JOURNAL OF LAW, ECONOMICS & POLICY

Cite as 19 J.L. ECON. & POL'Y 1

Journal of Law, Economics & Policy (ISSN: 1553-4367)
George Mason University School of Law
3301 Fairfax Drive
Arlington, Virginia 22201
jlep@gmu.edu
www.jlep.net

The *Journal of Law, Economics & Policy* is published quarterly at George Mason University School of Law, though additional issues may be added during a publication year. One issue is devoted to papers presented at our annual symposium, and the other issues feature peer-reviewed scholarly articles. The Journal focuses on the legal side of law and economics, an approach that de-emphasizes technical mathematics, while providing readers with greater access to economic insights about legal and policy issues.

Submissions:

The Journal invites authors to submit manuscripts for inclusion in our peer-reviewed issues. Authors should only submit articles that have not been previously published and are not under consideration for publication elsewhere.

Consistent with the Journal's stylistic goals, submitted articles should be clear and concise. Preference is given to submissions that are thirty pages or less, double-spaced, including footnotes. Citations must conform to the latest edition of *The Bluebook: A Uniform System of Citation*.

Authors may submit manuscripts in electronic or hard copy format. Electronic submissions should be sent to *jlepsub@gmail.com* in Microsoft Word or other compatible formats. For paper submissions, please mail in a hard copy, along with an electronic copy on a CD-ROM.

Subscriptions:

Subscriptions are available at the following rates:

Students: $25 per volume (Domestic addresses only – U.S. & Canada)
Domestic (U.S. & Canada): $40 per volume, $20 per issue
International: $60 per volume, $25 per issue

For subscription information, please contact the Journal by mail or e-mail.

All articles copyright © 2024 by the *Journal of Law, Economics & Policy*, except where otherwise expressly indicated. For any article to which the Journal holds the copyright, permission is granted to reprint or copy any portion of such article for classroom or other educational use, provided that: (1) proper notice of such copyright is affixed to each copy; (2) the author and the *Journal of Law, Economics & Policy* are clearly identified; (3) in the case of copies distributed in class, students are charged no more than the cost of duplication; and (4) the Journal is notified in writing of the use.

Copyright © 2024 by
The Journal of Law, Economics & Policy

EDITORIAL BOARD
2023-2024

SAM BELLET
Editor-in-Chief

JACK RING
Productions Editor

TALIA SEMAAN
Executive Editor

ASHLEY EVARO
Submissions Editor

HAILEY SYLVANDER (SR. ARE)
MIA WRIGHT REAL
QUYEN DANG
DANIEL PEREZ
Articles & Research Editors

ANTHONY SMITH
Publications Editor

CORY JACK
Symposium Editor

MATT PECORARO
MANASA CHILAKALA
NATHANIEL YU
HAYDEN DWYER
Notes Editors

MEMBERS

GABRIELLE BERANEK
KENDAHL BRIER
SOPHIA CABANA
ANTHONY CIRRI
LILLIAN CLARK
CALA COFFMAN
JACOB DERUE
JAKE DILLON
IAN DOW
JACOB FRANKSON
JAEHEE JUNG
KATHERINE KOON
ALEXANDER KRUPNICK
ALEXANDER LIUBINSKAS
ALEXANDRA MANNIX
ETHAN MCCARTHY
GARNER NOTTINGHAM
GREGORY PELLETIER
GABRIEL ROBERTSON
TRISTAN SILVA
JACK TIERNAN
AARON WATT
JIMMIE WEBB
ANDREW WILSON
CONNOR WISEMAN
ETHAN YANG

BOARD OF ADVISORS

HENRY BUTLER
HON. GUIDO CALABRESI
ROBERT D. COOTER
ROBERT C. ELLICKSON
RICHARD A. EPSTEIN
HON. DOUGLAS H. GINSBURG
MARK F. GRADY
BRUCE H. KOBAYASHI
A. DOUGLAS MELAMED
FRANCESCO PARISI
HON. ERIC POSNER
RICHARD A. POSNER
ROBERTA ROMANO
HANS-BERND SCHÄFER
STEVEN M. SHAVELL
HENRY E. SMITH
VERNON L. SMITH
THOMAS S. ULEN
W. KIP VISCUSI
TODD J. ZYWICKI

CONTENTS

ARTICLES

1 DISCOVERING OR SETTING AGGREGATE ROYALTIES AND FRAND RATES FOR SEP PORTFOLIOS
Keith Mallinson

41 NEW MONEY, OLD STATUTES: INFLATION AND STATUTORY DRIFT
Patrick Sullivan

95 NOTE: USING EMPIRICAL ANALYSIS TO ASSESS THE ACCURACY OF EXPERT TESTIMONY PREDICTIONS FROM THE AT&T-TIME WARNER VERTICAL MERGER
Lillian Clark

116 AGENTS AS UNDERWRITERS: INCORPORATING AGENCY LAW TO MAKE SECTION 11 OF THE SECURITIES ACT BETTER APPLICABLE TO NON-TRADITIONAL IPOS
Ethan McCarthy

