laureate Sir Ronald Coase was an early pioneer in considering property rights and transactions costs (including the costs to collect information, negotiate and monitor contracts). *See* Ronald H. Coase, *The Nature of the Firm*, 4 ECONOMICA 386 (1937).; Ronald H. Coase, *The Problem of Social Cost*, 3 J. L. & ECON., Oct. 1960, at 1.

[12] Some of the solutions to this problem include building a fence, a switch to all farming, or a switch to all cattle ranching.

[13] Examples include environmental groups like the Sierra Club. For a list of such firms, *see e.g.*, Wikipedia, *List of Environmental Organizations*, Wikipedia (Oct. 12. 2020), https://en.wikipedia.org/wiki/List_of_environmental_organizations. One of the techniques such clubs may utilize is to file lawsuits or to pay for (or partially pay for) class action law suits that a law firm may find insufficiently profitable to accept on a standard contingency basis. *See, e.g.*, Mark Harcourt et al., *The Role of Unions in Addressing Behavioural Market Failures*, Econ. & Industrial Dem. (June 2019), https://doi.org/10.1177/0143831X19853027.

[14] This can be done via a charity, a club, or a government agency. I believe most people, and most economists, find such action less intrusive than other government action.

[15] A class action lawsuit is more feasible where there is a single firm or entity that has caused the damages, even when there are many damaged parties (often consumers of a specific product). Multiple potentially damaging firms make assessing and proving liability more difficult. Robert D. Cooper & Ariel Porat, *Liability Externalities and Mandatory Choices: Should Doctors Pay Less?,* 1 J. TORT L. 1, 5 (2006).

However, the larger the number of relevant entities:[16] 1) the more difficult it is to assess liability; 2) the more likely it is that there are increases in transaction costs; [17] and 3) the larger the extent of free-rider problems.[18] Therefore, in some circumstances, these free market "solutions" to the externality problem are too costly or insufficient in their effectiveness. In such cases, government intervention may be warranted.

Non-economists may expect that the government should completely stop the activity that creates an external cost (for example, by stopping all pollution, or dictating a technology to reduce pollution). Economics generally suggests that even when government intervention may be warranted, introducing market mechanisms can more efficiently deal with external costs. One such method is a tax on pollution to try reaching a point where the marginal cost of pollution (the marginal external cost on residents) exactly equals the marginal cost of pollution control.[19] The modern practical application of this approach includes "emissions trading" in which a firm that is positioned to more efficiently reduce pollution is encouraged to do so.[20] Generally, providing incentives to push participants toward the socially-optimal result is superior to mandating a solution such as a particular technology for pollution control or mandating that the entire population receive a vaccine.[21] Even in cases in which it appears clearly that government

[16] "Entities" include producers, consumers, or simply people affected by other entities' decisions. *See id.*

[17] *See, e.g.*, William C. Apgar& H. James Brown Microeconomics And Public Policy 252 (1987). (discussing free-market solutions and issues relating to the number of entities and quantity's likely effect on transactions costs).

[18] A "free-rider" is an entity that receives the benefit of an action (such as a club reducing pollution) without contributing to the payment for the action. *See, e.g., id.* at 327. A free-rider problem often occurs when there is a condition of non-excludability; it is not possible to exclude entities from receiving the benefits of the action. *See, e.g.,*Oliver Kim & Mark Walker, *The Free Rider Problem: Experimental Evidence*, 43 PUB. CHOICE 3 (1984). ; Charles R. Plott, *Externalities and Corrective Policies in Experimental Markets*, 93 Econ j. 106 (1983). (stating that "[w]ithin the simple setting explored here, the traditional models found in the economics literature are amazingly accurate."). Non-excludability is one of the two characteristics of a "public good." Often, a public good is supplied below the socially-optimal quantity. Some studies indicate free-riders were less of a problem than logic and economics suggested they should have been. *See, e.g.*, Peter Bohm, *Estimating Demand for Public Goods: An Experiment*, 3 Eur Econ. Rev 111 (1972). Gerald Marwell & Ruth E. Ames, *Economists Free Ride, Does Anyone Else? Experiments on the Provision of Public Goods IV*, 15 J. Pub. Econ. 295 (1981). ; John W. Sweeney Jr., *An Experimental Investigation of the Free-rider Problem*, 2 Soc. Sci. Res. 277 (1973). ; Vernon L. Smith, *An Experimental Comparison of Three Public Good Decision Mechanisms*, 81 Scandinavian J. Econ. 198 (1979). Alison L. Booth, *The Free Rider Problem and a Social Custom Model of Trade Union Membership*, 100 Q. J. Econ. 253 (1985).

[19] *See, e.g.*, Allen et al., *supra* note 9, at 707. *See also* Paul J. Feldstein, Health Policy Issues: An Economic Perspective 370 (3d ed. 2003).(stating that "[i]mposing a tax directly on pollution is preferable to such indirect methods of control pollution."). Many non-economists are surprised that such a "pareto optimal" result will not eliminate pollution. *See* William Baumol, *On Taxation and the Control of Externalities*, 62 Am. Econ. Rev. 307 (1972). ; Varian, *supra* note 9, at 569–589

[20] *See, e.g.*, Frank J. Convery, *Reflections—The Emerging Literature on Emissions Trading in Europe*, 3 Rev. Env't. Econ. & Pol'y 121 (2009). For information on the United States Environmental Protection Agency's emissions trading programs, *see* United States Environmental Protection Agency, *Emissions Trading Resources*, EPA (Apr. 30, 2020), https://www.epa.gov/emissions-trading-resources.

[21] *See* Dagobert L. Brito et al., *Externalities and Compulsory Vaccinations,* 45 J. Pub. Econ. 69 (1991).(arguing that taxes, subsidies, and provisions of information are superior to mandatory vaccinations).

intervention is appropriate, one must also consider the costs of government intervention, which can include the direct costs of the government agencies involved, the costs of entities complying with a government decision, and the potential for unintended consequences.[22] Ultimately, this will involve a comparison of an imperfect market result with an imperfect government intervention. [23] An infectious disease, such as COVID-19, also has an externality dimension: those infected persons who are not careful can impose an external cost on others they infect. The cost to others (from an increased probability of infection and its consequences) is likely external to a person's decisions to wash their hands frequently, socially distance, to avoid large gatherings, and to wear a mask.[24] Alternatively, one can say that the person wearing a mask confers an external benefit (by reducing the chance of infection). Later in this manuscript, I consider in more detail how one might deal with the external costs caused by individuals who decide not to wear masks.

Good economics requires an estimate of the benefits and costs of any governmental intervention.[25] It is difficult to estimate the costs of business closures during the COVID-19 pandemic, in part since some reductions in business activity would have occurred regardless of government mandates.[26] However, given the reduction in Gross Domestic Product [hereinafter G.D.P.],[27] the expenditures by Congress, and actions by the Federal Reserve, the cost is easily trillions of dollars.

[22] For a very intuitive read on the economics of government intervention for the non-economist (but written by an economist), *see generally* THOMAS E. HALL, AFTERMATH: THE UNINTENDED CONSEQUENCES OF PUBLIC POLICIES (2014). For a discussion of the divergence between theoretically-optimal health policies (to account for externalities) and actual policies, see Paul J. Feldstein, *supra* note19, at 374–75 (citing the Clean Air Act 1977 Amendments).

[23] *See , e.g.,* CHARLES WOLF JR., MARKETS OR GOVERNMENT: CHOOSING BETWEEN IMPERFECT ALTERNATIVES (1st ed. 1986).

[24] These costs may be "internalized" (for example, considered during the decision), or partially internalized, for friends and family. For strangers, internalization into the individual's decision process is far less likely.

[25] I believe it is best to make a two-step evaluation. The first is a comparison of the world with and without the government intervention. Here the relevant costs are not the traditional economist's first derivative of the total cost function with respect to quantity. Instead, the calculation involves a focus on the decision of government intervention and the costs and benefits associated with that intervention. Often, real world choices are discrete and "lumpy," as opposed to continuous. In the second step, to the extent that there is a continuum of government interventions, it becomes necessary to compare the marginal (or incremental) benefits of additional intervention to the marginal costs of additional intervention.

[26] Some reduction in GDP would occur due to supply chain issues, reduced tourism and travel, voluntary reductions in activity and consumer uncertainty, regardless of government mandates.

[27] The International Monetary Fund revised its forecast for 2020 for the United States economy from down 8% to down 6.6%. *See* Paul Wiseman, *IMF: U.S. Economy Will Drop 6.6% in 2020 in Face of Pandem*ic, ABC News (July 17, 2020), https://abcnews.go.com/Business/wireStory/imf-us-economy-drop-66-2020-face-pandemic-71850554. This still leaves a net drop (from a typical year's growth in real GDP) of at least 8.6% (which is equivalent to about $1.87 trillion). *See* U.S. Bureau of Economic Analysis, *Gross Domestic Product, Fourth Quarter and Year 2019 (Advance Estimate)*, BEA (Jan. 30, 2020), https://www.bea.gov/news/2020/gross-domestic-product-fourth-quarter-and-year-2019-advance-estimate.

2. PRECEDENTS AND ANALOGIES FOR MASK MANDATES

There are two categories of precedents related to requiring the use of masks. First, the federal Occupational Safety and Health Administration [hereinafter O.S.H.A] requires certain occupations to wear surgical masks and other occupations to wear N-95 respirators. Those facing such requirements must also be tested regularly to ensure the ability to properly fit in and wear the mask.[28] For medical workers, there is an external cost (benefit to others) of wearing a mask. This serves as the rationale for mask requirements. For medical workers, masks may reduce the probability that the wearer becomes infected with an airborne disease, or (if infected themselves) the mask may reduce the probability that the wearer unwittingly infects patients and other staff. However, for industrial workers (also covered by O.S.H.A), masks are intended to reduce (inbound) exposure to particulate matter (such as asbestos, paint, coaldust, or sawdust) or chemicals (such as solvents) rather than to reduce the spread of infectious diseases. For industrial workers, there are virtually no external costs (non-complying workers generally can't infect others with their respiratory disease, such as asbestosis).[29]

There are other government requirements that are partially comparable to a requirement to wear masks. Smoking (including e-cigarettes) is banned on most flights around the world. In the United States, only twelve States had "not enacted any general statewide ban on smoking in workplaces and/or bars and/or restaurants."[30] Most of those twelve States have laws requiring non-smoking sections in such establishments. These prohibitions were based in part upon evidence of the health dangers of second-hand smoke — an external cost — and analogies can be made to the COVID-19 pandemic.[31] Americans seem to have generally accepted smoking prohibitions. However, the externality rationale is much stronger for COVID-19 than for second-hand smoke: if you contract lung cancer from second-hand smoke, you cannot spread the cancer to others you come in contact with.

Consider just one aspect of modern life for which there is substantial government intrusion: driving. Americans have generally accepted laws requiring driver's licenses, proof of insurance, seat belts, airbags, crumple zones, fuel efficiency

[28] *See* Occupational Safety and Health Administration, *29 CFR Part 1910: Additional Ambient Aerosol CNC Quantitative Fit Testing Protocols: Respiratory Protection Standard*, United States Department of Labor (Sept. 26, 2019), https://www.osha.gov/laws-regs/federalregister/2019-09-26.

[29] *See* Mayo Clinic, *Asbestosis*, Mayo Clinic (Dec. 27, 2019), https://www.mayoclinic.org/diseases-conditions/asbestosis/symptoms-causes/syc-20354637.

[30] Wikipedia, *List of Smoking Bans in the United States*, Wikipedia (July 17, 2020), https://en.wikipedia.org/wiki/List_of_smoking_bans_in_the_United_States.

[31] *See* Barry Schwartz, *Secondhand Smoke, Moral Sanctions, and How We Should Respond to COVID-19*, Behavioral Scientist (June 22, 2020), https://behavioralscientist.org/secondhand-smoke-moral-sanctions-and-how-we-should-respond-to-covid-19/.

levels, car safety standards, emissions tests, speed limits, car-pool only lanes, and blood alcohol limits. Similarly, life jackets are required in boats,[32] football and motorcycle helmets are required in some States, and government restrictions on the maximum rent you can charge on a home exist in some cities.[33] These requirements have either weak externality-based justifications[34] or none at all.

However, mask requirements have triggered some surprisingly emotional reactions.[35] Store employees often suffer the brunt of anti-mask reactions including broken arms,[36] being punched in the face,[37] or even being shot to death.[38]

3. MASK EFFECTIVENESS

Research on the effectiveness of masks is not new. In 2011, a scientific article reviewed sixty-seven studies and found that "[s]imple and low-cost interventions [hand washing and wearing masks] would be useful for reducing transmission of epidemic respiratory viruses."[39]

A recent cross-country comparison found mask-wearing is highly correlated with low per-capita COVID-19 mortality rates.[40] Those countries with high COVID-19 rates and lower instances of mask-wearing were: Brazil, Turkey, Spain, Italy, the United

[32] Federal law requires children under the age of 13 to wear a life jacket while on a boat at all times. *See* . Other laws require a life jacket to be available for all those on board. *See, e.g.*, Penny Kanable, *Life Jackets*, Wisconsin Department of Natural Resources (May 16, 2019), https://dnr.wi.gov/topic/boat/pfd.html.

[33] *See, e.g.*, Wikipedia, Rent Control in the United States, Wikipedia (Oct. 27, 2020), https://en.wikipedia.org/wiki/Rent_control_in_the_United_States.

[34] Emissions tests do have an externality justification. There is something similar to an externality justification for right-of-way laws. However, there is a weaker externality justification for carpool-only lanes and fuel-efficiency levels.

[35] *See* Occupational Safety and Health Administration, *Respiratory Infection Control: Respirators Versus Surgical Masks*, OSHA Fact Sheet (May 2009), https://www.osha.gov/Publications/OSHA3219.pdf; Tina Hesman Saey, *Why Scientists Say Wearing Masks Shouldn't Be Controversia*l, Science News (June 26, 2020), https://www.sciencenews.org/article/covid-19-coronavirus-why-wearing-masks-controversial.

[36] *See* Los Angeles Police Department, *Two Suspects Arrested for Assaulting a Security Guard*, LAPD Online (May 11, 2020), http://www.lapdonline.org/newsroom/news_view/66538.

[37] *See* Chief Robert Schurr, *Brown, Elijah S – (A)(2) Aggravated Assault and 5 Additionally Charges*, Bucks Crime Watch PA (May 11, 2020), https://bucks.crimewatchpa.com/perkasieboroughpd/36078/arrests/brown-elijah-s-18-2702-a2-aggravated-assault-and-5-additional-charges.

[38] *See* Alec Snyder et al., *Three Family Members Charged in Shooting Death of Security Guard Who Told a Customer to Put on a Face Mask*, CNN (May 5, 2020), https://www.cnn.com/2020/05/04/us/michigan-security-guard-mask-killing-trnd/index.html

[39] Tom Jefferson et al., *Physical Interventions to Interrupt or Reduce the Spread of Respiratory Viruses*, Pub Med (July 6, 2011), https://pubmed.ncbi.nlm.nih.gov/21735402/.

[40] *See* American Thoracic Society, *Countries with Early Adoption of Face Masks Showed Modest COVID-19 Infection Rates*, Medical Xpress (June 24, 2020), https://medicalxpress.com/news/2020-06-countries-early-masks-modest-covid-.html (citing Sunny H.Wong et al., *COVID-19 and Public Interest in Face Mask Use*, 202 Am. J. Respiratory & Critical Care Med. 453 (2020).

States, Russia, France, and the United Kingdom. Those countries with low COVID-19 rates but high mask-wearing included: Vietnam, Cambodia, Hong Kong, Thailand, and Sri Lanka. The United States has COVID-19 mortality rates that are fifty times higher than any of these countries in the second category.[41]

In examining the effects of mask-wearing, it is important to distinguish between the benefits to the person wearing the mask and the benefits to others around the person wearing the mask. Several recent scientific studies indicate that cloth masks provide only some protection to the wearer, but result in much greater effectiveness in reducing the spread of the virus to others.[42] This asymmetry in protection is due to a mask's ability to prevent outbound droplets from becoming aerosolized as microdroplets.[43] This indicates that the private benefit of wearing masks is much smaller than the public benefit of wearing a mask (the external benefit to others who might have been infected). Therefore, the externality effect of mask-wearing during COVID-19 is stronger than other public health activities for dealing with other infectious diseases, such as taking a vaccine (where there is an externality effect, but the external benefits exceed the private benefits by a smaller degree than for mask-wearing during COVID-19).

A study in the Proceedings of the National Academy of Sciences found that "among all the strategies for reducing transmission, wearing face masks may be *the* central variable that determines the spread of the virus."[44] A study released on July 19, 2020 traced 139 patrons that were exposed to two hair stylists who were COVID-19-positive at the time, but the stylists wore masks.[45] None of the 139 patrons were infected while four of six close contacts/family members for the stylists outside of work (where no masks were worn) were infected.

---

[41] *See* World Health Organization, *WHO Coronavirus Disease (COVID-19) Dashb*oard, WHO (July 20, 2020), https://covid19.who.int/.

[42] *See, e.g.*, Peter Szperling, *Should You Wear a Mask to Prevent the Spread of COVID-19?*, Ottawa News (Apr. 6, 2020), https://ottawa.ctvnews.ca/should-you-wear-a-mask-to-prevent-the-spread-of-covid-19-1.4885076. OSHA describes the idea that one of the uses of a surgical mask is to "place [the masks] on sick people to limit the spread of infectious respiratory secretions to others." Occupational Safety and Health Administration, *OSHA Fact Sheet*, OSHA (May 2009), https://www.osha.gov/Publications/OSHA3219.pdf.

[43] Without a mask, "a single cough or sneeze can produce 100,000 microdroplets, some less than 10 microns in size." In an enclosed space, without ventilation, these micro droplets can drift for 20 minutes. This increases the potential for viral spread. Douglas Broom, *This Japanese Experiment Shows How Easily Coronavirus Can Spread – and What You Can Do About It*, We Forum (Apr. 14, 2020), https://www.weforum.org/agenda/2020/04/coronavirus-microdroplets-talking-breathing-spread-covid-19/.

[44] Alice G. Walton, *Face Masks May be the Key Determinant of the COVID-19 Curve, Study Suggests*, Forbes (June 13, 2020), https://www.forbes.com/sites/alicegwalton/2020/06/13/face-masks-may-be-the-key-determinant-of-the-covid-19-curve-study-suggests/#1fb73bd56497.

[45] *See* M. Joshua Hendrix et al., *Absence of Apparent Transmission of SARS-CoV-2 from Two Stylists After Exposure at a Hair Salon with a Universal Face Covering Policy- Springfield, Missouri, May 2020*, CDC (July 14, 2020), https://www.cdc.gov/mmwr/volumes/69/wr/mm6928e2.htm.

Another study found "community mask use by well people could be beneficial, particularly for COVID-19, where transmission may be pre-symptomatic".[46] Still another large data multivariate regression approach found that "[i]n countries with cultural norms or government policies supporting public mask-wearing, per-capita coronavirus mortality increased on average by just 7.2% each week, as compared with 55.0% each week in remaining countries."[47] This same study found that lockdowns were negatively associated with mortality per capita, but that this effect was not statistically significant.[48] In mid-July of 2020, the Director of the Center for Disease Control, Dr. Robert Redfield, stated that "[i]f we could get everybody to wear a mask right now, I really do think over the next four, six, eight weeks, we could bring this epidemic under control."[49]

Health officials in the United States did not initially recommend mask-wearing for the general public, in part, due to a fear of redirecting supply away from medical professionals,[50] a concern that the public would be more likely to touch their faces while wearing a mask,[51] and an expectation that transmission was largely via viral particles on surfaces.[52] However, now, health officials universally suggest that the public wear a non-medical cloth mask when they can not socially distance themselves, especially indoors. Mask-wearing is now considered to be more important for at least four reasons: more evidence of airborne infection;[53] a high proportion of pre-symptomatic or

[46] C. Raina MacIntyre & Abrar Ahmad Chughtai, *A Rapid Systematic Review of the Efficacy of Face Masks and Respirators Against Coronaviruses and Other Respiratory Transmissible Viruses for the Community, Healthcare Workers and Sick Patients*, Int'l J. Nurs. Stud., Aug. 2020, at 1.

[47] Christopher T. Leffler et al., *Association of Country-wide Coronavirus Mortality with Demographics, Testing, Lockdowns, and Public Wearing of Masks*, MedRxIv (July 2, 2020), https://www.medrxiv.org/content/10.1101/2020.05.22.20109231v4. This effect was significant at the .001 level (which corresponds to a 99.9% confidence level). *See id.*

[48] *See id.*

[49] Fox Television Stations, *Widespread Wearing of Masks Could Get COVID-19 under Control within 4-8 Weeks, CDC Director Says*, Fox News (July 15, 2020), https://fox6now.com/2020/07/15/widespread-wearing-of-masks-could-get-covid-19-under-control-within-4-8-weeks-cdc-director-says/.

[50] *See* Elisabeth Buchwald, *U.S. Health Officials Say Americans Shouldn't Wear Face Masks to Prevent Coronavirus – Here are 3 Other Reasons Not To Wear Them*,Market Watch (Mar. 2, 2020), https://www.marketwatch.com/story/the-cdc-says-americans-dont-have-to-wear-facemasks-because-of-coronavirus-2020-01-30; Nina Bai, *Still Confused About Masks? Here's the Science Behind How Face Masks Prevent Coronavirus*, UCSF (June 26, 2020), https://www.ucsf.edu/news/2020/06/417906/still-confused-about-masks-heres-science-behind-how-face-masks-prevent. *See also* Mary Van Beusekom, *Data Do Not Back Cloth Masks to Limit Covid-19, Experts Say*, CIDRAP (Apr. 09, 2020), https://www.cidrap.umn.edu/news-perspective/2020/04/data-do-not-back-cloth-masks-limit-covid-19-experts-say.

[51] *See* Melissa Quinn, *Surgeon General Says Administration "Trying to Correct" Earlier Guidance Against Wearing Masks*, CBS News (July 12, 2020), https://www.cbsnews.com/news/coronavirus-surgeon-general-jerome-adams-wearing-masks-face-the-nation/ (providing the explanation by the U.S. Surgeon General on the reversal of the position on wearing masks).

[52] *See generally* Center for Disease Control and Prevention, *How COVID-19 Spreads*, CDC (June 16, 2020), https://www.cdc.gov/coronavirus/2019-ncov/prevent-getting-sick/how-covid-spreads.html?deliveryName=USCDC_2067-DM31064.

[53] *See, e.g.*, Eden David & Dr. Mark Abdelmalek, *Why Scientists Think COVID-19 May Be Spread through Particles in the Air*, ABC News (July 8, 2020), https://abcnews.go.com/US/scientists-covid-19-spread-particles-air/story?id=71665634.

asymptomatic infectious people;[54] continued shortages of antigen or antibody tests;[55] and long delays in processing tests.[56] This is important, because "this will help people who may have the virus and do not know it from transmitting it to others."[57] Recent scientific evidence supporting mask-wearing is voluminous.[58]

The effectiveness of mask mandates is likely also influenced by the manner by which localities and their law enforcement choose to effectuate and enforce such mask mandates. States have varied in their enforcement of mask mandates.[59] For example, in Florida, the Miami-Dade Police Department has "cited hundreds of businesses and individuals for not following face mask rules, and the county has collected nearly $300,000 in fines."[60] On the other hand, in Austin, Texas, the police department rarely levies fines, although fines can potentially be as high as $2000 per day for individuals.[61] Many department leaders have said that "punishing people for not wearing masks – which have come to symbolize the pandemic's political divide – would put officers at the center of yet another fraught controversy."[62] Most police departments have focused on education rather than enforcement *per se*. Therefore, some have suggested that the mask mandate enforcement is mostly in the hands of private actors: businesses.[63]

---

[54] *See* Seyed M. Moghadas et al., *The Implications of Silent Transmission for the Control of COVID-19 Outbreaks*, Proceedings of the National Academy of Sciences of the United States of America (July 6, 2020) https://doi.org/10.1073/pnas.2008373117. *See also* Lauren C. Tindale et al., *Evidence For Transmission of COVID-19 Prior to Symptom Onset*, eLife Sciences (June 22, 2020), https://elifesciences.org/articles/57149 (stating "[w[e found that the majority of incidences may be attributable to silent transmission from a combination of the pre-symptomatic stage and asymptomatic infections. Consequently, even if all symptomatic cases are isolated, a vast outbreak may nonetheless unfold.").

[55] *See* Morgan McFall-Johnsen, *The US is in the Middle of Another Coronavirus Testing Crisis – on a Far Larger Scale than Before*, Business Insider (July 10, 2020), https://www.businessinsider.com/another-coronavirus-testing-shortage-has-hit-us-2020-7.

[56] *See* Austin Kemker, *La. Experiencing Lengthy Delays in Getting COVID-19 Test Results Back*, WAFB9 (July 10, 2020), https://www.wafb.com/2020/07/10/la-experiencing-lengthy-delays-getting-covid-test-results-back/; Ken Alltucker, *'Pushing the Frontiers': Long Lines for COVID Tests, Stressed Labs Delay Results as Demand Spikes*, USA Today (July 11, 2020), https://www.usatoday.com/story/news/health/2020/07/11/covid-19-test-results-delayed-labs-struggle-cases-surge/5406936002/.

[57] U.S. Food and Drug Administration, *N95 Respirators, Surgical Masks, and Face Masks*,FDA (June7,2020), https://www.fda.gov/medical-devices/personal-protective-equipment-infection-control/n95-respirators-surgical-masks-and-face-masks%: :text=N95%20respirators%20and%20surgical%20masks%20are%20examples%20of,airborne%20particles%20and%20from%20liquid%20contaminating%20the%20face.

[58] *See, e.g.*, Center for Disease Control and Prevention, *Considerations for Wearing Cloth Face Coverings: Help Slow the Spread of COVID-19*, CDC (July 16, 2020), https://www.cdc.gov/coronavirus/2019-ncov/prevent-getting-sick/cloth-face-cover-guidance.html#recent-studies.

[59] I am co-authoring a piece currently with Carly Gibbons, in which we detail the differing COVID-19 governmental responses by state. Further information on the scope and span of mask mandates can be studied in that article, titled *Gubernatorial Party Affiliation & COVID-19*, which will be forthcoming in an upcoming publication.

[60] Kristine Phillips, *Many Face Mask Mandates Go Unenforced as Police Feel Political, Economic Pressure*, USA Today (Sept. 16, 2020), https://www.usatoday.com/story/news/politics/2020/09/16/covid-19-face-mask-mandates-go-unenforced-police-under-pressure/5714736002/.

[61] *Id.*

[62] *Id.*

[63] *See* Carol Thompson, *Mask Up or Face a Misdemeanor: Michigan Stores Called on To Enforce Mask Rules*, Lansing State Journal (July 10, 2020), https://www.lansingstatejournal.com/story/news/2020/07/10/michigan-enforce-coronavirus-mask-misdemeanor-penalty-fine/5414862002/.

It is also difficult to estimate the optimal or efficient punishment for violation of a mask mandate when taking into consideration a general lax level of enforcement across the States.[64]

4. COSTS AND BENEFITS OF MASK MANDATES

Attempting to quantify the benefit of wearing masks is more difficult than simply recognizing the strong evidence of their ability to reduce the spread of COVID-19.[65] A model developed by the University of Washington's Institute for Health Metrics and Evaluation[66] projected (on July 10, 2020) approximately 44,800 fewer deaths in the United States by November 1, 2020, provided that 95% of the population wore masks. This reflected an incremental increase in mask use from 64% (the United States rate of mask use on July 10, 2020) to 95%.[67]

At that time, a mandatory mask mandate was estimated to be approximately four times more effective than retaining other existing mandates in terms of controlling the spread of COVID-19. On October 28, 2020, the projection was that approximately 62,800 fewer deaths would occur (by February 21, 2021) with 95% of the population engaging in mask-wearing (rather than the continued 69% of the population engaging in mask-wearing which existed on October 28, 2020).[68]

---

[64] *See* Scott Tang, *Will it Work to Fine People Who Refuse to Wear a Mask?*, Market Place (Sept. 16, 2020), https://www.marketplace.org/2020/09/16/penalties-behavioral-economics-fine-people-who-refuse-wear-mask/.

[65] A sound measure of costs and benefits would carefully specify the decision and action in question, and then measure the costs and benefits associated with the decision as compared to the absence of the decision. *See*, R.H. Coase, Business Organization and the Accountant, Econ Lib (Oct.-Dec. 1938), https://www.econlib.org/book-chapters/chapter-coase-business-organization-and-the-accountant/ (stating that "[t]he first point that needs to be made, and strongly emphasized is that attention must be concentrated on the variations which result if a particular decision is taken."). ALASTAIR M. GRAY ET. AL., APPLIED METHODS OF COST-EFFECTIVENESS ANALYSIS IN HEALTHCARE (1st ed. 2001).

[66] *See* Institute for Health Metrics and Evaluation, *United States of America: COVID-19 Projections*, IHME (July 14, 2020), https://covid19.healthdata.org/united-states-of-america.This source explains that the "current projection" scenario assumes that social distancing mandates will continue to be lifted but will be re-imposed for six weeks if daily death rates reach 8 per million. The "mandates easing" scenario assumes that mandates will continue to be lifted and will not be re-imposed. The "universal masks" scenario assumes that mask wearing will reach 95% in 7 days and social distancing mandates will continue to ease but will be re-imposed for six weeks if daily death rates reach 8 per million. Mandates assumed are: "educational facilities closed, non-essential businesses closed, people ordered to stay at home, and large gatherings banned." *See id.*

[67] *See id.* This is assuming the same scenario descriptions as detailed in note 70. The 44,800 lives saved were incremental (marginal) to the increase in mask-wearing from 64% to 95%. It does not include the inframarginal benefits of mask wearing up to 64%.

[68] *See id.* This does not include the inframarginal benefits from mask wearing up to 69%.

However, easing of all mandates (but with mask-wearing at 69%) would lead to approximately 100,000 additional deaths by February 21, 2021.[69] It does appear, therefore, that compared to forecasts on July 10, 2020, variations in the other mandates (such as school closures) have become more important.

The University of Arkansas for Medical Services, College of Public Health's model of COVID-19 also considered the effects of mask use. Comparing a base case of "if conditions do not change" to "almost complete compliance with mask-wearing in public," infection rates and deaths drop by 70%.[70] Another recent study found that "the benefits of *each additional cloth mask* worn by the public are conservatively in the $3,000-$6,000 range due to their impact in slowing the spread of the virus."[71] A recent study by Goldman Sachs found that wearing masks as a partial substitute for business closures could save the United States $1 trillion.[72]

What are the costs of wearing masks? By October 2020, mask prices had dropped from levels observed earlier in the year. I estimate that the federal government could negotiate the purchase of seven double-layer cotton masks per person for less than $1.5 billion. Even this relatively-low estimate is likely not necessary, as there are already a large number of masks in the United States "market." The relevant issue is not the price of masks, but rather whether Americans will wear masks at all. There is some evidence of adverse effects caused by wearing medical-grade surgical masks and N-95 respirators (for example, headaches).

However, this literature largely focuses on N-95 respirators. There does not appear to be any meaningful scientific evidence of any dangers of wearing a non-medical grade cloth mask unless one has a pre-existing respiratory problem or claustrophobia.[73] Monica Gandhi, a professor of medicine and an infectious-disease expert at the University of California at San Francisco stated, "common surgical and cloth masks have 'zero impact' on oxygenation and quality of breathing." However, one could add another $50 billion to this analysis to account for inconvenience costs or other costs.

[69] The death forecasts as of October 28, 2020 for February 21, 2021 can be described as follows: base case/current projection = 385,610; 95% mask wearing (and continued other mandates) = 322,360; mandates eased = 485,607.

[70] *See* Fay W. Boozman, *Modeling the COVID-19 Pandemic in Arkansas*, UAMS (June 15, 2020), https://publichealth.uams.edu/search/COVID%20modeling (illustrating that this would cause cases to drop from 20,000 to 6,000 per day).

[71] Jason Abaluck et al., *The Case for Universal Cloth Mask Adoption and Policies to Increase Supply of Medical Masks for Health Workers*, SSRN (April 1, 2020), http://dx.doi.org/10.2139/ssrn.3567438.

[72] *See* Claire Gillespie, *Does Wearing a Face Mask Reduce Oxygen – and Can It Increase CO2 Levels? Here's What Experts Say*, Health (May 13, 2020), https://www.health.com/condition/infectious-diseases/coronavirus/does-wearing-face-mask-increase-co2-levels).

[73] *See* Nur Ibrahim, *Is It Dangerous to Wear a COVID-19 Protective Mask for Too Long?*, Snopes (May 8, 2020), https://www.snopes.com/fact-check/masks-dangerous-health/; Adrienne Dunn, *Fact Check: Wearing a Face Mask Will Not Cause Hypoxia, Hypoxemia or Hypercapnia*, USA Today (May 30, 2020), https://www.usatoday.com/story/news/factcheck/2020/05/30/fact-check-wearing-face-mask-not-cause-hypoxia-hypercapnia/5260106002/ (providing evidence through interviews with medical experts).

A recent article concludes that "[d]espite the public health benefits of mask usage, due to mask mandates likely being enforced discriminatorily, we advise caution against mask mandates."[74] While I can't quantify the costs of discriminatory enforcement of a mask mandate, the much higher rates of COVID-19 infection and mortality for minority populations suggest that if a cost-benefit analysis were done only for non-whites, the results would even more strongly support a mask-wearing mandate.[75] Additionally, a study (largely focusing on the costs and benefits of controlling the spread of influenza) found the benefits of 60% of the population engaging in mask-wearing to be between five and thirteen times greater than the cost.[76] I would expect much larger multipliers for COVID-19 with 95% compliance in mask-wearing.[77]

5. OPTIONS FOR "SOLVING" THE MASK WEARING EXTERNALITY ISSUE

In the presence of an infectious disease such as COVID-19, someone wearing a mask confers an external benefit on others (or alternatively, someone not wearing a mask can create external costs for others). Here I begin by discussing, in a general context, some of the ways in which externalities can be reduced. I, then, consider which of these techniques may (or may not) work for mask-wearing externalities. In a general sense, the possible externality problem "solutions" to consider are: 1) collect and disseminate information; 2) negotiation between parties; 3) pure liability (law suits); 4) taxes or subsidies; 5) having businesses decide; or 6) a government-instituted mask mandate.

---

[74] *See* Robert Gatter & Seema Mohapatra, *COVID-19 and the Conundrum of Mask Requirements*, 77 Wash. & Lee L. Rev. 17 (2020).

[75] *See*, Brian P. Dunleavy, *CDC Data Highlight Racial Disparities in Spread, Scope of COVID-19 Pandemic*, UPI (July 10, 2020), https://www.upi.com/Health_News/2020/07/10/CDC-data-highlight-racial-disparities-in-spread-scope-of-COVID-19-pandemic/6211594399597/; *See also* Jonathan M. Wortham, *Characteristics of Persons Who Died with COVID-19 – United States, February 12 – May 18, 2020*, CDC (July 10, 2020), https://www.cdc.gov/mmwr/volumes/69/wr/mm6928e1.htm.

[76] *See* Shohini Mukerji et al., *Review of Economic Evaluations of Mask and Respirator Use for Protection against Respiratory Infection Transmission*, BMC Infectious Diseases (Oct. 13, 2015), https://bmcinfectdis.biomedcentral.com/articles/10.1186/s12879-015-1167-6. One might suspect that part of the measured benefit was inframarginal (that is, measuring the benefit from those already wearing masks). However, in the United States, unlike some other countries, such as South Korea, there was virtually no mask-wearing by the general public. Virtually the entire measurement is marginal or incremental to the standard of 60% mask-wearing.

[77] The multiplier should be greater, because, compared to influenza, COVID-19 is more infectious, has a higher mortality rate, is physically somewhat larger (1.2 microns as opposed to 0.8 -1.0 microns), and is more likely to be trapped. Moreover, the multiplier should be greater after taking into consideration the presumption of a higher compliance rate.

Before discussing the possible solutions to the mask-wearing externality problem, it is useful to note that there is a market transaction (the purchase of masks) that I find to comprise a trivial part of the issue. Masks are now inexpensive and widely available. I argue that the absence of mask-wearing by a segment of the population in the United States, today, is not because these individuals lack masks or access to masks, but rather because they do not want to wear masks. For example, political rallies for President Trump generally provided facemasks to attendees.[78] However, the vast majority of attendees did not wear facemasks.[79] Therefore, the externality dimension of mask-wearing is not (in any meaningful way) related to a market transaction (such as the sale of masks), but rather to non-market decisions and actions by people interacting with other people, often in public places.

In Section I, I listed collection and dissemination of information as one of the ways clubs or governments may attempt to reduce the impact of externalities. There have been substantial efforts to disseminate information on the importance of mask-wearing, hand-washing, limiting the sizes of gatherings, and social distancing in reducing the spread of COVID-19. However, there is also substantial "misinformation,"[80] all of which is presented by those who are not epidemiologists or infectious disease experts.[81] Moreover, President Trump and some others in his administration have downplayed the severity of the pandemic[82] and have not encouraged mask-wearing.[83]

[78] *See, e.g.*, Courtney Subramanian & David Jackson, *Trump Campaign to Provide Temperature Checks, Face Masks to Tulsa Rally Attendees*, USA Today (June 15, 2020), https://www.usatoday.com/story/news/politics/2020/06/15/tulsa-rally-trump-campaign-provide-face-masks-temperature-checks/3191388001/.

[79] *See* J. Edward Moreno, *Most Trump Rally Attendees Opt Not to Wear Face Masks*, The Hill (June 20, 2020, 8:09 PM), https://thehill.com/homenews/campaign/503752-most-trump-rally-attendees-opt-not-to-wear-face-masks.

[80] Here, I define "misinformation" as that which is not supported by the majority of peer reviewed scientific literature.

[81] *See, e.g.*, Denis G. Rancourt, *Masks Don't Work: A Review of Science Relevant to COVID-19 Social Policy*, River Cities Reader (June 11, 2020), https://vaccinechoicecanada.com/wp-content/uploads/masks-dont-work-denis-rancourt-april-2020.pdf (written by a metals expert in physics that lost his academic position). To the layman, this could appear to be a peer-reviewed medical article. *See also*, the statements and tweets by Dr. Scott Atlas (a radiologist and former Fox News commentator, appointed as a White House Science advisor in August 2020). One of Dr. Atlas's tweets was removed by Twitter as misinformation. Cathy Bussewitz, *Twitter Blocks Tweet from Scott Atlas*, Washington Times (Oct. 18, 2020), masks/https://www.washingtontimes.com/news/2020/oct/18/twitter-blocks-scott-atlas-tweet-masks/ .

[82] *See, e.g.*, Grace Segers & Kathryn Watson, *Trump Admitted to Woodward He Downplayed Coronavirus Threat in Early Days of Outbreak*, Cbs News (Sept. 10, 2020, 11:22 AM), https://www.cbsnews.com/news/trump-woodward-book-claims-downplayed-covid-19-threat/; Juana Summers, *Timeline: How Trump Has Downplayed the Coronavirus Pandemic*, NPR (Oct. 2, 2020, 8:09 AM), https://www.npr.org/sections/latest-updates-trump-covid-19-results/2020/10/02/919432383/how-trump-has-downplayed-the-coronavirus-pandemic.