139 A POLITICAL QUESTION SOLUTION TO THE ALARMING RISE OF NATIONWIDE INJUNCTIONS
Daniel Perez

DISCOVERING OR SETTING AGGREGATE ROYALTIES AND FRAND RATES FOR SEP PORTFOLIOS

Keith Mallinson[*]

I. INTRODUCTION

Fair Reasonable and Non-Discriminatory (FRAND) licensing[1] for Standard-Essential Patents (SEPs)[2] has worked well in communications technologies that have provided soaring performance and interoperability among networks, applications, and devices over the last 30 years. Market-based royalty rates have been established largely through bilateral negotiations in cellular communications with 2G, 3G, and 4G technology standards, and with patent pooling of video and audio compression technology standards such as AVC/H.264 and AAC.[3] Total royalties paid for all SEP licensing are no more than around five percent of mobile phone product revenues, and rather less

---

[*] Keith Mallinson is the founder of WiseHarbor. He has more than 25 years of experience in the telecommunications industry: as a research analyst, consultant and testifying expert witness. Complementing his industry focus, he has a broad skill set including technologies, market analysis, regulation, economics, and finance. He has published numerous articles and speaks publicly at industry events on a wide variety of topics including market developments and competition in next generation mobile network technologies, semiconductors, technology standards, patents, and licensing.

WiseHarbor is a global analyst and consulting firm serving companies, industry associations and government clients. Founded in 2006, WiseHarbor has remained focused on the ever-expanding and changing ecosystem in wireless and mobile communications as it connects people and an increasing array of things — from wearables and connected buildings to autonomous vehicles and industrial robots — and in transformation of many different markets and industry verticals. For more information, visit https://www.wiseharbor.com.

[1] For example, according to standard setting organization (SSO) ETSI, as applicable to various cellular standards including 4G and 5G, "IPR holders whether members of ETSI and their AFFILIATES or third parties, should be adequately and fairly rewarded for the use of their IPRs in the implementation of STANDARDS and TECHNICAL SPECIFICATIONS." ETSI Intellectual Property Rights Policy, Annex 6, Section 3.2 (Nov. 2022).

[2] These are patents that read on standards. In other words, such standards cannot be implemented without infringing those numerous patents owned by many different companies. An SSO will not include an SEP technology in a standard without a corresponding FRAND commitment.

[3] Market-based royalty rates are those negotiated with due regard for i) how valid patented technologies confer value to applications, *see* Georgia-Pacific Corp. v. U.S. Plywood Corp., 318 F. Supp. 1116, 1120 (S.D.N.Y. 1970), ii) established royalty-rate benchmarks including the extent to which these are underpinned through volumes and values of licensed trade over the years, Karl Fink, *Where Is the Federal Circuit on Using Comparable Licenses to Prove Reasonable Royalties and Apportionment in Patent Cases?*, JD SUPRA (Feb. 2022), and iii) the extent to which parties are "similarly situated" in the market, *see* Dennis Carlton and Allan L. Shampine, *An Economic Interpretation Of Frand*, 9(3) J. COMPET. L. ECON. 531-52 (2013).

than that on other devices and applications.[4] Aggregate amounts paid have remained rather flat over the last decade despite the introduction of new technologies and standards such as 5G, since 2019.[5] The technology transferred through licensing from SEP owners such as Ericsson, Interdigital, Nokia, and Qualcomm to implementers including smartphone Original Equipment Manufacturers (OEMs) Apple, Samsung, Sony, and Xiaomi has brought widespread commercial success and consumer satisfaction across a large and ever-expanding ecosystem.[6] Mobile voice and texting revolutionized personal communications with massive adoption globally in the 1990s and 2000s. Smartphones, which provide fast data connections to networked applications such as Instagram in social media, Netflix in video streaming, and Google's Maps in navigation, have prevailed worldwide since the mid-2010s. By yearend 2022, 5.4 billion people subscribed to a mobile service, including 4.4 billion who also used the mobile Internet.[7] Mobile connectivity is extending revolutionary change beyond personal communications to the Internet of Things (IoT) with a total now of more than 16 billion cellular devices.[8]

Notwithstanding the evident efficacy and efficiency of standards development and SEP licensing that has enabled the improvements in technical performance, commercialization, and consumer adoption, some European and US interests are lobbying for rate-setting of aggregate and individual licensors' royalties.[9] In absence of economic logic or supporting evidence that royalties are harmful, unfair, or excessive, major implementers in Big Tech and automotive industries are self-servingly seeking to reduce their royalty

---

4 Keith Mallinson, *Cumulative Mobile-SEP Royalty Payments No More Than Around 5% of Mobile Handset Revenues*, IP FIN. (Aug. 19, 2015), http://www.ip.finance/2015/08/cumulative-mobile-sep-royalty-payments.html [hereinafter Mallinson, *Cumulative*]; Alexander Galetovic, Stephen Haber & Lew Zaretzki, *An Estimate of the Average Cumulative Royalty Yield in the World Mobile Phone Industry: Theory, Measurement and Results*, 42 TELECOMM. POL'Y 263 (Apr. 2018); J. Gregory Sidak, *What Aggregate Royalty Do Manufacturers of Mobile Phones Pay to License Standard-Essential Patents?*, 1 CRITERION J. INNOVATION 701 (2016).