[83] *See, e.g.*, Robert Farley, *Trump Has Not Been 'Clear' in Support of Masks*, Fact Check (Sept. 25, 2020), https://www.factcheck.org/2020/09/trump-has-not-been-clear-in-support-of-masks/ .

This is likely a major cause of lower willingness to wear masks by Republicans than by Democrats.[84]

As noted in Section I, negotiations may reduce externalities when property rights are well-defined and transaction costs are low. However, disagreements exist over property rights with respect to mask wearing. A study by the Brookings Institution found that "64 percent of Americans believe that their right to not have to wear a mask or a scarf over their face is more important than reducing the likelihood of contracting the virus or spreading it to others."[85] This included a substantial portion of people who do wear a mask in public. In addition, transaction costs are extremely high. In contrast to the number of steel mills in the United States, every person is a potential "polluter" (potentially infecting other people). Moreover, the number of externality events is very large. If a person only goes out of their home twice per week to shop, and comes within less than six feet of another person ten times each trip they still have twenty incidents per week in which they might try to negotiate having other people wear masks. I expect that during the time spent in attempting to negotiate with the ten people, more than ten new people would come into the store. Others, say those forced to take public transit to work, who then work, and then make the trip back home, may be exposed to hundreds or thousands of people per week. Negotiation is simply not a viable alternative.

Next, consider a pure liability solution. In theory, if one could identify the entity that caused an infectious disease and prove liability, a pure liability solution is possible via individual lawsuits or class action lawsuits. However, except in rare circumstances, lawsuits involving infectious diseases are not successful.[86] Moreover, as noted above, a

[84] *See, e.g.*, James Joyner, *Republicans Refuse to Wear Masks*, Outside The Beltway (July 13, 2020), https://www.outsidethebeltway.com/republicans-refuse-to-wear-masks/ (stating that "[t]he one exception [to frequent mask wearing] is Republicans, among whom a majority say they wear masks infrequently — either sometimes (18%), rarely (9%) or never (27%)"). Moreover, a Gallup poll in July found that "27% of Republicans say they 'never' wear a mask outside their home, versus 1% of Democrats." *See* William Saletan, *Republican Voters Have Come Around on Masks*, Slate (July 31, 2020), https://slate.com/news-and-politics/2020/07/republicans-masks-coronavirus-polls.html. Another contributing factor may be differences in the level of trust in science between the political parties. *See* Felix Richter, *Has Trust in Science Become a Partisan Issue?*, Statista (Oct. 21, 2020), https://www.statista.com/chart/23248/trust-in-scientists-by-political-leaning/.

[85] Korin Miller, *Many Americans Say It's 'Their Right' Not to Wear a Mask. Experts Say it's a 'Threat' to the Country*, Yahoo! Life (Sept. 1, 2020), https://www.yahoo.com/lifestyle/many-americans-right-not-to-wear-mask-experts-say-threat-country-193153762.html?guccounter=1&guce_referrer=aHR0cHM6Ly93d3cuYmluZy5jb20v&guce_referrer_sig=AQAAAN6fF1FSy776JN-9Oa3FWoYO5posx7meGO_2uAooxrghjWUi55ACf7_2BxQAgy3NBeTXuxXaaMTvmjnnwT204PBsAg7Lv2ioH-EutVITtoeHPDtaYY3W_GZmoqXQ5rFRA-_HrFkakTYGlNO1UOdbJTYTfRlfCntccbHWhHyCgTRv.

[86] One of the rare exceptions involves HIV and STD criminal laws. During the early years of the HIV epidemic, several states implemented HIV-specific criminal exposure laws that applied even to behaviors that could not transmit HIV and applied regardless of actual transmission. *See HIV and STD Criminal Laws*, Centers For Disease Control And Prevention, https://www.cdc.gov/hiv/policies/law/states/exposure.html (last visited Dec. 21, 2020).

high proportion of the population that is infected are pre-symptomatic and asymptomatic. They have no knowledge of their infection and proving the source of one's infection will be virtually impossible.

It may also be possible to use a subsidy (a "carrot" rather than a "stick") to deal with an externality. One could theoretically consider taxing those not wearing a mask or subsidizing those wearing a mask (similar to the subsidies provided for immunizations).[87] There are three key differences between mask-wearing and immunization externalities. First, the cost of a mask is trivial. On the other hand, the cost of immunization can be meaningful to some people, especially low-income people. This means that subsidizing the market transaction of purchasing an immunization may be effective, but it is irrelevant for mask wearing. Second, in a lifetime, a parent may schedule a half dozen childhood immunizations. In contrast, many people may make a choice to put on (or not put on) a mask a dozen times a day. It is far more difficult to monitor mask-wearing than childhood immunization. Similarly, someone cannot "take off" an immunization. In addition, one of the studies that suggested subsidizing immunizations was often superior to mandated immunizations also found that when the interest rate was zero or when deficit spending was employed (such that future period tax payers would finance the program), government mandatory immunization was superior.[88] These findings approximate the current conditions through actions by the United States Federal Reserve Bank. As a practical matter, attempting to subsidize mask wearing, or tax those not wearing masks, is simply not practical.

There is a body of literature addressing free-riding (where the free-rider receives benefits but does not contribute to reducing a problem as mentioned in Section I). Some studies indicate that free-riders were less of a problem than logic and economics would suggest they should have been.[89]

[87] Childhood immunizations have been subsidized for many years in the United States as a method by which to reduce the externality problem associated with unvaccinated individuals. For a simple description, *see* JJAMES W. HENDERSON, HEALTH ECONOMICS AND POLICY 77 (5th ed. 2012). *see also* Feldstein, *supra* note19, at 371. Modern immunizations include: Hepatitis B, Rotavirus, Diphtheria, Pertussis and Tetanus (DTaP), Pneumococcal disease, Polio, Influenza, Varicella (chickenpox), MMR measles, Rubella, Mumps; Measles; and Meningococcal disease. *See generally, Vaccine History: Developments by Year*, CHILDREN'S HOSPITAL OF PHILADELPHIA https://www.chop.edu/centers-programs/vaccine-education-center/vaccine-history/developments-by-year (last updated Mar. 30, 2021); *see also Childhood Immunizations in the United States*, WIKIPEDIA, https://en.wikipedia.org/wiki/Childhood_immunizations_in_the_United_States (last visited May 19, 2021, 1:09 PM).

[88] *See* Pierre-Yves Geoffard & Tomas Philpson, *Disease Eradication: Private Versus Public Vaccination*, 87 AM. ECON. REV. 222, 227 (1997).

[89] *See, e.g.*, Bohm, *supra* note 18; Marwell & Ames, *supra* note 18; Sweeney Jr., *supra* note 18; Smith, *supra* note 18; Booth, *supra* note 18.

This implies that less intrusive approaches may be superior. In contrast, there is literature that indicates that economic agents behave as expected in the face of free-rider opportunities or external costs.[90] The implications that stem from the free-rider literature are mixed.

One approach could be to allow individual businesses to decide whether to require masks. This could lead to voluntary separation into two societal groups: mask-wearers and non-mask-wearers.[91] Unfortunately, many businesses would be unable to survive with only a portion of the total consumers. This is particularly true during the pandemic. Moreover, in many areas, there may be only one practical business location alternative for customers, particularly in low-income areas, urban-commercial islands, and in rural areas.

I believe, as a practical matter, mask mandates aimed to help to reduce the spread of COVID-19 are economically valid for the following reasons. First, the evidence of mask effectiveness is overwhelming. Second, a mask mandate may eliminate the need for more draconian measures, such as business closures, and masks are far more cost-effective than such measures. Third, one survey found that 82% of Americans supported a national mask mandate.[92] Other surveys suggest that the majority of Americans are in favor of mask mandates with penalties for those who do not comply and that this proportion is higher than those who actually wear masks.[93] This is consistent with a concept of Americans recognizing the potential for free-riders; the percentage support for mask-wearing appears to be higher when there is a mandatory mechanism to reduce free-riders.

Fourth, with approximately 325 million Americans over the age of five years, and with such Americans often coming into close contact with others several times a day, the number of potentially-infectious events (literally billions per day) makes non-mandatory mechanisms simply impractical.[94]

---

[90] *But see, e.g.*, Kim & Walker, *supra* note 18. Charles R. Plott, Externalities and Corrective Policies in Experimental Markets, 93 Econ. J. 106 (1983) (stating that "[w]ithin the simple setting explored here, the traditional models found in the economics literature are amazingly accurate.").

[91] There would still need to be rules for public places, including public transit.

[92] Gabriela Schulte, *Poll: 32 Percent of Voters Support a National Face Mask Mandate*, The Hill (Aug. 3, 2020), https://thehill.com/hilltv/what-americas-thinking/510317-poll-82-percent-of-voters-support-a-national-mask-mandate. *See also* Brian Yermal Jr., *Nearly 3 in 4 Voters Support State Face Mask Mandate with Penalties For Those Who Don't Comply*, Morning Consult (July 22, 2020, 6:00 AM), https://morningconsult.com/2020/07/22/face-mask-polling/ (stating that 86% of Democrats support a mandate in their State that could result in a fine or jail time for people who don't wear masks, while 58% of Republicans also back such a measure and 35% opposing it).

[93] *See e.g.* sign.com, *Fines or Not Wearing a Mask* (July 29, 2021), https://www.signs.com/fines-for-not-wearing-a-mask/.

[94] In contrast, childhood vaccinations occur a small number of times in a person's life (depending on the diseases targeted).

CONCLUSION

The existence of an externality is one of the few potentially-valid rationales for government intervention in the affairs of business and citizens. The scientific evidence is overwhelming that mask-wearing can be effective in reducing the spread of COVID-19. The benefits of someone wearing a mask are largely external (the benefits accrue beyond the person deciding to wear a mask). The externality effect would occur even if masks only protected the wearer, but it is particularly strong since masks are better at stopping outbound viral spread (my mask protects you) than stopping inbound viral spread (my mask protects me). I find government-ordered business closures and stay-at-home orders far more onerous than mandatory mask orders. I estimate the cost per life saved via business closures and stay-at-home orders may be at least twenty times greater than mask-wearing. Moreover, surveys suggest a relatively high rate of proportion of Americans favor mandatory mask mandates with enforcement, and such preference is higher than the degree to which the population actually wears masks. I also find that non-mandatory methods of encouraging mask-wearing are largely impractical and insufficient to better "solve" (or at least reduce) the externality problem.

VOLUME 6 ISSUE 1 ARTICLES & ESSAYS

UNIVERSITY OF BOLOGNA LAW REVIEW

https://doi.org/10.6092/issn.2531-6133/13622

Received: 8 Sept. 2020 | Accepted: 13 May 2021 | Published: 6 Oct. 2021

The Price of Transitional Justice: A Cost-Benefit Analysis of its Mechanisms in Post-Revolution Phase

EMAN M. RASHWAN

Eman M. Rashwan (LL.B., LL.M., LL.M.) is Lecturer of Public Law at Cairo University (Egypt), and currently is a Visiting Lecturer at the Hamburg University (Germany) where she's pursuing her Ph.D. in Constitutional Law and Economics. She has worked as an international consultant for leading international organizations, think tanks, and governmental sectors. This research was completed with a generous fund by the German Academic Exchange Service (DAAD). I would like to thank Michael Faure, Stefan Voigt, Edoardo Martino, Nada Maamoun, and the participants of the EDLE seminar at Bologna University, and of the 14th Annual Conference of the Italian Association of Law and Economics, for their valuable comments. All possible mistakes are however my own.

@ eman-rashwan@cu.edu.eg

0000-0002-8830-4169

ABSTRACT

Transitional Justice [hereinafter T.J.] in the post-revolution phase refers to the policies that aim to deal with the autocratic past-regime violations against its people to achieve accountability and democracy and promote human rights and the rule of law. To achieve these goals, the United Nations, within its Rule of Law Initiative, issued in 2010, a set of five mechanisms that work as guidelines for nations recovering from conflicts. I argue that whatever the mechanism or combination selected by a society transforming from an autocracy into democracy is, the nature of these mechanisms requires a trade-off between multiple considerations. To explain this inevitable trade-off, I go through each mechanism in detail, analyze it from both legal and economic perspectives, and then provide a basic cost-benefit analysis. I suggest that transitional justice as a constitutional arrangement requires a holistic approach in its adoption and application because this initial cost-benefit analysis cannot be standardized for all cases. I also suggest that transitional justice policies that take into account proportionality, a combination of different mechanisms, customization of the mechanisms upon the relevant case, and adopting these policies in the formality of basic or organic laws may be expected to have the most effective outcomes achieving the goals of T.J. with the least legal complications.

KEYWORDS

Transitional Justice; Revolutions; Constitutional Law; Economics; Cost-Benefit Analysis

JEL CODES

K40, K41

TABLE OF CONTENTS

INTRODUCTION

Transitioning from autocratic to democratic regimes through revolutions[1] is a path full of hard choices. One of the most challenging choices that the parties to the post-revolution phase have to make is how they deal with the concept of transitional justice [hereinafter T.J.]. The United Nations [hereinafter U.N.] guidelines on T.J., which are part of its *Rule of Law Initiative,* define T.J. as "the full range of processes and mechanisms associated with a society's attempt to come to terms with a legacy of large-scale past abuses, in order to ensure accountability, serve justice and achieve reconciliation."[2] This article focuses on violations committed by autocratic regimes against their people. Consequently, other violations that could be subject to T.J., including those in the context of national and international conflicts, colonization, or any other form of regime transformation that does not include a revolution against an autocratic regime, are beyond the limits of this article.

[1] This article defines revolution as a movement by the people that led to a regime change through abolishing the throwing the existing government. This definition is similar to the one used by many prominent scholars including Huntington, Walt, and Colgan: *see* Jeff Colgan, *Measuring Revolution*, 29 Conflict Mgmt. Peace Sci.444, 446 (2012).Consequently, all other forms of internal or external disrupts that do not 1) have a group of the state citizens as their primary component, 2) rise against an autocratic national regime, and 3) manage to change the governing regime even if temporarily, are beyond the setup of this analysis. These other forms could include rebellions that did not succeed in throwing the government, secession from the state to form a new one, civil war between different national groups, aimless popular implosion, resistance against a foreign colonizing power, or a coup d'etat by the military. The limits of this article is the analysis of the post-revolution phase in cases of revolutions against autocratic regimes that aim to democratization. There could be two scenarios of the regime change in this case: Regime collapse or rupture, and negotiated transition. However, the differentiation between these two scenarios is beyond this article's analysis. For more on the regime ending types, *see* Tricia D. Olsen et al., Transitional Justice In Balance: Comparing Processes, Weighing Efficacy 155–59 (2010).

[2] United Nations, *Guidance Note of the Secretary-General: United Nations Approaches to Transitional Justice* 2 (2010), https://www.un.org/ruleoflaw/files/TJ_Guidance_Note_March_2010FINAL.pdf.

Let us assume that a revolution succeeded against an autocratic regime that had committed human rights violations against its people - and probably against the revolutionaries themselves. Moreover, the revolutionaries, the new-rulers (whomever they are), and the victims (who are every citizen who was directly or indirectly suffering disutility by the corrupted policies of the past-autocratic regime) share the same preferences. These preferences are achieving T.J.'s goals, i.e., accountability, justice, conciliation, and of course, the transformation into a democracy. Obviously, this is the most optimistic scenario of a post-revolution set-up that rarely applies in reality. First, both history and rational choice theory tell us that the policy-makers and the victims would have divergent preferences in most cases. Even if the policy-makers are not part of the past regime and wish to achieve a radical break with it, this does not mean they would necessarily aspire to achieve the revolution goals. The new rulers' ultimate interest would probably be to care for their self-interest rather than the social welfare and consequently try to maximize their powers and profits. Second, even when the new policy-makers have the same preferences as the victims and aspire for social welfare-enhancing policies, there would still be behavioral biases to expect from both parties emerging from the asymmetric information between them; The policy-makers are expected to possess better information about the applied laws, involved parties, and technical details than the victims (the public).

These concerns and primary dilemmas of T.J. are, however, not the topic of this research; they constitute the first-degree T.J. dilemmas. This article deals with the second-degree dilemmas, i.e., the selection among the offered T.J. mechanisms and not resorting to T.J. in the first place as a notion. It is logical to think that these first-degree dilemmas would spill over the second-degree mechanisms' selection process. This spillover can be traced in part through the analysis given to each mechanism. However, as long as these second-degree dilemmas are the core focus of this research, the author performs the following analysis under the previously mentioned assumptions to discuss the mechanisms in a specific manner away from the primary concerns referred to earlier. The argument is that even under these assumptions used to simplify the analysis setup, the tough questions hold; Which mechanisms of T.J. to adopt? Amnesties or prosecutions? Lustration or integration? Financial reparations or symbolic compensation through finding the truth?

The Guidance Note of the Secretary-General on the United Nations Approach to Transitional Justice (2010) sets five T.J. mechanisms.[3] These mechanisms are: prosecution initiatives, facilitating initiatives in respect of the right to truth, delivering reparations, institutional reform, and national consultations.[4]

This article presents a cost-benefit analysis [hereinafter C.B.A.] of each mechanism to explain the trade-off that the policy-makers need to do in the post-revolution phase. The objective of this C.B.A. is to present comprehensive guidance to the expected trade-off between social costs and benefits that each mechanism involves derived from a law and economics analysis and actual examples. Using this basic C.B.A. would then help future informed decisions by policy-makers dealing with T.J. mechanisms' selection processes by identifying the calculations' inputs, but not their outcome that depends on the case.

The main parties to the tackled post-revolution phase are: first, the past-regime members, including the high-ranked members like the president or the chief of the governing political party, and the less-ranked members like governorates or members of that party. This group's primary concern is how to keep influence on the new regime and avoid any punishing mechanisms. Their impact on the policy-making in the transitional phase differs upon the type of transformation and the second group's preferences, i.e., transitional rulers. Second, transitional rulers are the group governing the post-revolution phase. This group does not necessarily have to be the same group that started the revolution, but it is the group that has sufficient power to control the state after removing the last government. How this group weighs the costs and benefits of any T.J. mechanism depends in the first place on how far their preferences are aligned with the first or the third group. Third, the victims, who are the citizens directly or indirectly negatively impacted by the past-regime policies and violations. Among these victims, there are those supporting the revolution and those favouring the past regime.

[3] The guidelines use the terms "components" and "elements" to refer to the mechanisms of T.J. For the sake of consistency and clarity through this article, I only use one term "mechanisms" to refer to them, to not confuse the reader. The literature uses "mechanisms", "components", "elements", "policies", "tools", and "measures" alternatively. However, they are all referring to the same issue, being the processes through which T.J. is applied.

[4] Consequently, the guidelines do not include amnesties as one of the T.J. mechanisms, although both literature and practice do, *see* e.g., Carlos H Acuña & Catalina Smulovitz, *Guarding the Guardians in Argentina: Some Lessons about the Risks and Benefits of Empowering the Courts*, *in* Transitional Justice and the Rule of Law in New Democracies 93 (A. James McAdams ed., 1997); Jack Goldsmith & Stephen D. Krasner, *The limits of idealism*, 132 Daedalus, no. 1, 2003, at 47; Tom Hadden, *Punishment, Amnesty and Truth: Legal and Political Approaches*, *in* Democracy and Ethnic Conflict: Advancing Peace in Deeply Divided Societies 196 (Adrian Guelke ed., 2004); Mark J. Osiel, *Why prosecute? Critics of punishment for mass atrocity*, 22 Hum. Rts. Q. 118 (2000). ; Stephen John Stedman, *Spoiler Problems in Peace Processes*, 22 Int'l Sec., no. 2, 1997, at 5. ; Leslie Vinjamuri & Jack Snyder, *Advocacy and Scholarship in the Study of International War Crime Tribunals and Transitional Justice*, 7 Ann. Rev. Pol. Sci. 345 (2004); Paul W. Zagorski, *Civil-Military Relations and Argentine Democracy: The Armed Forces under the Menem Government*, 20 Armed Forces Soc. 423 (1994). As this research depends on the U.N. guidelines as its reference point, it will be restricted to the five mechanisms they uphold.

However, as long as the popular revolution succeeded in throwing out the past regime, one can presume that the preferences of most of this group are affiliated with the T.J. purposes, including achieving democracy, accountability, and reconciliation, at least in the early stages of the transitional phase when the T.J. policies are still under design. As referred to earlier, this C.B.A. performs under the assumption that both the second and third groups share the same preferences of achieving T.J. goals. At first thought, at least in theory, this should ease the T.J. policy design process and minimize any principal-agent problems to a great extent. However, even under these optimal assumptions, the calculations of the selected policies/mechanisms are complicated.

My argument is that no matter what the adopted mechanism is, there will always be costs for this adoption, including the opportunity cost of adopting a different mechanism, especially since they do not adhere to the same philosophy (approach) in dealing with past crimes or future arrangements. The different entries quantification, multiplied by their probability, in each mechanism C.B.A. depends on the context these mechanisms are applied within. As a result, this quantification cannot be standardized. Consequently, although the compared costs and benefits may be constant, their weighing and the maximization of their outcome is not. This also suggests that not only holistic approaches are expected to be the most efficient in achieving T.J. goals, but in specific, the balanced approaches that combine different mechanisms are expected to promote this efficiency. By efficiency, we refer to Kaldor-hicks efficiency, which C.B.A. is based upon.[5]

Although T.J. deals with past violations, it is an *ex-ante* constitutional arrangement. First, it is a legal arrangement with a constitutional nature because it regulates the state authorities, human rights of the citizens, and institutional reforms. In some cases, this nature meets formality when T.J. laws are generated in the constitutional text or organic laws. Second, it is an *ex-ante* arrangement because it also sets rules for future institutions besides dealing with past human rights violations. This is normal in a legal arrangement in transitional times, also as Teitel puts it: "Law in transitional periods is both backward-looking and forward-looking, retrospective and prospective, continuous and discontinuous."[6] In the T.J. context, political actors are deciding under uncertainty and imperfect information on a future trade-off between the expected utility and the expected cost, in the shadow of a risk margin.

[5] A situation is considered Kaldor-hicks efficient when the economic gains of that situation exceed the economic losses to whomever they occur, *see* Cento G. Velianovski, Economic principles of law 33 (2007). This is the same rationale used to run the C.B.A.. Accordingly, in our context, a mechanism is considered efficient if the benefits out of its adoption exceed the costs of its application, even if, for instance, the benefits are all accumulated by the victims while the past-regime members incur all the costs.

[6] Ruti G. Teitel, Transitional Justice 215 (2000).

Their decisions will influence their positions in the new political order. Consequently, the options in the C.B.A. should be evaluated based on *ex-ante* efficiency.

This quantification is expected to be made using an inspired utility equation of the *Becker model*,[7] according to which. As a constitutional arrangement, T.J. is not a contract, and it does not have a third-party enforcer who would guarantee its application. The judiciary that enforces the constitution and other laws in the typical situations when the government abstains or precludes their application is themselves a party to the conflict. The courts of an authoritarian regime would most probably need institutional reform. Moreover, some judiciary members might benefit from a form of reparation and might serve as witnesses in the truth commissions. In other words, the typical enforcer of the law is here a party, either as a victim, perpetrator, or a policymaker, who has their own costs and benefits of any relevant mechanism. Consequently, T.J. policies, like other constitutional arrangements, need to be self-enforcing. A self-enforcing T.J. policy would be the one in which its benefits (*b*) exceed its costs (*c*) for the different parties, or that the costs of its absence are so high, which gives them the incentives to apply it. The rest of this article tries to present this equation's potential entries within each of the T.J. mechanisms.

The C.B.A. is a methodology that is used differently in each field. In the sphere of public policy making, there are many debates over how far we should rely on C.B.A. results and concerning their limits. This may especially be an issue when C.B.A. is applied to topics that involve considerations that are difficult – or even impossible – to be monetized.[8] This debate, however, is not the subject of this research. However, it is necessary to mention that quantifying the relevant entries in the C.B.A. presented here, and even adding to or deleting from them, is both flexible and inevitable. The given value to each entry, the kind of costs and benefits of each mechanism, and the priority of goals over other goals vary from case to case. The following provides a basic analysis that is meant to be customized through the relevant policy-makers who wish to take it as guidance.

The central research questions this article answers are: What is the economic rationale behind T.J. mechanisms? By "economic," I do not mean only the financial

[7] Becker model for criminal deterrence is a mathematical model presented by Gary Becker (1968) that treats criminals as rational actors who desire to maximize their wellbeing but through illegal means rather the legal ones. Accordingly, Becker suggests that for a criminal sanction to achieve deterrence there should be an equation that takes in account the expected gain of the crime to the perpetrator and the cost resulting of the severity and probability of punishment. For more on Becker's seminal work on crime and dealing with it as a economic concept, *see* Velianovski, *supra* note 5; Christine Jolls et al., *A Behavioral Approach to Law and Economics*, *in* Behavioral Law & Economics 1471 (Cass R. Sunstein ed., 2000); Elena Kantorowicz-Reznichenko, *Any-Where-Any-Time: Ambiguity and the Perceived Probability of Apprehension*, 84 UMKC L. Rev. 27 (2015) ; Gary S. Becker, *Crime and Punishment: An Economic Approach*, 76 J. Pol. Econ. 169 (1968).

[8] For more on the discussion over and strategies of cost-benefit analysis in the sphere of public policy, *see* Cass R. Sunstein, The Cost-Benefit Revolution (2018).

aspects, but rather the economic theory that explains the logic of these mechanisms and their necessity, including both financial and non-financial components. Moreover, what are the expected costs & benefits of each of these mechanisms, including their constitutional complexities? In doing so, the U.N. Guidelines on T.J. will be used as the model of T.J. mechanisms for their international impact and given legal value as a rule book for nations transitioning from autocracy to democracy.

The article proceeds as follows: section one presents five sub-sections, each containing the C.B.A. of the five studied mechanisms. Section two gives general notes and policy implications. The last section is the conclusion.

1. THE MECHANISMS OF TRANSITIONAL JUSTICE

The T.J. mechanisms are channels set to achieve their goals. In this section, I explain them from a legal perspective, provide for the economic intuitions behind them, as well as the challenges faced. Afterward, I give a C.B.A. for each of these mechanisms. Although this C.B.A. is inspired by both theory and case studies available on the subject, it is still a positive theoretical analysis of these mechanisms. It aims at predicting the possible costs and benefits of each mechanism, but only the empirical analysis informs us about the prevailing approaches. However, because of T.J.'s contextual nature that I have referred to earlier, the reader will notice that some arguments could work as costs and/or benefits at the same time. Moreover, some mechanisms may work negatively and positively in different cases, which was also proven by the empirical studies available so far.[9]

1.1. PROSECUTION INITIATIVES

This mechanism entails that those involved in committing the addressed crimes are to be trialed and, where appropriate, punished. The measures of these trials, according to the guidelines, are the following:

[9] *See e.g.,*Oskar N.T. Thoms, James Ron & Roland Pariss, *State-Level Effects of Transitional Justice: What Do We Know?*, 4 Int'l J. Transit. Just. 329 (2010); Roman David, *What We Know About Transitional Justice: Survey and Experimental Evidence*, 38 Pol. Psych. 151 (2017) ; Olsen et al., *supra* note 1; Geoff Dancy et al., *Behind Bars and Bargains: New Findings on Transitional Justice in Emerging Democracies*, 63 Int'l Stud. Q. 99 (2019).

- The trialed crimes include *"the serious violations of international humanitarian law and gross violations of international human rights law."*[10] The word "include" infers the meaning that these crimes can also include other crimes under national laws.

- These trials have to be undertaken according to the *international standards of fair trials.*[11] There is no single comprehensive source of international law that defines what a fair trial entails. Instead, there are different measures mentioned in regional and international instruments, such as the Universal Declaration of Human Rights, the International Covenant on Civil and Political Rights [I.C.C.P.R], the Geneva Conventions, the African Charter on Human and Peoples Rights [A.C.H.P.R], the European Convention on Human Rights, and the American Convention on Human Rights. Amnesty International provides a practical manual on free trials provisions that have been approved either as treaty obligations binding on their parties or as part of the customary international law that reflects the international community's collective will and then binding over all its members.[12] The fair trial standards include a wide range of rights that have to be observed pre and during trials.[13] However, the rules governing these standards also acknowledge the exceptional nature of emergency cases. The first condition is that the state of emergency already and lawfully exists. *"Under international human rights treaties, a state of emergency can be declared only if there is an exceptional and grave threat to the nation, such as the use or threat of force from within or externally that threatens a state's existence or territorial integrity."*[14] Although they may vary in extent, revolutions and transitional phases are usually an emergency state by nature and law. Some treaties gave the states the right to derogate from some of these fair trial guarantees under the conditions of being temporary, necessary, and proportional. However, some of them are absolute and cannot be derogated from under any condition. These absolute rights include:[15]

  - The prohibition against torture or other cruel, inhuman, or degrading treatment or punishment;

[10] International Human Rights Law [hereinafter I.H.R.L.] and International Humanitarian Law [hereinafter I.H.L] are two different complementary branches of Public International Law. Although they have the same aim of protecting the lives and rights of humans, they differ in their origins, and scopes of application. Most significantly, while I.H.L. applies only on the armed conflicts, I.H.R.L applies at all times. *See* International Committee of the Red Cross, *What is the difference between IHL and human rights law?*, *in* International Humanitarian Law: Answers to Your Questions (2014). *See* United Nations, *supra* note 2, at 7.

[11] *Id.*

[12] Amnesty International, Fair Trial Manual 1 (2 ed. 2014), https://www.amnesty.org/en/documents/POL30/002/2014/en/.

[13] For a full review of these rights, *see id.*

[14] *Id.* at 232.

[15] *Id.* at 229–338.

- The right of people deprived of their liberty to humane treatment;
- The prohibition of enforced disappearance;
- The prohibition of arbitrary arrest or detention, including unacknowledged detention;
- The right to be recognized as a person before the law;
- The right to petition a court challenging the legality of detention;
- The right to proceedings before an independent, impartial and competent court;
- The right to a public trial, in all but exceptional cases which are warranted in the interests of justice, the requirement of clear and precise definitions of offenses and punishments;
- The prohibition of retroactive application of criminal laws (including the imposition of a heavier penalty than was applicable at the time of the crime), and the right to benefit from a lighter penalty;
- The obligation to separate people held in pre-trial detention from those who have been convicted and to treat them in line with their status as unconvicted;
- The presumption of innocence;
- The right to legal aid for those without adequate financial resources;
- The prohibition of collective punishment;
- The principle that the essential aim of punishment involving deprivation of liberty is reform and rehabilitation;
- The prohibition against double jeopardy, judicial guarantees, such as *habeas corpus*[16] and *Amparo*[17], to protect non-derogable rights;
- The right to effective judicial remedies for violations of other human rights;
- The right to compensation for individuals whose innocence is established by a final judgment, in addition to the non-derogable rights guaranteed in death penalty cases.

[16] "*Habeas Corpus*" is a Latin expression that translates into "You have the body". Legally it refers to "A writ (court order) that commands an individual or a government official who has restrained another to produce the prisoner at a designated time and place so that the court can determine the legality of custody and decide whether to order the prisoner's release" habeas corpus, West's Encyclopedia of American Law (2 ed. 2008), https://legal-dictionary.thefreedictionary.com/habeas+corpus (last visited Oct 12, 2018).

[17] "The writ of Amparo is a remedy for the protection of individual or constitutional rights, found in certain jurisdictions including Mexico, Spain, the Philippines and parts of Latin America" amparo, The World Law Dictionary Project (2017), https://www.translegal.com/legal-english-dictionary/amparo (last visited Oct 12, 2018).

- The trials should be held in a *non-discriminatory and in an objective and a timely manner*.[18]

- The *jurisdiction* over these crimes is primarily the state's responsibility, so they will need to develop the necessary capacity to perform the prosecutions according to the previous standards. However, the international community can also provide help in conducting such processes. This point entails the need for national laws that conform to International Human Rights Law [hereinafter I.H.R.L.] and International Humanitarian Law [hereinafter I.H.L.][19]

- In the same context as the previous point, the guidelines also suggest, first, that *systematic monitoring* over the judicial performance within these trials can guarantee their effectiveness and fairness. Second, that in cases where the states are *unable or unwilling* to undertake *effective investigations and prosecutions*, international hybrid criminal tribunals[20] can perform a concurrent jurisdiction.[21] The guidelines assure that in such cases, such tribunals have to give *priority consideration to their legacy in the country and their exit strategy*; how will their performance influence the treated country by the completion of the prosecution? Moreover, how can one help the national authorities to improve their capacities to contribute to bringing the alleged perpetrators to justice?[22] However, this conception of international prosecutions seems problematic in many cases. On the one hand, the crimes that the past-autocratic regime members are supposed to be prosecuted for are not always under the crimes subject to the international courts' jurisdiction according to public international law. For example, financial corruption crimes, or elections manipulation, are not considered war crimes or crimes against humanity. On the other hand, even if this point was – supposedly - solved by creating a special international court – which has no precedent for similar crimes - the main obstacle lies in considerations of the principle of national

[18] United Nations, *supra* note 2, at 7.

[19] *Id.*

[20] Hybrid international criminal tribunals are tribunals that consist of both national and international elements on both the level of subjective (applying mixed laws) and procedural (having both national and international judges) levels, *see* Sarah Williams, Hybrid and Internationalised Criminal Tribunals: Selected Jurisdictional Issues (2012).

[21] International courts in this case work as a subsidiary jurisdiction when the national courts fail to perform their jurisdiction under the principle of universal jurisdiction, like the case for International Criminal Court for example. For more on the principle of complementarity in international law, *see* Xavier Philippe, *The principles of universal jurisdiction and complementarity: How do the two principles intermesh?*, 88 Int'l Rev. Red Cross 375 (2006). I think, however, that in case of T.J. international hybrid criminal tribunals can be less problematic than purely international courts for the reasons explained in the main texts, generally related to the past-revolution circumstances and concerns.

[22] United Nations, *supra* note 2, at 7–8.

sovereignty. Efforts by international tribunals to settle national disputes, and policies on international monitoring over the national judiciary, can be rejected based on breaching the state's national sovereignty and independence.[23] Especially since any revolution, at least in the beginning, faces accusations of conspiracy and foreign interference.

Initially, the trials and prosecutions of the past-regime members who committed violations of I.H.R.L. and/or I.H.L. may aim to correct the negative externalities caused by their crimes.[24] However, in this context, the following is noted.

- The crimes of murder, mass killing, torture, and other violence crimes, were either committed during the rule of that regime or the revolutions' incidents. The other crimes, which include political or economic corruption, were committed during the past regime rule. Consequently, the probability of collecting the legally sufficient evidence, according to the international standards of criminal procedures laws, against the direct committers or against the regime's chief leaders who gave the orders is significantly low. Either because the criminals had ample time and necessary authority to destroy the evidence, which happened, for example, in Japan after World War II,[25] and in Egypt after the 2011 revolution,[26] or

[23] *See, e.g.*,Chandra Lekha Sriram, *Revolutions in Accountability: New Approaches to Past Abuses*, 19 Am. Univ. Int'l L. Rev. 301 (2003);Robert Cryer, *International Criminal Law vs State Sovereignty: Another Round?*, 16 Eur. J. Int'l L. 979 (2005).;Atul Bharadwaj, *International criminal court and the question of Sovereignty*, 27 Strategic Analysis 5 (2003) ; Daniel Partan & Predrag Rogic, *Sovereignty and International Criminal Justice*, 1 Int'l L: Revista Colombiana Derecho Internacional, num. 1, at 53 (2003).

[24] Externalities in microeconomics are a form of market failure that happens when a third party to a voluntary action incur either a cost (negative externalities) or a benefit (positive externalities), which creates a divergence between the privately and socially optimal equilibria. The value of the uncompensated external effects to that third party is called externalities. Internalization of externalities refers to the instruments created to induce people to take account of their actions' external effects. *See* Francesco Parisi, The Language of Law and Economics: a Dictionary 114-6 (2013). Negative externalities of crimes refer to the cost that other individuals and society as a whole suffered because of these crimes. For example, the negative externalities of torture include first, the physical, psychological, and monetary losses to the tortured person and also his family; second, the costs born by the society as a result of losing the output of a citizen and the potential benefit of other citizens who would contribute more actively to the society if it was not for fear of being tortured as well; and finally, the damage incurred by the society due to lack of the rule of law, human rights, and democracy which can be cultural, political, and also economical. In this example, and most of the other examples in the sphere of transitional justice context, these negative externalities can reach infinity if we quantify them because of the complexity of included calculations as explained in the example. For more on the negative externalities of the crime, *see also* Graham Farrell & John Roman, *Crime as pollution: proposal for market-based incentives to reduce crime externalities*, *in* Crime Reduction and the Law 135 (Stephens & Moss eds., 2006); John Roman & Graham Farrell, *Cost-Benefit Analysis for Crime Prevention: Opportunity Costs, Routine Savings and Crime Externalities*, 14 Crime Prevention STUDIES, no.1, 2002, at 53.

[25] Zachary D. Kaufman, United States Law & Policy on Transitional Justice: Principles, politics, and pragmatics (2017).

[26] *See* Watan, وطن, وثيقة المخابرات المصرية تكشف تدمير أدلة إدانة «مبارك» والداخلية نقتل المتظاهرين [The Egyptian Intelligence Document Reveals Destroying "Mubarak" and the Interior Ministry Condemnation Evidence of the Protestors Mass Killing], "Watan" , (2014) https://www.watanserb.com/2014/12/04/وثيقة-المخابرات-المصرية-تكشف-تدمير-أد/ (last visited June 29, 2020).

because of the chaotic nature of the revolution's incidents. This concern entails a high rate of error, i.e., criminals not to be punished while non-criminals may be punished. An example of the first case is a top official who misused the public resources for his private interests (rent-seeking) but could destroy all the documents that could prove so. An example of the second case could be a policeman defending his police station against violent attacks during the protests, it could be that he would never kill an innocent citizen in the usual conditions, but in the current case, he should be in a legitimate defense situation.[27] However, the probability of proving so in the middle of a chaotic, politically tensioned, and unstable security situation like a revolution,[28] could be very low. In case these difficulties led to adopting pre-decided punishment based on presumed responsibilities of the position the accused occupied without counting on the usual legal standards, this may as well result in over-deterrence for public posts holders.[29] Officials could stop doing their job duties because they are afraid of being subject to legal accountability in the heat of the moment of the aftermath of revolution and the "war" on corruption, past-regime, police violations, or whatever is targeted. This reaction can reduce the public officer's level of activity below the optimal level to social welfare while exceeding the level of care to socially disturbing levels. Such a behavior has been witnessed already in some cases and will be referred to in more detail later. In other cases, these difficulties resulted in clearing all the sentences because the authorities were unable to collect the necessary evidence, which may result in under-deterrence. The rulers of the transitional phase, the rulers of the new regime, and the public officers of these two regimes would realize that they can get away with abuses against the citizens because the probability of sanctioning them is too low due to evidence difficulties. While performing public duties, the level of care would then go lower than the optimal level for social welfare.[30] The Egyptian case, for example, witnessed the

[27] For more about the doctrine of legitimate self-defense in criminal law, *see* Boaz Sangero Boaz Sangero, Self-Defence in Criminal Law (2006).

[28] *See* Luc Huyse, *Justice After Transition: On The Choices Successor Elites Make In Dealing With The Past*, 1 Transitional Justice 337, 345 (1995).

[29] *See* Richard A. Posner, *An Economic Theory of the Criminal Law*, 85 Colum. L. Rev. 1193, 1221 (1985). For more on the debate about over-deterrence, *see also* Keith N. Hylton, *Economics of Criminal Procedure*, *in* Oxford Handbook of Law and Economics: Public Law and Legal Institutions 325 (Francesco Parisi ed., 2017). ; Jolls et al., *supra* note 7; Richard Craswell, *Deterrence and Damages: The Multiplier Principle and Its Alternatives*, 97 Mich. L. Rev. 2185 (1999).

[30] Concepts of "optimal level of care vs. optimal level of activity" and "overdeterrence vs. underdeterrence" are borrowed from the tort law and economics literature that witnessed lengthy discussions to reach the most efficient liability rule for the social welfare. For more on these discussions, *see* Steven Shavell, Foundations of Economic Analysis of Law (2004); Michael Faure et al., Tort Law and Economics (2009).; Velianovski , *supra* note 5; Jolls et al., *supra* note 7; Craswell, *supra* note 29.

failure of many of the prosecution initiatives because of the adoption of the traditional evidence collecting methods that are not effective in the case of the T.J. process, which led to the insufficiency of the collected evidence.[31]

The difficulties of collecting the necessary evidence are maximized when we take into consideration that even in case this task is allocated to international authorities to avoid any conflict of interests, the national authorities are still involved in the process and are the primary parties who provide this evidence, because of logistic reasons.

- The difficulties in collecting evidence are not the only hardship that prosecution initiatives may face, although the T.J. literature usually overlooks it. Other legal and political obstacles include other concerns that may be present when judging the past regime because of the dispute's political nature. These concerns could include: Politicized courts,[32] the partiality of the judges,[33] the pressure of the public opinion on the judiciary,[34] sanctioning the adoption of specific political opinion, or adhering to a political party that infringes the constitutional right of freedom of opinion, speech, or association, and most importantly retroactive justice.[35]

  If the courts follow the past-regime substantial criminal laws, there is a chance that a number of the perpetrators may escape punishment because the laws were tailored to serve the goals of that regime. An example of this is the acts of the Nazis, which were lawful under the Nazi laws, or at least adopted by the then applicable legal techniques.[36] There are two broad strategies to avoid this:

  – The first and typical way includes the classic techniques adopted by the transitioning systems to finesse the long debate between the moral

[31] *See* Abdullah Khalil, 25 2011) ([The Map of Transitional Justice in Egypt Since 25 January Revolution (The Track - The Challenges - The Policies)] 440 (2017), https://books.google.de/books?id=tkihDgAAQBAJ&pg=PA1&lpg=PA1&dq= 25 source=bl&ots=l7kOfE90uE&sig=ACfU3U3B5q3acNAwOHVKi-UjPZJqRNhGXA&hl=en&sa=X&ved=2ahUKEwjjyb-u4dHqAhXR_KQKHSBjAR8Q6AEwD3oECAoQAQ#v=onepage&q= (last visited Jun 27, 2021).

[32] *See* Ellen Emilie Stensrud, *New Dilemmas in Transitional Justice: Lessons from the Mixed Courts in Sierra Leone and Cambodia*, 46 J. Peace Rsch. 5 (2009). ; *see, e.g.*, Susan Thomson, *The Darker Side of Transitional Justice: The Power Dynamics Behind Rwanda's "GACACA" Courts,* 81 J. Int'l African Inst. 373 (2011).

[33] *See* Richard Lewis Siegel, *Transitional Justice: A Decade of Debate and Experience,* 20 Hum. Rts. Q. 431 (1998). Thomson, *supra* note 32.

[34] *See* Siegel, *supra* note 33.

[35] *See* Eric A. Posner & Adrian Vermeule, *Transitional Justice as Ordinary Justice,* 117 Harv. L. Rev. 761 (2004). *see, e.g.*,Marek M. Kaminski et al., *Normative and Strategic Aspects of Transitional Justice,* 50 J. Conflict Resol. 295 (2006).

[36] *See* Hans Petter Graver, Hans Petter Graver, Judges Against Justice: On Judges When the Rule of Law Is Under Attack (2014).

considerations against retroactive laws, or alternatively amnesties. Eric A. Posner & Adrian Vermeule classify these techniques into: [37]

(a) The appeal to higher pre-existing law. Either this law is the constitution, international law, or natural law.[38] Although it may be advocated that these rules, even in the absence of national law, signal the illegality of the accused's actions in breach of these legal principles, the historical application of this norm did result in genuinely controversial legal and political consequences.

(b) Taking nominal law seriously, i.e., the exact words of the laws. This methodology entails the rigid literal interpretation of the old-regime rules, which could lead in many cases to better results than its implicit goals.

(c) Interpretive statutes to the old laws. Unlike the first two techniques that courts usually practice, this technique is followed by the legislature to smooth over the conflict between retroactive laws and procedural legality. Accordingly, the legislatures may enact interpretive statutes that proclaim an understanding of the past-regime laws that, despite appearances, is different from the one these laws authorized or even allowed.[39]

(d) Retroactive extension of statutes of limitations.[40]

– The second strategy is to explicitly adopt *ex post facto* criminal legislation, which means "a law is made after the doing of the thing to which it relates."[41] This strategy initially contradicts the constitutional and international established legal principle *nullum crimen nulla poena sine lege*, which means that an act can be neither criminalized nor penalized without a pre-existing law. However, some decisions were delivered by the European Court of Human Rights addressing the cases of lustration laws against the past communist in the east European countries that approved such laws against the allegations of retroactivity. These decisions' reasoning depends basically

[37] Posner & Vermeule, *supra* note 35.

[38] The definitions and debates over what constitutes "natural law" are a subject of a voluminous legal literature. For more about its definitions and aspects, *see* Graver , *supra* note 36, at 143–151; Claus OffeClaus Offe, Varieties of Transition: The East European and East German Experience 89 (1996).

[39] An example of this technique is found in the Belgian case after World War II when the narrow scope of the treason crime in the penal code limited only to the military was widened via interpretive statues to include other forms of indirect collaboration. *See* Henry L. Mason , Henry L. Mason, The Purge of Dutch Quislings: Emergency Justice in the Netherlands 129–30 (1952).

[40] For more about these techniques, *see* Posner & Vermeule, *supra* note 35, at 791–800.

[41] William Winslow Crosskey, *The True Meaning of the Constitutional Prohibition of Ex-Post-Facto Laws,* 14 Univ. Chicago L. Rev. 539, 539 (1947).

on two pillars: First, the emphasization of the justice considerations and their contribution to the establishment of good governance. The court, in one of its decisions, explicitly stated: *"it is not a case of the retroactive application of criminal law but of an inexcusable mistake of law;"*[42] Second: that rule of law and justice considerations, and the trade-offs between them, should be interpreted and weighted in their historical, temporal and political contexts.[43]

Besides these legal concerns, the political consequences of prosecutions are also questioned. Fears of backlash by the past-regime members or supporters, especially if it was a military regime, are highly credible. The stronger and more controlling the past regime was, the wider is the range of the persons threatened with prosecutions' initiatives, the more violent is the "counter-revolution." It is like building an army against the new regime, which is still not adequately controlling the state because of the nature of the phase. Another result can be the political and social isolation of a group of society. Moreover, breaching the rule of law and free societies' principles mentioned earlier may also threaten the new democracy.[44] All of these are possible costs of prosecutions.

- As mentioned in the last point, including all the suspects in the prosecution initiatives can mean prosecuting hundreds or even thousands of people depending on every regime's government and police structure. The administrative costs of collecting the evidence and processing the trials against such a large number of accused persons will be remarkably high. These costs also include the time costs, as one of the guarantees of a fair trial is the right to appeal;[45] the prosecutions usually take years of litigation. A suggestion of selecting a limited number of perpetrators could seem problematic for equality reasons, on the one hand, and the dilemma of "whom to select" on the other. Bruce Ackerman raises these pragmatic concerns about the transitional prosecutions.[46] He argues that selecting the leading figures of the past regime will face the procedural problems of evidence since usually, their orders were implicit or unwritten. While selecting the minor figures faces the questions of liability in case they were executing

[42] Cynthia M. Horne, *International Legal Rulings on Lustration Policies in Central and Eastern Europe: Rule of Law in Historical Context*, 34 Law& Soc. Inquiry 713, 735 (2009). https://www.jstor.org/stable/pdf/40539376.pdf?refreqid=excelsior%3Ae62287ca317d373f232a93ac478af73f (last visited Jun 27, 2021).
[43] *Id.* at 734–37.
[44] *See* Huyse, *supra* note 28.
[45] Amnesty International, *supra* note 12, at 182–91.
[46] Bruce Ackerman, The Future of Liberal Revolution 69–98 (1992).

higher orders, and also causes rage against the new regime for only "haunting the small fish." However, Eric A. Posner & Adrian Vermeule reply to these arguments that this does not mean that: First, these are not typical difficulties of other organized crimes' prosecutions, and second, they do not entail that the optimal number of prosecutions should be zero.[47] In light of this debate, I think that the optimal number of prosecutions should be a result of a careful calculation of the anticipated costs and benefits of the different scenarios of the prosecuting range, which will differ from case to case. How to define and evaluate the costs and benefits for each scenario? The following cost-benefit matrix in table (1) will illustrate the entries that each C.B.A. of a prosecutions' scenario should include. This is the first step, i.e., defining. The second step is evaluating; how to put weight on these entries to solve the different equations and compare the outputs? This process would ultimately be done through national consultations. Consulting national and international experts, civil society organizations, victims' associations, and other institutions representing the pro-past regime preferences shall help each society put approximate numbers on these entries multiplied by their probability. Eventually, comparing each scenario's outcome shall help each society choose the most efficient range of prosecution. For example, legal and economic experts would help to estimate 1. the expected time and administrative costs of a scenario that includes prosecuting only the past-regime first-line members, i.e., heads of authorities and ministers, and 2. the amount of financial resources illegally accumulated by these personnel that could be restored to the state treasury and the probability of its restoration. In the meantime, national and international organizations can then advise, based on previous comparative experience, what the extent and probability of this limited range's impact are on public deterrence in the long run. This example continues for the rest of the entries and scenarios. The same strategy applies too to evaluating C.B.A.s for the other T.J. mechanisms.

- The error costs are also remarkably high. In case of false convictions, the prosecutions usually lead to either death, a huge fine, or a prison penalty. While the second could mean a non-optimal allocation of financial resources, the first and the third mean a loss of human capital. In the case of wrong acquittals, the ex-perpetrators will have space by the law, the necessary expertise, and sufficient incentives to work against the new regime. Taking into account the high probability of error, the costs in this case are multiplied.[48]

[47] Posner & Vermeule, *supra* note 35, at 800.
[48] For more about administrative and error costs, *see* Hylton, *supra* note 29.

- Back to the point of over-deterrence, the costs of mass prosecution, which means prosecuting every suspected perpetrator, as long as they somehow belonged to or served the past regime, are not only the loss of human capital or financial resources. It also may stop police officers, public employees, and even political officials from taking responsibility and doing their job duties because they are afraid of being punished. This attitude entails minimizing the level of activity far under the optimal level while pushing the level of care to be far over the optimal level in regards to social welfare.

From these notes, I do not infer that the outputs of prosecution initiatives are entirely negative. I rather argue that different measures have to be selected for different criminals and crimes and that prosecuting should not be for everyone to avoid the previously mentioned hazards.[49] Such a balance should vary from one case to another.[50] In order to reach the optimal design for each case, the policy-makers will have to make a C.B.A.

For example, let us assume that the presented case involves a revolution over a regime that had been committing severe human rights violations towards the majority of its people for long decades to the extent that this policy became inherited in the state's deep system. Consequently, the benefits of prosecutions against members of that regime, including achieving accountability for the majority of the population and deterring any possible practice of this inherited policy, are overwhelming. In this case, if the benefits could possibly either weigh or equal the costs, the policy-makers could think of controlling the costs by minimizing the scale of prosecution. For example, in this case, policy-makers may decide to prosecute only the heads of the involved authorities in these violations while choosing less costly mechanisms to deal with the lower level of criminals who constitute the more significant number. In this case, fewer prosecutions limited only to major criminals can minimize the administrative, monetary, and time costs significantly. Moreover, decreasing the number of prosecutions avoids a wide-ranged polarization in the society, and also limits the error costs that multiply each time a new accused enters the prosecution circle through complicating the cases and adding new details to them.

Table (1) sums up the previous notes in a cost-benefit matrix to indicate the weighted factors:

[49] A similar conclusion was reached by Jeremy Sarkin, *The Tension Between Justice and Reconciliation in Rwanda: Politics, Human Rights, Due Process and the Role of the Gacaca Courts in Dealing with the Genocide*, 45 J. AFRICAN. L. 143 (2001).

[50] Customization upon the case is the same conclusion reached empirically by OLSEN ET AL., *supra* note 1.

COSTS	BENEFITS
Administrative costs	Deterrence
Time costs	Feeling of justice
Error costs	Building trust in the new regime
Incentivizing a backlash	Respecting the international legal standards
Social and political polarization and isolation	Achieving accountability
Politicizing the judiciary	Avoidance of costly direct interaction between the victim and the wrongdoers
Concerns of Unconstitutionality and Breach of Due Process Guarantees[51]	
Over or under-deterrence (depending on the adopted strategy)	

Figure 1: Table (1) Cost-Benefit Matrix of Prosecution Initiatives as a T.J. Mechanism

It should be noted, though, that the number of inputs in the matrix under the Costs and Benefits columns do not qualify to judge the weighing result. Each input of this shall have an equivalent number that refers to its weight depending on the studied case, as indicated earlier. The decision can be taken afterward by estimating the total output of this mechanism alone, on the one hand, to decide to what extent it should be applied. On the other hand, a comparison between the output of this mechanism and the other T.J. mechanisms' outputs could then lead to choosing some of them over others. This comparison can be made using models that qualify for the expected costs and benefits of every mechanism. This weighting or evaluation process cannot be constant, i.e., there cannot be a standard evaluation that will always lead to prevailing prosecutions over truth commissions, for instance, or limited prosecutions over medium-ranged prosecutions. The equation's entries as explained in the C.B.A. are probably constant in every case, but the number that shall be put on these entries to solve the equation, evaluate its output, and compare it to other potential outputs shall differ depending on the circumstances of each case. As mentioned above, this evaluation or quantifying process is to be done through multiple parties, including victims, experts, policymakers, etc.

[51] Besides the scenarios referred to earlier, the debate over the legality and constitutionality of the post-autocratic regimes is huge. *See* for example, Graver, *supra* note 35, at 144. Consequently, adopting a specific philosophy of what counts as "constitutional" will decide the weight of this cost.

Consequently, there is no way the process will not be subjective and independent of these parties' personal views. For example, some people would have a personal preference for "justice" considerations over "practical" considerations. However, the more parties are involved, and the more diverse the background they come from, the more likely this evaluation process can get closer to an accurate result that fits each given society's preferences and nature. This methodology shall also apply to weighing the other mechanisms.

1.2. FACILITATING INITIATIVES IN RESPECT OF THE RIGHT TO TRUTH

This category of mechanisms aims to assist post-conflict societies in investigating the truth behind human rights violations. This process is usually undertaken by truth commissions [hereinafter T.C.], which are "non-judicial or quasi-judicial investigative bodies, which map patterns of past violence and unearth the causes and consequences of these destructive events."[52] The mission can also be done by commissions of inquiry, or other fact-finding committees, which are similar to the truth commissions but usually operate under narrower mandates. Publishing the final results and recommendations is part of the process. Moreover, the documentation and archiving of the related evidence and materials help reveal the truth about the past, achieve the justice necessary for the transition, and keep the conflict's history.[53]

Priscilla B. Hayner gave in 1994 four primary elements for defining truth commissions: 1) they focus on the past; 2) they are not focused on a specific event but on investigating the violations of I.H.R.L. or I.H.L. over a specific period; 3) they are temporal and cease to exist by the submission of their report; 4) they are endowed with a sort of authority by their sponsor, which gives them the access to more information, protection, and impact for their report.[54] In 2011, she revisited these parameters and added a fifth element: 5) they engage broadly and directly with the affected population and gather information upon their experience.[55] In her new approach, Hayner emphasizes that what differentiates T.C. from other similar phenomena is that they intend to improve the social understanding and acceptance of the past events to influence the future, not just to resolve specific facts. The T.C.'s aim, according to

[52] United Nations, *supra* note 2, at 8.
[53] *Id.*
[54] Priscilla B. Hayner, *Fifteen Truth Commissions-1974 to 1994: A Comparative Study, in* 16 Hum. Rts. Q. 597.
[55] Priscilla B. Hayner, Unspeakable Truths: Transitional Justice and the Challenge of Truth Commissions 11–2 (2010). The other four elements were subject to marginal amendments that shall not change their nature or affect.

Hayner, is to change policies, practices, and relationships in the future in a manner that respects and honors the people who went under the past's experience.

It is argued that not only do T.C. help to find the truth behind crimes for accountability purposes, but also they give recognition to violations of the past, which is necessary to both the victims and the other sectors of the society who are pro the old regime. On the one hand, for victims, revealing the truth about the violations against them, confessing the crimes by perpetrators, and the interaction between victims and perpetrators that might also end with apologies, this whole process involved having a public voice for the victims, can have a healing effect for them.[56] On the other hand, when supporters of the past regime hear from both victims and regime-members the details and facts regarding the violations which they were subject to, this could make them understand the catalysts of the revolution and the necessity to T.J. process,[57] which mitigates their opposing behavior toward the transition process. Both these effects together not only help the society to make a clear break with the past but also to minimize the polarization between its forces and the probability of a counter-revolution. Similar reasonings motivated the formation of T.C. in many countries, getting out of violent conflict and crises, including South Africa, Haiti, Guatemala, and others.[58] However, these motivations do not necessarily turn to actual outcomes for many constraints that relate in part to the costs of T.C.s that will be explained in greater detail and to the difficulties of power-sharing in post-crisis societies generally. In the last three examples, concerns related to race biases in Guatemala, time and financial constraints, shortage of using qualified experts, limited publication of the final report in the case of South Africa, the past-regimes destroying most of the archive detailing their crimes, and the failure sometimes in formulating a clear conception of the "truth" these commissions are looking for assuming that it would be an automatic result of their mandate application, all constrained the complete achievements of these motivations.[59]

[56] Martha Minow, *The hope for healing: What can truth commissions do?*, *in* Truth v. Justice: The Morality of Truth Commissions 235 (Robert I. Rotberg and Dennis Thompon eds., 2010). ; Holly L. Guthrey, Victim Healing and Truth Commissions: Transforming Pain Through Voice in Solomon Islands and Timor-Leste (2015).

[57] Olsen et al., *supra* note 1 at 155.

[58] Audrey R. Chapman & Patrick Ball, *The Truth of Truth Commissions: Comparative Lessons from Haiti, South Africa, and Guatemala*, 23 Hum. Rts. Q. 1 (2001). Hayner, *supra* note 54.

[59] Chapman and Ball, *supra* note 58.

The economic rationale of the truth commissions and other truth revealing mechanisms is divided into four main directions:

1. T.C.s signal the acknowledgment of the victims' losses, the criminals' accountability, and the commitment to human rights and the rule of law to all the parties of the transitional phase: the victims, the past regime, the new regime, and the judicial authorities.

2. The truth commissions can have a minimizing effect on the polarization in the society between the supporters and oppositions to revolutions, and also the bystanders. When victims speak up about the violations they have been subject to, this gives a forum within which the other parties can, for the first time, acknowledge the victims' stories and their rights.[60] Consequently, through T.C., society can have a clean break with its past.

3. In case the truth commissions precede the prosecutions, they minimize the information costs for conducting these trials. This advantage also applies to institutional reforms as they give clear information about the past system, the mechanisms were used to achieve its hostile policies, and the personnel involved in these policies.

4. In case amnesties are given to the witnesses or perpetrators who give aid to these commissions, under the condition that they are not accused of gross human rights violations, this will influence the whole T.J. process in different ways. First, it will minimize both the error and administrative costs of the prosecution thanks to the information and cooperation they will provide; Second, it will incentivize the allies of the past regime and its supporters to shift positions and take the side of the new regime as they see that they can have a place in this new regime if they cooperate with it. To maximize this influence, though, amnesties' payoffs need to be guaranteed even for perpetrators who did not give information until a late stage, so they do not abstain from positively responding to such initiative even if they resisted in the beginning due to low trust levels.

[60] Olsen et al., *supra* note 1 at 155.

However, the truth commissions also have constraints that may limit their success. Priscilla B. Hayner lists these constraints,[61] and they include:

1. The mandate: The way these commissions are called for and formed is problematic. Because of the fragile political context and the time constraints, they are usually not formed through a sufficient public debate, discussion, or referendum, which contradicts many human rights advocations.[62] This reflects a principal-agent problem[63] between the public and the policy-makers who run these commissions, where the information between the two parties is always asymmetric, and their preferences might align or not. This situation may lead to various consequences. Among these possibilities, there could be paternalistic behavior by the policy-makers that ignore the people's desires or expectations because the firsts see them as irrational or inefficient. There also could be a rent-seeking behavior of the policy-makers that disregard the original T.J. goals and abuse the people's ignorance of the sophisticated details or techniques of the mandate to achieve personal or elite payoffs. Another scenario could be heavy pressure from the public opinion to eliminate any form of amnesties even if the process could not be completed without an incentive for past perpetrators to cooperate. The possibilities go on.

2. The political constraints: Despite what has been referred to in the benefits of the truth commissions as a minimizer of the polarization in the society between pro and against past-regime sides, the later reveal of the truth can also work in the opposite direction. There is a risk that revealing the crimes heats the hatred against the members of that regime and, consequently, makes integrating them into the new regime more difficult. The answer to the question in which direction this mechanism will work can only be given through future empirical studies. The political constraints may also include the desire to preserve the image of some of the perpetrators mentioned in these investigations. In this case, the policy-makers' payoffs from keeping at least part of the truth hidden are higher than the payoffs of signaling the new regime's control to the past one.

3. Restricted access to information: Some of the information needed for T.C.s to achieve their goals may be sensitive to national security, destroyed by competent persons, or blocked because of a conflict of interests. In any case, information costs can be too high.

[61] Hayner, *supra* note 53.
[62] *Id.*
[63] For more on agency problems in the law and economics sphere, see Parisi, supra note 24, at 5–9.

4. Lack of resources: The process of investigation and collection of past violations is financially costly. Not all nations have the necessary financial, human, and time resources necessary for achieving such a process. In other terms, administrative costs can be quite substantial. This obstacle may be overcome through international aids, and there are already comparative experiences that involved so. However, accepting these aids and making the best use of them depends on the country's international relations and relevance, the acceptance of the national parties of receiving international aids and involving international parties in the national T.J. process, and how far the international experts can understand the national context they aim to help within and communicate effectively with its parties.

The interaction and sequence between prosecutions and truth commissions are problematic and complicated. While the first aim at imposing criminal sanctions on the guilty accused, the second aim at recording the truth and acknowledging it, to keep a record of the history, learn from it, and make sure that the victims are heard. They usually complement each other; however, in some cases, the truth commissions play as a substitute for prosecutions in case they were prevented because of amnesties or by force.[64] The difficulty is, however, to find the best sequence of them. Should trials precede truth commissions or the opposite, or should they be simultaneous? Alexander Dukalskis presents these three possible scenarios and the pros and cons of each of them.[65] The output of each scenario depends on the different contexts of every transition. Scenario (1) is when the truth commissions precede the prosecutions; Scenario (2) is when prosecutions precede truth commissions; Scenario (3) is when they work simultaneously. Depending on the literature's concerns and remarks on the truth commission's performance generally and their interaction with trials in specific, in addition to the precedent normative analysis, one can draw a picture of the pros and cons of each scenario. I will integrate each scenario's expected costs and benefits in the cost-benefit analysis of the truth commissions.

From the previous analysis, a cost-benefit matrix of the truth commissions can be presented as follows in table (2):

[64] Douglass W. Cassel & Jr., *International Truth Commissions and Justice*, *in* Transitional Justice; Volume I: General Considerations (Neil J. Kritz ed., 1995); Alexander Dukalskis, *Interactions in Transition: How Truth Commissions and Trials Complement or Constrain Each Other*, 13 Int'l. Stud. Rev. 432 (2011); Hayner, *supra* note 54.

[65] Dukalskis, *supra* note 63.

General Costs			General Benefits		
Time costs			Signaling commitment to the rule of law		
Administrative costs			Archiving the history		
Information costs			Achieving accountability		
Political polarization			Adherence to international standards		
Standard error costs			Reducing the costs of institutional reforms		
Costs by Scenario			Benefits by Scenario		
Scenario 1	Scenario 2	Scenario 3	Scenario 1	Scenario 2	**Scenario 3**
Error costs by providing amnesties before prosecutions	Probability of lack of incentive to cooperate in case the accusations were declared	Too high adm. costs for managing the conflict between the two mechanisms	Minimizing adm. costs of the prosecutions	Probability of incentivizing the relevant perpetrators to cooperate in case of conviction of the principal perpetrators + offering amnesties or reduced sanctions	Minimizing the costs of the other scenarios
	Multiplying the time costs		Minimizing info. costs of the prosecutions	Minimizing adm. costs of truth commissions in respect to the trialed violations	
	Higher info. costs in respect to the untrialed violations because of the time gap			Minimizing info. costs of the truth commissions in respect to the trialed violations	

Figure 2: Table (2) Cost-Benefit Matrix of Truth Commissions as a T.J. Mechanism

The general costs and benefits refer to the variables present in any mechanism of truth commissions regardless of its type of combination with prosecution initiatives. The costs and benefits by scenario refer to, in addition to the previous general variables, the costs and benefits of the truth commissions that change depending on the selected scenario. Just as in the prosecution initiatives, these variables are abstract in this analysis, and they would take different values depending on the party quantifying them and their

expected payoffs. Consequently, the more multilateral the process is, the more successful it is expected to be.

Note also that the "standard error costs" refer to the margin of error possible in any mechanism, as different from the "error costs," referring to the errors in judicial decisions.

1.3. DELIVERING REPARATIONS

The reparation programs seek to redress the victims of human rights violations through material or symbolic benefits.[66] The right to reparation of the victim in case of violations of I.H.R.L. and I.H.L. is a well-established principle under international law.[67] This reparation may take different forms.

The forms of reparation include: 1) *Restitution* which aims at restoring, to the victim, the situation before the violation, to the possible extent; 2) *Compensation* which provides financial recognition for the purpose of redressing violations; 3) *Rehabilitation* which entails providing services for the victims to restore their dignity, health, and reputation, including legal, medical, and psychological care services; 4) *Satisfaction* which aims at restoring the dignity and reputation of the victims through different measures like judicial decrees or official declaration, public apology, the acknowledgment of violations against the victims, commemorations, and tributes to the victims; 5) *Guarantees of non-repetition* by ensuring effective civilian control over the military and the police, and the obligation of the different parties by the international legal standards of due process and fairness.[68]

Reparations as a T.J. mechanism have the same economic reasoning as tort law with respect to internalizing the harm caused by the wrongful party to the victim and inducing other persons to invest in taking cautions to prevent such harm. It is also a way through which law manages the bargaining between parties who have relationships with relatively high transaction costs, taking into consideration the possible difficulties of direct bargaining between the victim and the wrongdoer.[69] These difficulties are

[66] United Nations, supra note 2, at 8–9.

[67] *See generally* Theo Van Boven, *Study Concerning the Right to Restitution, Compensation and Rehabilitation for Victims of Gross Violations of Human Rights and Fundamental Freedoms*, *in* Transitional Justice; Volume I: General Considerations (Neil J. Kritz ed., 1995).; Lisa J. Laplante, *The Plural Justice Aims of Reparations*, *in* Transitional Justice Theories 66 (Susanne Buckley-Zistel et al. eds., 2014). ; Luke Moffett, *Transitional Justice and Reparations: Remedying the Past?*, *in* Research Handbook on Transitional Justice: Research Handbooks in International Law Series 377 (Cheryl Lawther et al. eds., 2017;Conor McCarthy, Reparations and Victim Support in the International Criminal Court (2012).

[68] United Nations, *supra* note 2, at 9; Van Boven, *supra* note 67, at 548–9.

[69] *See* Robert Cooter & Thomas Ulen, Law & Economics 189–90 (2011).

supposed to be more intense in the case of a transition from an autocratic rule that involved mass violations against the victims' human rights.

However, in many cases, reparations cannot, alone, achieve the internalization and the avoidance of the costly direct negotiation between the victim and the wrongful party, for two main reasons:

1. By their nature, as mentioned earlier, reparations include other T.J. mechanisms as well, like investigating the truth and Institutional Reforms. For example, the fifth T.J. mechanism, which requires the guarantee of non-repetition, will necessarily entail an institutional reform.

2. Most importantly, some damages either cannot be assessed, or their assessment costs will be significantly high compared to the costs of other mechanisms. For example, what could be the correct way to assess the harm caused by a corrupted election's authority that was in power for several years or even decades? Moreover, in case of an error in the reparations' allocation, this will result in an inefficient allocation of the already significantly limited resources because of the phase nature, as indicated earlier.

Additionally, the more powerful the past regime, the longer it ruled, and the more severe and common its violations, the more victims it should have, and the more resources are needed to cover their redress, especially if the financial costs of these reparations are not covered by the past-regime members.

However, in case the state will be responsible for delivering the reparations, rather than the wrongful persons directly as a consequence of their criminal liability, the reparations mechanism has a competitive advantage. The probability of proving only the damage that the victims suffered from is significantly higher than the probability of proving the criminalized acts against specific persons and the causation linkage between these acts and the damage. Many criminal evidence rules and criminal law procedures are lifted in case of seeking the proof of damage rather than seeking the criminal liability of specific persons. This advantage is especially significant in big-scale crimes, which affect a wide range of victims and crimes in which the probability of the proof of criminal liability is low. The United Nations Declaration of Basic Principles of Justice for Victims of Crime and Abuse of Power states under its section addressing the "Victims of Crime", that:

> 2. A person may be considered a victim, under this Declaration, regardless of whether the perpetrator is identified, apprehended, prosecuted or convicted and regardless of the familial relationship between the perpetrator and the victim. The term "victim" also includes, where appropriate, the immediate family or dependents of the direct victim and persons who have suffered harm in intervening to assist victims in distress or to prevent victimization.[70]

However, there will still be costs for proving the harm that happened to the presumed "victim," beyond this there are potential erroneous costs in case someone was proved to be a victim, or vice versa; a standard error in any possible mechanism.

This separation between the criminal conviction and the reparations could also be an effective way to internalize the harm of the crimes with less polarization in the society because it does not require holding specific persons accountable. Moreover, it saves human and financial capital that is needed for prosecutions.[71] However, the value of these advantages depends on the source of these reparations. Is it covered from the state budget or by international actors, or through payments that the wrongdoers secure in exchange for guarantees given to them to escape the criminal liability or the possibility of political isolation? In both cases, there will be costs.

This issue brings us back to the error costs point. Besides the standard error costs, there is a form of reparations that could involve an additional margin of error. This form is the suspension of the custodial or financial sanctions against the members of opposition to the past-regime. The assumption that whoever was prosecuted by the past regime for a crime that has a political aspect, like: terrorism, violence crimes, attempt to change the regime (in the systems that have such a crime), membership in an illegal organization, or any other form of a relevant crime, can be either proven or falsified. However, given the time element and the high administrative costs, this can be either highly costly or, in some cases, impossible. In case the sanctions against these "victims" are not suspended or re-evaluated, this disturbs the goals of the whole process and adds new enemies to the new regime. In contrast, in case the sanctions were mistakenly suspended, the harm that can be caused by these "victims" can be very severe. For example, there are cases where members of past illegal organizations were subjects of amnesties in the context of T.J. processes, who were proven later to have partaken in terrorist behavior.

[70] *See* United Nations: Declaration of Basic Principles of Justice for Victims of Crime and Abuse of Power; General Assembly Resolution 40/34 (November 29, 1985), , *in* Transitional Justice; Volume III: Laws, Rulings, and Reports 646 (Neil J. Kritz ed., 1995).

[71] Note that in this part there is no comparison between reparations and prosecutions as two mechanisms of T.J., but an explanation of the possible advantages of separating reparations from criminal liability.

The Egyptian case could be a useful reference in this regard.[72] In other words, the lack of due process towards the past-regime combats is not by itself a guarantee that they are entitled to a declaration of the accusations against them or of other possible accusations. This is a probable error-cost related only to this form of reparations. I call this the "False Benefit of Doubt Error."

Table (3) shows the C.B.A. of reparations as a T.J. mechanism:

Costs	Benefits
Administrative costs	Possibility of eliminating part of the proof costs (in case of comparison to mechanisms that require proof, like prosecutions or lustration)
Assessment barriers	Inducing an optimal level of caution and activity
Standard error costs	Internalizing the harm to the victim
Probability of a significant number of victims	Adherence to the international legal standards
Financial resources limitations	Avoidance of costly direct interaction between the victim and the wrongdoers
Probability of false benefit of doubt error	

Figure 3: Table (3) Cost-Benefit Matrix of Reparations as a T.J. Mechanism

Note that this C.B.A. represents the case where reparations are granted on the basis of the harm, as explained earlier, and not based on a criminal conviction or tort liability. Consequently, no probability of error costs is added to the costs because reparations, in this case, are not dependent on other judicial decisions other than decisions related to the reparation itself, if any. Thus , the error that could happen is just the standard error in any mechanism. However, if the reparation mechanism was designed in a way that requires a judicial trial that identifies and convicts a perpetrator, then the judicial procedures' error costs should be included in the calculation as well, among the other costs and benefits of the prosecutions.

[72] *See* Hossam Bahgat, *Who Let the Jihadis Out?*, Mada Masr (Feb. 16, 2014), https://www.madamasr.com/en/2014/02/16/feature/politics/who-let-the-jihadis-out/.

1.4. INSTITUTIONAL REFORM

The institutions involved in causing the conflict by breaching human rights have to be transformed into institutions that respect these rights and the rule of law's values to prevent the recurrence of these violations. This transformation should include both lustration and training on applying human rights law and humanitarian law standards.[73]

Institutional reforms are a mechanism that can have a broad or limited meaning. Its limited one refers to lustration, purges, and vetting, which are processes that aim to change the corrupted personnel existing within the state institutions. Its broad meaning refers to changing the state institutions in a way that prevents the repetition of the past violations; in other terms, the repetition of policies and state actions that represent the same philosophy and approach of the past regime. In the case of revolutions over autocratic regimes, this includes human rights violations and autocratic behavior. Under this interpretation, constitutional change, structural change of the authorities' performance, mentoring institutions, and any other effort to stop the culture and the institutional cycle that produced the past, are all considered as part of T.J. Although the second meaning sounds more crucial to achieve T.J. goals, it is hard to be specified in specific policies, and it is too broad to be tackled within these research limitations. Consequently, I here adopt the narrow meaning of institutional reforms.

Because of its legal complications and mixed-blessings, lustration is usually the most problematic aspect of institutional reforms; hence , it will be the focus of this sub-section. Lustration is "the disqualification and, wherein office, the removal of certain categories of office-holders under the prior regime from certain public or private offices under the new regime."[74] Lustration has many complications and various advantages and disadvantages. I will first discuss its legal complications and then present the economic reasoning behind lustration laws and their social costs and benefits.

Some of the legal complications of the application of lustration are the typical challenges of prosecution initiatives referred to earlier, including retroactivity, hardships of collecting sufficient evidence, and other due process considerations.[75]

[73] United Nations, *supra* note 2, at 9.

[74] Herman Schwartz, *Lustration in Eastern Europe, in* Transitional Justice: How Emerging Democracies Reckon with Former Regimes. Volume I: General Considerations 461, 461 (Neil J. Kritz ed., 1995).

[75] *Id.*; *See* Susanne Y. P. Choi & Roman David, *Lustration Systems and Trust: Evidence from Survey Experiments in the Czech Republic, Hungary, and Poland* 117 Am. J. Socio. 1172 (2012).

However, besides these complications, there are other specific legal challenges of applying lustration laws.[76] These challenges include:

1. Considerations of inequality and discriminatory treatment as a cause of the breach of I.H.R.L.[77] and unconstitutionality of lustration laws accordingly;

2. The scope and basis of lustration or vetting processes.[78] For example, dismissing the collaborators of the past regime is faced with problems of collective responsibility and guilt, and the legality of the orders given to them by the previous legal system;[79]

3. In case lustration was applied directly through the legislature or an executive act without prior prosecutions, this could be challenged in a court as a denial of the right to a fair hearing guaranteed by I.H.R.L.;[80]

4. Lustration laws may also violate individuals' right to work.[81]

These considerations, however, could be countered by the following arguments:

1. The right to equal treatment and to work guaranteed by I.H.R.L. and most of the constitutions do not prevent the state from applying sanctions on its citizens or listing specific conditions for some of its offices;

2. These mechanisms can be justified by the protection of the state against threats of its order;[82]

3. Lustration laws as administrative or labor laws have privilege over the criminal laws used in the ordinary prosecutions. This advantage is the application of the "Presumed Liability" principle, or *"responsabilité objective / responsabilité sans faute"* as known originally and established by the French state council and legal

[76] Beside these challenges, a typical argument against lustration laws is that some people were just following the orders, especially the minor officials. There is a huge debate on this point. For more on this, *see* Graver, *supra* note 36, at 148; Posner & Vermeule, *supra* note 35, at 778. Whatever side of the debate one can take, this still shows how it could be better to lustrate only the main officials, not the minors as well.

[77] *See* Roman Boed, *An Evaluation of the Legality and Efficacy of Lustration as a Tool of Transitional Justice*, 37 Columbia J. Transnatl. Law 357, 357–402 (1999).

[78] Lustration and vetting are usually used as synonyms in the literature. Although they are both forms of personnel institutional reforms after crises, there is a scholarly work on differentiating between them on the base that the last is more general than the first. According to this differentiation, the lustration policies are more gravitated to the extra-regional cases. For more on this, *see* Cynthia M. Horne, *Transitional Justice: Vetting and lustration, in* Research Handbook on Transitional Justice 424 (Cheryl Lawther, Luke Moffett & Dov Jacobs eds., 2017).

[79] Schwartz, *supra* note 74, at 463–4.

[80] Boed, *supra* note 77.

[81] *Id.*

[82] *Id.* at 399.

literature, on the leading officials of the past-regime. This principle dispenses with the "fault" element of proving the liability and takes into account only the elements of "harm" and "causation."[83] Consequently, it could be an effective strategy to solve the low probability of collecting sufficient evidence in these cases. The principle of Presumed Liability is originally used for establishing monetary damages. However, up to my knowledge, there is no legal barrier to apply another aspect of the administrative liability, which is the qualification conditions of the state offices, as long as it is not a penal sanction, to avoid the unconstitutionality concerns based on the breach of the principle of the "Presumption of Innocence."[84]

4. Finally, the rulings, decisions, and opinions of the competent international legal bodies, mainly the European Courts of Human Rights, International Labor, and the Office of the United Nations High Commissioner for Human Rights [hereinafter O.H.C.H.R.] in lustration cases are not anti-lustration per se. They forgo the challenges of retroactivity, discrimination, and employment barriers, as a part of the whole democratization structure. As referred to earlier, they also advocate for protecting and interpreting the rule of law and justice considerations in their historical context and accordingly, favoring the compelling interests of achieving justice and strengthening the new democracy, in such cases of exception like transitioning from crises or autocratic regimes. The problem with lustration laws, according to these rulings, is not with the philosophy of these laws, but with their implementation.[85] There should be legal guarantees for lustration processes to be aligned with international legal principles. These rulings, added to other literature on lustration as a mechanism of T.J.,[86] advocate for two essential considerations to be taken into account to avoid the costs and illegal concerns of lustration:

   (1) The guarantee of fair hearing, due process, legal certainty, and clarity;

   (2) The individualization of the process, which means that a differentiation should be made between the different officials of the past regime. According to that differentiation, the lustration decisions should be made after a proper investigation, hearing, and evaluation of the relevant official's fault and his/her possible threat to the new order. In other terms, an evaluation of

[83] *See* Raed Mohamed Adel Bayan, *Al Asas Al Qanooni Lelmasooleya Al Edareya Bedoon Khataa : Derasa Moqarna [The Legal Basis for the Administrative Responsibility without a Mistake : A Comparative Study]*, 43 Dirasat Shari a Law Sci. 289, 289–304 (2016), http://platform.almanhal.com/CrossRef/Preview/?ID=2-90655.

[84] For more about the "Presumption of Innocence," *see* Kenneth Pennington, *Innocent Until Proven Guilty: The Origins of a Legal Maxim*, 63 The Jurist 106 (2003).

[85] Horne, *supra* note 42; O.H.C.H.R., *See Rule-of-Law Tools for Post-Conflict States: Vetting: an operational framework* (2006), https://www.ohchr.org/Documents/Publications/RuleoflawVettingen.pdf.

[86] Boed, *supra* note 77; Choi & David, *supra* note 75; Schwartz, *supra* note 74.

whether the costs of their violations could be cost-justified. However, is this possible? This takes us to economic reasoning.

The question then remains: why do states need to apply lustration processes in the first place? Unlike other institutional reforms, lustration and similar measures have a punishment aspect that is in the core basis of the critics mentioned above. Lustration policies not only reform the institutions but also punish the personnel involved in past violations by preventing them from the benefits of one of their constitutional rights, i.e., the right to hold public offices. Consequently, a significant aspect of the lustration's rationale could be found in the original debates over the reasoning, effectiveness, and efficiency of punishment per se. The economic reasoning of crime and punishment is subject to lengthy debates.[87] Some scholars reason the notion of punishment by the external cost that crime causes, which needs to be redressed.[88] Others give more attention to the unconditional deterrence that punishment should achieve to force criminals not to substitute a market transaction.[89] However, besides criminal laws' job in pricing the harm of the crime to internalize it and achieving deterrence to other people to not commit the same crimes, some actions are criminalized to prevent their repetition by the same perpetrators.[90] Accordingly, in some cases, the aim of the punishment is not the typical mission of public deterrence or pricing, but to prevent the repetition of the same crime by the same persons through their elimination from positions they can use to cause the harm, or in other words, the incapacitation of the criminal.[91] The literature refers to this distinction as specific deterrence vs. general deterrence.[92] An example of that is imprisoning a criminal, forcing him/her to stop their criminal activity.[93] This reasoning applies to the category of crimes, of which a small proportion is socially cost-justified. An example of a socially cost-justified crime is theft under the dire of deadly hunger. Nevertheless, this is a low probability and might be non-criminalized at all. So, the smaller the proportion of such a socially cost-justified crime is, the smaller the social cost is , and the higher the social benefit of its prevention is, and that is why such crimes tend to be prevented instead of priced.[94]

[87] Veljanovski, *supra* note 5, at 241–262;Samuel Cameron, *The Economics of Crime Deterrence: A Survey of Theory and Evidence*, 41 Kyklos Int'l Rev. Soc. Sci. 301 (1998).
[88] Becker, *supra* note 7.
[89] *See, e.g.*, Richard Posner, *Criminal Law, i*n Economic Analysis of Law 273 (9th ed. 2014).
[90] Posner, *supra* note 29; *See also* Charles H. Logan, *Deterrence and Incapacitation: Estimating the Effects of Criminal Sanctions on Crime Rates*, 9 Contemp. Socio. 389 (1980).
[91] A. Mitchell Polinsky & Steven Shavell, *The Theory of Public Enforcement of Law, in* Handbook of Law and Economics 403.
[92] For a review, *see*Mark C. Stafford & Mark Warr, *A Reconceptualization of General and Specific Deterrence*, 30 J. Rsch. Crime & Delinq. 123 (1993).
[93] Imprisonment as an incapacitation measure is, though, debatable. For more, *see* Isaac Ehrlich, *Crime, Punishment, and the Market for Offenses* J. Econ. Persp., Winter 2996, at 43.
[94] Posner, *supra* note 29, at 1214–5.

the common crimes involved in T.J. processes, including genocide, torture, corruption, and other human rights violations, be cost-justified? The cost of these crimes is too high that only a small proportion of them could be cost-justified, and then prevention via lustration can be argued to be more efficient than just punishing the act if it repeats.

Incapacitation can be achieved through different mechanisms, not only imprisonment but also by nullifying a license to perform a specific activity, for example.[95] In the case of T.J., the same reasoning applies to the rationality behind lustration laws. It is not about internalizing the harm of the crimes, because this can be done through reparations. It is also not only about deterrence, because this can be achieved through criminal prosecutions. However, the main reasoning is preventing the same corrupt persons or the persons who adopted past-regime strategies from corrupting or disturbing the new regime.

The question of the social cost and benefit of such laws and whether incapacitation should be executed through lustration or imprisonment, or even the death penalty, is however, subject to the calculations of every different society, depending on other variables. As presented by A. Mitchell Polinsky & Steven Shavell, the basic equation is that the harm caused by the prospective criminal is larger than the cost of his/her incapacitation.[96] Some studies find that this harm is reduced with age,[97] i.e., the older the criminal is, the less is the harm expected to be caused by him/her to society if he/she is incapacitated. This tendency could be due to the lifetime left for his/her potential criminal activity or health reasons. In our case, though, it can be argued that it is reduced by time, given that the more time passes since a transition, the new regime increasingly consolidates, and consequently, the ability of the past-regime members to spoil the transformation and repel the new rules decreases. Accordingly, the incapacitation should continue as long as the cost of the harm caused by the relevant member of the past regime is more than the cost of his/her incapacitation. These costs are to be assessed by the policy-makers, with technical experts' help, at the given point in or after the transition. For the side of the equation reflecting the cost of the harm caused by the considered member, there should be:

- A variable standing for a rough estimation of the costs resulted from his/her past behavior divided by his/her years of service to reach an estimation of these costs per year. The word "rough" is used because many of these past-violations costs are

[95] *See* Polinsky & Shavell, *supra* note 91, at 443; *see also* James L. Nichols & H. Laurence Ross, *The Effectiveness of Legal Sanctions in Dealing with Drinking Drivers,* 22 J. Safety Rsch. 117 (1991).
[96] *See*McNollgast, *The Political Economy of Law*, *2 in* Handbook of Law and Economics Volume 2, 1651-1738 (2007).
[97] Polinsky & Shavell, *supra* note 91, at 443.

hard to estimate and have long-lasting domino effects. Consequently, what would be feasible is assessing only the direct impact of the rent-seeking activity; for example, evaluating the financial losses caused to the national treasury because of allocating lands to well-connected investors without fair compensation to the state. Some harms would be easier to evaluate than others, and the evaluation process itself is costly. That is why these calculations are not widely common.

- The previous variable would be multiplied (1) with the probability of repeating the past behavior given the extent of the new-regime consolidation and control over the authority where this member operates. (The variable would be multiplied (2) also with the number of years remaining in his/her service until retirement.)

For the equation's side concerning the cost of incapacitation, the exact general costs of lustration indicated in table (4) apply (G). However, an added variable would be the probability of the availability of an efficient member of the new regime to replace the considered past-regime member(P_o).
Accordingly, to apply incapacitation, $C_y \times P_r \times N$ should exceed $G - P_o$.

However, Richard A. Posner refers to a critical qualification that should be entered into the calculation, which is the effect of the offenders' elasticity of supply.[98] If this elasticity is "very high," the effect of taking one criminal out of the scene would be making room for another criminal to get in.[99] This theory may not apply to the crimes committed amid the revolution incidents because of the mass violation and chaotic nature of the situation, which influences a rational actor's standard calculation. However, taking into account the high payoffs that politicians can expect of the rent-seeking activity manifested in autocratic policies, this could be possible in the case of T.J. also. Lustration policies are, indeed, ex-post arrangements dealing with the old-rulers; they will be removed from power and replaced by new-rulers adopting the revolution's goals. However, lustration alone does not guarantee the non-repetition of the past. The new-rulers themselves, or the officials who were removed from their positions in the first place, could still have tempting incentives to reproduce their own autocracy, especially that this behavior would be *per se* in their self-interest. This rationale justifies the combination of this mechanism between "lustration" and "training" under one mechanism labeled "institutional reform". If it is only for removing the past regime's criminals, there will be no effective impact of this mechanism. The aim is to prevent the repetition of the crime, or reducing its probability to be more specific, because a zero-crime probability would be too costly[100] and reducing the supply of

[98] Posner, *supra* note 29.
[99] *Id.* at 1217.
[100] *Id.* at 1215.

criminals. The way to do so is by accompanying the lustration laws with training, policy reforms, and monitoring that could guarantee the transformation of the institutions themselves.

On the one hand, the use of lustration laws may also result in a loss of human capital by diminishing the dismissed persons' technical and administrative expertise.[101] The more persons lustration is applied to , the more is the loss of human capital, and consequently, the weaker the institutions may become; thus, the higher is the probability of polarization in the society and an organized backlash by the dismissed officials. These drawbacks could apply to other mechanisms as well, like criminal prosecutions. As mentioned earlier, the weighing of the effectiveness, cost, and benefits will differ by case depending on other variables. For example, the range of members targeted by lustration systems and/or prosecutions.

On the other hand, keeping the key past-regime officials increases the probability of the repetition of their past performance and leads the people to suspect the transparency and loyalty of the new regime to the values of the revolution, i.e., the distrust in the prior regime because of its corrupted policies could spill over to the relationship between the people and the new-regime.[102] Some systems, however, tried to escape this dilemma by adopting what is called "lustration systems," which differentiate between three strategies of lustration: dismissal, exposure, and confession. Dismissal aims to purify the government by sacrificing its "tainted officials", but by demeaning them socially; exposure aims to increase the transparency of the government, but inadvertently stigmatizes its "tainted officials" socially; confession is an act of self-purification by the "tainted officials" to be "morally re-born" under new conditions [103]. However, a systematic empirical study is still lacking to measure these mechanisms' effectiveness in achieving the original purposes of lustration laws.[104] For example, although they can save the loss of human capital, their impact on preventing new corruption crimes can be questionable.

Gordon Tullock briefly referred to institutional reforms as one of the possible smooth democratization processes out of autocracy.[105] He argues that when a tyranny thinks of retirement, most probably, the main reason that would make him abstain

[101] Schwartz, *supra* note 74, at 464.
[102] Choi & David, *supra* note 75, at 1173–4.
[103] *Id.* at 1173.
[104] For experimental evidence from Hungary, the Czech Republic, and Poland on the different effect of these three strategies on trust *see* Choi & David, supra note 75. However, a general theory can't be derived from such a study. Moreover, it measures only the effect on trust, while other purposes of the lustration laws are not discussed.
[105] *See* Gordon Tullock, *Revolution and Its Suppression; "Popular" Uprisings, in* The Selected Works of Gordon Tullock (Volume 8): The Social Dilemma of Autocracy, Revolution, Coup d'Etat, and War 219 (2005).

would be worrying about his safety after retirement. Consequently, a good strategy for him would be creating a democratic constitution and holding elections to choose a new government. Tullock thinks that in this case, the new government will both be grateful to the past tyranny and too busy sorting the new system out to harm him, which would be the best case for the autocrat. However, most of the examples of this approach in South America involved a reversal back to autocracy after a few years.[106] The reason, in my opinion, is that the case which Tullock presented is neither a case of revolution nor of democratization. The reason is that there is only one way for the autocrat to be sure that his own autocratic government will not turn behind him and attack both him and the new democratic government, defending their benefits from the autocratic regime: namely by ensuring that the new democratic government being a new version of the same past officials and beneficiaries. In this case, although this approach is a safe exit for the autocrat and his government that could save the country lots of blood and costs, it does not represent any form of real institutional reform. First, the past abuses were not recognized, and no one asked for forgiveness or promised the non-repetition of these violations. Second, it is probably a mere change in the *de jure* without a change in the *de facto* application of constitutional democracy principles. Consequently, any thinking of the safe exit principle for autocrats to de-incentivize them to hold on to power should also include two necessary conditions: a genuine institutional reform, ridding the dictator of his top officials so they may not reproduce his government; and recognition of his violations in return to giving him, and probably his family, a blanket amnesty. This definition of institutional reforms entails lustration policies against the first line personnel, who usually include the governing party's directory board, the cabinet, and heads of the leading state authorities, e.g., the parliament's head. Although these persons , in addition to the autocrat, do not perform alone, and they are usually connected to a complex and broad net of beneficiaries and collaboratives, they are the most critical because they control this network. Consequently, minimizing their influence could ease controlling the rest of the system.

Upon the previous legal/economic analysis, the costs and benefits of lustration will depend on the adopted legal system. However, to build only a basic matrix of the cost-benefit analysis, I will assume that the previously indicated arguments counter the legal considerations and that the two broad legal guarantees required by the international rulings and literature are applied. Accordingly, the cost-benefit analysis would be as follows in table (4):

[106] *Id.*

Costs	Benefits
Administrative costs	Recovering the trust in the government
Error costs	Adherence to international standards
Loss of human capital	Deterrence
The threat of backlash by the past regime and its collaborators	Prevention of repetition of the crimes (depending on combination with training, new policies, and monitoring)
Polarization in society	
Constitutional considerations	

Figure 4: Table (4) Cost-Benefit Matrix of Lustration as a T.J. Mechanism

Note that the probability of adopting complementary mechanisms of institutional reforms, i.e., training, new policies guaranteeing the rule of law and I.H.R.L., and monitoring, is added as a prerequisite to the benefits of lustration generally. The reason is that other benefits, including deterrence, trust, and international standards, depend on them too.

1.5. NATIONAL CONSULTATIONS

This mechanism involves public participation in laying down the principles and mechanisms of transitional justice and interaction with their application. This participation also includes the necessity of the outreach of transitional justice, its knowledge, details, and implementation in society.[107] The mechanism then takes two ways: The first is integrating the public's feedback into the T.J. mechanisms' design, and the second is outreaching to the public to teach them about T.J. mechanisms.[108]

It has been argued in the literature and the analysis provided above that the design and selection of T.J. mechanisms differ and depend on the case and the context in which these mechanisms are applied. The more these mechanisms are context-oriented, the more they are likely to succeed. National consultations are a way through which the T.J. policy-makers can seek a better understanding of this context.[109]

[107] United Nations, *supra* note 2, at 9.
[108] Anna Triponel & Stephen Pearson, *What Do You Think Should Happen? Public Participation in Transitional Justice*, 22 Pace Int'l L. Rev. 103, 103-14 (2010).
[109] Phuong Pham & Patrick Vinck, *Empirical Research and the Development and Assessment of Transitional Justice Mechanisms*, 2 Int'l J. Transitional Just. 231 (2007). Triponel & Pearson, *supra* note 107.

In other terms, it contributes to solving the principal-agent problem between the public (the principal), including the victims and the policy-makers (the agents), through two means. First, involving the agents in the design of the T.J. mechanisms; second, providing more information in the market on both sides, i.e., policy-makers could know more about the needs and priorities of the victims and concerns of the pro-past regime or neutral parties, and the victims and other involved parties could understand more about how the T.J. mechanisms work and at what cost. This "low information cost"-effect that national consultations cause also minimizes other possible problems in the market, including a gap between the supply and demand of T.J. mechanisms and concerns of paternalism practiced by the policy-makers. Moreover, the integration of the public and relevant actors in the negotiation phase increases the probability of their cooperation in the implementation phase.[110]

The techniques through which national consultations are practically applied vary. They may include:

1. Doing empirical research to assess the needed mechanisms and the preferences of the public. The effectiveness of these researches and their methods is presented in a study by Phuong Pham & Patrick Vinck.[111] This empirical research may be quantitative, qualitative, or through mixed methods.[112] This is a way of using the scientific research methods, which are distinguished by being objective, to reach the most possible accurate estimations of the public needs, concerns, and preferences. Better results could be expected when these studies are run by neutral scientific institutions instead of governmental bodies;

2. In case T.J. laws are incorporated in the interim constitutions, subjecting them to national consultations and popular approval over the constitution as general. The first, in the drafting phase, happens through deliberations among the constituent assembly, which is either elected or formed by an elected authority, and the political elites generally. The public also takes part in this phase not only through electing their representatives but also through civil society organizations, national bars, media polls, and other common methods of involving society in the constitution-making process. The second process is usually conducted via national referenda;

[110] *See* Triponel & Pearson, *supra* note 108.
[111] Pham & Vinck, *supra* note 109.
[112] *See* Office of the United Nations High Commissioner for Human Rights, Rule-Of-Law Tools For Post-Conflict States: National consultations on transitional justice (2009).

3. Integrating national actors, civil organizations, non-governmental authorities, and the other relevant actors in the discussions over the T.J. mechanisms' design,[113] and the implementation and follow-up phases afterward. For example, in some cases, these entities can collect information from victims regarding their expectations from the reparations programs and communicate these expectations, victims' numbers, and their cases' details to the authorities handling T.J.. Grass-root and regional organizations are specifically helpful in this regard because of their familiarity with some levels, regions, and contexts that the high-level policymakers may not be specifically familiar with. This method was applied recently in Tunisia,[114] for instance, the only system that completed a T.J. process among the first-wave Arab Spring cases.

4. Facilitating open-access national hearing sessions regarding T.J. policies, whether in-person or aired on TV and radio or online. In addition, the media coverage of the T.J. processes, and the cultural and educational activities undertaken by the state or the non-governmental institutions, are all forms of integrating the public into the process which promotes the national ownership of it.

The two possible costs of national consultations are: First, increasing the time costs of the T.J. process by increasing its length, before and after its start; second, adding extra administrative costs to the process. Table (5) shows a matrix of the costs and benefits of National Consultations as a T.J. mechanism:

Costs	Benefits
Extra administrative costs	Minimizing the info-gap between the victims and the policymakers
Extra time costs	Raising awareness among the public
	Adherence to international standards
	Facilitating the implementation of the other mechanisms (high probability of cooperation)
	Reducing the probability of risk in adopting the other mechanisms, and consequently, reducing the costs of possible failures (time, administrative, and moral costs)

Figure 5: Table (5) Cost-Benefit Matrix of National Consultations as a T.J. Mechanism

[113] Triponel & Pearson, *supra* note 108.

[114] *E.g.*, The Tunisian Authority to Truth and Dignity (TDA), *الملخص التنفيذي للتقرير الختامي الشامل الهيئة التونسية للحقيقة والكرامة* [*The Executive Summary of the Final Total Report of the Tunisian Authority to Truth and Dignity*] 38, IVD (2019), http://www.ivd.tn/rapport/index.php.

2. GENERAL REMARKS AND POLICY IMPLICATIONS

In this section, I give some notes on the previous analysis of T.J. and its mechanisms. Some of these notes might be useful for the policy-makers, and others suggest research questions for future studies, especially empirical studies. The objective of this research was to use economic thinking in order to explain T.J. mechanisms as presented by the U.N. Guidelines in the context of revolution over autocratic regimes. A C.B.A. of each mechanism was provided to achieve this objective, drawn from the available literature and empirical findings, and following rational choice theories, especially public choice. Although this research represents a starting point for using C.B.A. in the field of T.J., there are general remarks that still have to be taken into account when doing so both by researchers, practitioners, and policymakers. These remarks include the following:

- Although the U.N. guidelines ignore amnesties as one of T.J. mechanisms and promote the value of accountability, the previous C.B.A. suggests a thorough consideration for this mechanism. Amnesties and lack of T.J. are two different things. When applying amnesties, the society and the government recognize that there were violations against the victims in the first place, that they are not acceptable, and that they will not be repeated, but they will be forgiven because seeking "justice" for them would be too costly to the extent that it would obstacle achieving T.J. goals themselves. This is different from ignoring what happened in the past altogether. I referred to this distinction earlier when discussing Tullock's thoughts on the safe exit to autocrats.

- Being, after all, a political arrangement, a C.B.A. of T.J. policies could change from one stage to another because of considerations of dynamic efficiency.[115] For example, in the beginning, starting prosecution initiatives to please the public and obtain their trust and thereby seize power may lead to over-deterrence. People holding public offices, whether top or less senior officials, could fear the legal responsibility over continuing past policies and the "fever" of chasing mistakes in the public sector. This fear can be fed by little knowledge of the applicable laws and the scope of legal liability, which finally leads to a level of activity below the optimal level for social welfare and levels of care far above the desired. After some time, however, failing to avoid or intentionally becoming stuck in the procedural and legal complications, which then lead to T.J. policies' failure or their

[115] For more on dynamic efficiency vs. static efficiency, *see* Veljanovski, *supra* note 5, at 35–6 Or as Teitel puts it: "Law in transitional periods is both backward-looking and forward-looking, retrospective and prospective, continuous and discontinuous". Teitel, *supra* note 6, at 215.

discontinuity, could lead to under-deterrence. Accordingly, the possibilities are always open, and the calculations regarding over/under deterrence are mercurial. Therefore, a careful trade-off between *ex-ante* and *ex-post* efficiency should be made when designing T.J. policies from the beginning.

- The measurement of the given variables in the presented models of T.J. mechanisms should not depend only on the universal quantitative weights given to them but also on the national and local weights to provide the most possible accurate evaluation. This local evaluation can be achieved through the pre-empirical studies referred to earlier. Even if some data will be challenging to collect, "*lawmakers would do better to use imperfect empirical analysis than perfect non-empirical analysis.*"[116] Some of the main reasons why empirical analysis is not still commonly used in the law-making processes, especially in less developed systems, are first, that most legal scholars and practitioners lack the necessary training for empirical analysis, and second, many law thinkers tend to trust the doctrinal analysis more than the empirical one. This skepticism is probably back to the conviction that "not everything can be quantified or measured." Although this might be true, still using the imperfect empirical findings to accompany, not replace, the pure legal classic doctrinal methods can be expected to give better guidance of legal design and application, especially where the momentum is critical and can not be reversed like in transitional policies. Finally, one can expect that these transitional systems would tend to lack the necessary expertise, human capital, and financial resources to run such analyses. International cooperation could help in this regard.

- One way to overcome the possible shortcomings of the pre-application empirical analysis is to consider T.J. pilot projects and phased approaches to minimize the costs of mechanisms that contribute negatively to the anticipated goals.[117]

- Although the U.N. guidelines are clear in affirming that the T.J. mechanisms should be victim centered,[118] the previous analysis shows how in many cases this focus is not realistic. Other variables in the process sometimes outweigh the "victims" considerations, like the consolidation of the new system or the practical limitations.

- The adherence to the international legal standards is a vital element in weighing the benefits of the adoption of T.J. mechanisms because: first, it endows legitimacy

[116] Robert D. Cooter, The Strategic Constitution 5 (2000).
[117] *See* Office of the United Nations High Commissioner for Human Rights, *supra* note 112.
[118] *See* United Nations, *supra* note 2.

over the new regime by giving it credibility from both the international and national actors; and second, this credibility may secure the international financial and political support which is highly needed in the transitional phase.

- The unconstitutionality considerations' concern is present in all of T.J. mechanisms because of the retroactivity of laws ban and inequality arguments. However, taking into account the international rulings on this matter, some concerns are less worrying in the contemporary and anticipated legal and political scholarship than in the past.

However, besides the already mentioned legal reasonings, one of the legal finesses that can be used to resolve the constitutional objections, which – up to my knowledge - was not mentioned in the available legal literature, is adopting T.J. mechanisms in the form of basic, i.e., organic, law. The basic laws – also known as fundamental or organic laws - are those laws that stipulate the regulation of constitutional matters, such as the rules governing the state authorities, elections, formation of the judiciary, . . . etc. However, they are not included in a constitutional document. In another phrasing, they are constitutional by nature but not by formality. Usually, such laws are briefly referred to in the constitution, and given a specific procedural framework to be issued or amended by the legislative. Accordingly, these laws have supremacy over other ordinary laws and regulations in the state, just like the constitution itself. They are considered complementary laws to that constitution. The only difference is that they are not a part of that constitutional document.[119] In France, which has a separate written constitutional document, these laws are called *Lois Organiques*.[120] In other jurisdictions, basic law is a term that is used alternatively to refer to the constitution, but with inferring the meaning that it is a temporary measure without formal enactment; however, it can last for a long time, like in the case of Germany. This codified or uncodified form of constitutions may be used for transitionary circumstances or for avoiding the claim of being the highest law for religious reasons.[121]

In the case of adopting basic laws in the first meaning, i.e. constitutional by nature, legislative by formality, the reasoning of issuing and differentiating them from both the constitutional text and the ordinary laws could have two arguments.

[119] *See* Sabry El Senousy, القانون الدستوري؛ شرح لأهم المبادئ الدستورية في المرحلة الانتقالية وأحكام دستور 2014 [Constitutional Law; An explanation for the most important general constitutional principles in the transitional phase and the provisions of the Constitution] 17, 17–23 (1 ed. 2014).

[120] *See generally* George Burdeau, Droit Constitutionnel et Institutions Politiques [Constitutional Law and Political Institutions] (20 ed. 1984).

[121] *See*David M. O'Brien, Constitutional Law and Politics: Struggles for Power and Governmental Accountability (9th ed. 2014).

First, minimizing the administrative and time costs of issuing and amending them in the usually relatively complicated constitutional process for their relatively flexible and/or urgent nature, while, second, maximizing the costs of their surpass for their vitality. In the case of adopting T.J. mechanisms in the form of laws, can these laws be considered as basic laws by nature? This can be debatable, and it is not the topic of this study to go through this legal debate. In some cases, the constitution-makers can avoid any suspicion or debates by explicitly stating that the laws of transitional justice in a specific phase will be considered or regulated by a basic law. For example, one of these cases is the Basic Law of Transitional Justice of Tunisia.[122] However, in brief, the legal opinion that this research adopts is: indeed , transitional justice laws are by their nature basic laws. The reason is that although they are genuinely relevant to criminal law because they regulate the penalizing and sanctioning of specific acts, they still regulate constitutional matters. These matters are such as: the formation of the state authorities during and after the transition, the limitation of the civil rights of some of its citizens, partial organization of the electoral rules, and most importantly, extra-ordinary judicial procedures that could not be valid under the "usual" constitutions. This last detail is utterly vital, specifically for this study, as it relates to one of the dilemmas that T.J. usually faces, which are constitutional challenges. These dilemmas were presented and analyzed through the discussion of the different mechanisms of T.J.. This formality could then work as a shield against any judicial challenge of the adopted T.J. mechanisms because the constitutions are not challenged before the supreme courts; they are the state's highest laws. The effectiveness of this solution ranges though, depending on the relevant mechanism. For example, in the case of reparations and truth commissions, because they lack the punishment aspect, unconstitutionality concerns are not as strong as in the case of prosecutions and lustration. For example, being a member of the past regime's party, or following orders that may lack moral reasoning even if they were legal, could not be sanctioned without retroactive laws or laws that could be struck down in constitutional courts based on inequality considerations between citizens.[123] Interesting enough, amnesty can be subject to the same concerns of unconstitutionality on the basis of the absence of due process, which happened in the Nepali or South African cases, for example.[124]

[122] *See* الرائد الرسمي [The Tunisian Gazette], *Organic law n° 2013-53 dated 24 December 2013, establishing and organizing the transitional justice* (2013), http://www.legislation.tn/detailtexte/Loi-num-2013-53-du-24-12-2013-jort-2013-105__2013105000531.

[123] A contemporary example of this direction is the position of the Egyptian Supreme Constitutional Court regarding the Political Lustration Law after the 2011 revolution. For more on this position and its circumstances, see Khalil, *supra* note 31, at 212.

[124] Amanda Cats-Batil, *Moving Beyond Transitions to Transformation: Interactions between Transitional Justice and Constitution-Building* 10 I 19–20 (2019), https://www.idea.int/publications/catalogue/moving-beyond-transitions-to-transformation.

Moreover, generating the T.J. policies in the form of a basic law guarantees that more of relevant actors are represented in the process of their designation because their delivery usually requires more sophisticated procedures and more representative authorities than those in the case of ordinary laws or regulations. This strategy reduces information asymmetry, which could lead to policies that are more likely to be self-enforcing. Finally, in this way, these laws can work as a pre-commitment device that includes all the interested parties, which could make feasible the movement forward through the political transformation.

Generally, a T.J. policy that takes into account: Proportionality + Combination of different mechanisms + Customization of the mechanisms upon the relevant case + Basic laws of T.J. should have the most effective outcomes achieving the goals of T.J. with the least legal complications. A balance between two considerations is needed to be done, not only by lawmakers but also by the constitutional courts. These considerations are: 1) The necessity of a radical break from the authoritarian past and the consideration of the context in which the principles of the rule of law, democracy, and due process are applied, which can justify a less strict application of their measures, and 2) The rejection for giving up the human rights considerations at all in the first step of the democratization, which could have a "killer effect" on the process.[125]

- The second step of the cost-benefit modeling should be comparing the mechanisms solely or when combined with other mechanisms. This step is, however, beyond the limits of this research and is left to future research. These models are only the basic models, and they are to be used and customized according to the different cases.

- Different combinations can be chosen of these mechanisms, they can reinforce each other, and they are not mutually exclusive, but they have to conform with the international standards and obligations.[126] These combinations depend on the outcomes of the cost-benefit analyses done by each case according to their inputs.

- Achieving deterrence regarding violations that involve political crimes is tricky. The point is that to design efficient ex-ante laws which can prevent the repetition of the past-autocratic policies, the calculations of the sufficient punishment should include the calculations of the regime itself of how to exploit the state

[125] *See* Marek Safjan, *Transitional Justice: The Polish Example, The Case of Lustration*, Eur. J. Legal Stud., Feb. 2008, at 235.*E.g.*, Horne, *supra* note 42; MMark A. Drumbl, *Prosecution of Genocide v. the Fair Trial Principle: Comments on Brown and Others v. The Government of Rwanda and the UK Secretary of State for the Home Department*, 8 J. Int'l Crim. Just. 289 (2010).Kieran McEvoy, *Beyond Legalism: Towards a Thicker Understanding of Transitional Justice*, 34 J. L. & Soc'y 411 (2007). For more on the benefits of constitutionalizing T.J. policies, see, e.g.,also Cats-Batil,, *supra* note 124.

[126] United Nations, *supra* note 2, at 3,10.

resources without getting the people to the point of revolution. Moreover, the offenders should be comparing the gain from the violation with the cost if they are apprehended and punished, i.e., the probability of being sentenced is a part of the deterrence calculation.[127] This probability in T.J.'s case is very low for a couple of reasons: 1) As indicated earlier, there will be a lack of evidence and other procedural difficulties. Consequently, a rational authoritarian regime would take the necessary arrangements not to be sentenced, either by destroying evidence or through bargaining with the new regime to reach a compromise; 2) In case that the regime heads' calculations reach the point that they think they will not be able to get away with what they committed, they will not be deterred, but will exploit the people until this point, and afterward they will usually make some improvements if possible, or flee the country.[128]

- Despite the last point, T.J. policies' negative effect on the probability of repetition of past crimes, which is the same goal of deterrence, could be explained through other rationales. These rationales include: 1) The institutional reforms in terms of changing the laws and enforcing them; 2) Institutional reforms in terms of lustration and vetting policies, which should eliminate a number of the leading past criminals; 3) The national consultations can contribute to this goal by spreading awareness among the people about human rights, the rule of law, and justice, and strengthening the civil society, which is expected to increase the costs of trying to repeat the past violations.

- The combination between prosecution initiatives and/or lustration for only the head members of the past regime and reparations and/or truth commissions could be an effective strategy to avoid a counter-revolution, society polarization, over-deterrence, stigmatization of possibly innocents, and loss of human capital and qualities. At the same time, it achieves the accountability and prevention of the top criminals from continuing their criminal activities. In other words, giving the heads of the past-regime amnesties could be dangerous, unlike the inferior personnel who worked for them. This policy can be justified through the same economic rationality used by Richard A. Posner in explaining the economic reason behind the Multiple Offender Laws.[129] The first part of the justification is that the

[127] Posner, *supra* note 29, at 18.
[128] For more on the dictator choices regarding succession and retirement, *see also* Tullock, *supra* note 105; *see, e.g.*, Gordon Tullock, *The Goals and Organizational Forms of Autocracies; the Problem of Succession, in* The Selected Works of Gordon Tullock (Volume 8): The Social Dilemma of Autocracy, Revolution, Coup d'Etat, and War 82 (Charles K. Rowley ed., 2005).
[129] Posner, *supra* note 29.

leading criminals who put the policies, commanded, and directed the ordinated officials to apply them, showed a higher propensity to commit similar crimes in the future. The second is that the sanction value should be raised on people who, through their past, had shown that the return of the crime is higher for them than for other possible criminals. Accordingly, this variation in justice mechanisms should achieve a higher probability of preventing crimes and deterrence. Further systematic empirical research is needed to measure the effects of the different combinations of T.J. mechanisms

- This strategy, however, has a low probability of success in the case of mechanisms against the army and the police as a part of what is called in the T.J. literature as "Security Sector Reform [hereinafter S.S.R.]. While most of the pro-past-regime persons can find a place for themselves in the new regime and obey the new system of democracy and the rule of law, the police, the persons who used to protect the past regime and get the highest rewards in the society to do so, may find more difficulty in getting integrated into the new system. The reason is that any change in the regime incurs a loss for them; they will be deprived of the extra-payoffs they used to gain under the old regime. At the same time, if all or most of them were vetted or prisoned for a while, this will be like forming an army against the state, a trained army that lacks any incentive to work for the new regime, and many incentives to destroy it. The S.S.R. as a part of T.J. is already under-studied. The available studies on S.S.R. focus on establishing the connection between it and T.J. at all, reasoning S.S.R., its limits and challenges, and exploring some of the case studies that witnessed its application.[130] However, up to my knowledge, none of these studies analyzed this specific concern of S.S.R.. This concern forms a dilemma. A solution, however, might be replacing the leading staff of the police and the army with influential leaders who are pro the new regime. It could be successful if the big heads are pro the new regime and can strictly monitor these troops. The empirical analysis will be needed to test such a suggestion and

[130] *E.g.*, Eirin Mobekk Geneva , *Geneva Centre for the Democratic Control of Armed Forces (DCAF) Transitional Justice and Security Sector Reform: Enabling Sustainable Peace*, DCAF (2006), https://www.dcaf.ch/sites/default/files/publications/documents/OP13_Mobekk.pdf; Ana Cutter Patel, *Transitional Justice, DDR and Security Sector Reform*, Research Brief (2010), https://www.ictj.org/sites/default/files/ICTJ-DDR-S.S.R.-ResearchBrief-2010-English.pdf; Christopher Gitari Ndungú, *Failure to Reform; A Critique of Police Vetting in Kenya*, ictj (2017), https://www.ictj.org/sites/default/files/ICTJ-Briefing-Kenya-PoliceVetting-2017.pdf; Laura Davis, *Transitional Justice and Security System Reform*, Initiative for Peace Building (2009), http://www.peacewomen.org/sites/default/files/recon_transjusticessr_ictj_2009_0.pdf; Sumit Bisraya & Sujit Choudhry, *Security Sector Reform in Constitutional Transitions,* ICTJ (2020), https://www.idea.int/publications/catalogue/security-sector-reform-constitutional-transitions.

investigate whether this concern was existent in any of the cases that applied S.S.R. and the impact of the different strategies – if any – to deviate.

- Each of the C.B.A. presented in this article assumed that the relevant mechanism would be applied for a specific period; i.e., it is not open-ended. However, that analysis would change if this time cap is lifted. For example, one could expect that the administrative and time costs would be multiplied in this case. This argument is not a suggestion of time-limiting T.J.'s application, and despite that, any typical reasoning of the legal principle of *lapse by prescription* could be presented here; it can be faced by a counter-analysis of the exceptional nature of the subject of the violation of T.J.. The point is: To decide what the time cap of a relevant T.J. mechanism should be if there should be any, the following will be needed:

(a) an empirical assessment of the legal complications and multi-aspect practical consequences of applying the open-ended T.J. mechanisms compared to the time-limited ones.[131]

updated C.B.A. analysis that adds the aspect of the time limitations to the assessment of the costs and benefits of the relevant mechanism to be compared to the time-capped original C.B.A. .

(b) considering the possibility of innovative ways to apply the classical forms of T.J. mechanisms that could minimize the costs and maximize the benefits of lifting the time limits when applied to them. Also, considering that not all the mechanisms – and sub-mechanisms - will respond equally to the variable of "time-cap." Consequently, a comparison between the updated C.B.A. should be kept in mind.

(c) despite the last point, attention should be given to the fact that lifting the time limits has an effect that will apply to all of the cases and all of the mechanisms. This effect will open the negotiation over the truths, information, and amnesties with the past-regime member endlessly. The possible costs of negotiation, both legal and illegal negotiation, will always be present.

- Although I am not conducting a quantitative analysis using the presented models of the cost-benefit analysis of T.J. mechanisms here, I think that such an analysis would be interesting to conduct or read in the future.

[131] See for example some of the notes on the German case of the open-ended application of some of the T.J. mechanisms provided by Thomas Weber, *Time Appears to Have Run Out on the Last Nazi War Crimes Trials. But There Are Other Roads to Justice*, Time (Apr. 3, 2019), https://time.com/5563615/nazi-trials-over/ .

CONCLUSION

Systems need a radical break between the past iteration and its new form , a process which aims to internalize the past harm and promote respect of democracy and human rights to achieve their transformation from autocracy to democracy. One of the tools to achieve this transformation is Transitional Justice, which is done by adopting different mechanisms.

Each of the mechanisms provided by international law has its challenges, costs, and benefits. The decision to adopt one or more of these mechanisms should be informed by a careful analysis of each mechanism's costs and benefits and each possible combination. This article outlined a theoretical cost-benefit analysis derived from the literature on the Transitional Justice of case studies, reports, empirical studies, laws, rulings, and theoretical analyses. However, the weights given to these inputs will vary from one case to another, and that is how the output will vary as well. Consequently, an effective mechanism for one case may be ineffective to another, depending on many other possible variables. This process is costly and complex due to the nature of factors considered and the different – and even opposing - stakeholders involved. However, deliberative and multilateral solutions, including democracy, are always more costly in the short-term than authoritative and mono-designed and controlled processes. Arbitrary plans are also expected to be less costly in the short term than informed study-based plans and decisions. However, the deliberative, multilateral, and study-based decisions are expected to be less costly in the long run, especially in a critical and broad project like T.J. after a revolution.

On another note, by their nature, from a constitutional legal perspective, transitional justice laws are basic organic laws. In case their formality is aligned with this nature, this may solve some of the legal complications T.J. faces. Future empirical research can tell us more about the effectiveness of this solution and compare the efficiency of the different T.J. mechanisms and their combinations.

University of Bologna Law Review

Department of Legal Studies
Via Zamboni, 42
40126 Bologna (Italy)

Email: bolognalawreview@unibo.it
Website: bolognalawreview.unibo.it
Facebook: www.facebook.com/ubolognalr
Twitter: twitter.com/unibolawreview