5 Keith Mallinson, *The Smartphone Royalty Stack: A long-term look*, IAM (Mar. 2, 2022), https://www.wiseharbor.com/wp-content/uploads/2022/04/Special-Report-2022-Q1_-Patent-Dealmaking-IAM-Smartphone-royalty-stack.pdf.

6 Keith Mallinson, *Don't Fix What Isn't Broken: The Extraordinary Record of Innovation and Success in the Cellular Industry Under Existing Licensing Practices*, 23 GEORGE MASON L. REV. 967 (July 2016) [hereinafter Mallinson, *Don't Fix What Isn't Broken*].

7 GSMA Intelligence, The Mobile Economy 2023, 3 (2023).

8 The Radicati Group, *Forecast Number of Mobile Devices Worldwide from 2020 to 2025*, STATISTA (Apr. 2021), https://www.statista.com/statistics/245501/multiple-mobile-device-ownership-worldwide.

9 Gordon G. Change, *Why is Europe Helping China Decimate U.S. Tech Leadership*? NEWSWEEK (Sept. 6, 2023), https://www.newsweek.com/why-europe-helping-china-decimate-us-tech-leadership-opinion-1825029.

costs.[10] Misguided legislative proposals are based on poorly supported and dubious assertions that there is insufficient transparency in royalty rates and that rates offered by some licensors are supra-FRAND. While increasing disclosures on existing licensing would improve transparency, to instead set rates anew will harmfully upset what has been proven to work well with no sign of market failure. Proposed legislative changes are attempting to abandon or diminish well-established market-based mechanisms in determining royalty charges. It seems that processes of commercial negotiation in establishing rates and applying comparable licensing benchmarks derived from existing licenses could be replaced by the "top-down approach" in which a notional aggregate royalty is apportioned among SEP owners based on their respective applicable patent counts.[11] While there are many legal, economic, and commercial reasons why the proposed regulation should not be pursued, if aggregate rate setting and apportionment of royalties is to be employed it is essential that governance, organizational processes, and analytical methodologies are fit for purpose. While this article touches on many different important issues, it focuses principally on the economics and commercial factors in the methodologies and metrics to be used in deriving figures for aggregate royalties and individual FRAND rates using the top-down approach. My objective here is to highlight issues including shortcomings and to prescribe how — if aggregate rate setting and top-down apportionment are to be used at all — reasonably accurate, reliable, fair, and consistent rates can be set. These are necessary to ensure ongoing successful development, implementation, and consumer adoption of standard-essential technologies in the anticipated widening array of applications.

This article is largely based on the two submissions I made to the European Commission in response to its request for feedback on its draft legislation and impact assessment report, published April 27, 2023.[12] I have also

---

10 Brooke Masters, *What the Great EU Patent Fight Means for Global Competition*, FIN. TIMES (Aug. 2023), https://www.ft.com/content/ebd533a7-b8d1-4d51-bd2e-8288c60490d1.

11 An early appearance and judicial implementation of such a technique was in Judge Holderman's 2013 Opinion in *In re Innovatio IP Ventures, LLC*, 956 F. Supp. 2d 925 (N.D. Ill. 2013). Significantly different implementations of a top-down approach have also been applied in other judgements, including in *Unwired Planet Int'l Ltd. v. Huawei Tech. Co. Ltd.*, [2020] UKSC 37 and in *TCL Comm'n. Tech. Holdings, Ltd. v. Telefonaktiebolaget LM Ericsson*, 2017 U.S. Dist. LEXIS 214003 (C.D. Cal. Mar. 9, 2018) (unanimously and entirely reversed on appeal).

12 Feedback on draft EU legislation for SEPs by Keith Mallinson: Keith Mallinson, Comment Letter on Proposal for a Regulation of the European Parliament and of the Council on Standard Essential Patents and Amending Regulation (June 14, 2023), https://www.wiseharbor.com/wp-content/uploads/2023/06/Mallinson-SEP-consultation-response-14-June-2023.pdf [hereinafter Mallinson June]; Mallinson, Comment Letter on Proposal for a Regulation of the European Parliament and of the Council on Standard Essential Patents and Amending Regulation (Aug. 8, 2023), https://www.wiseharbor.com/wp-content/uploads/2023/08/Aggregate-rate-setting-Mallinson-WiseHarbor-2023.08.08.pdf [hereinafter Mallinson August]; European Commission, *Regulation Of The European Parliament And of The Council,* (EU)2017/1001 (Apr. 2023), https://single-market-economy.ec.europa.eu/system/files/2023-04/COM_2023_232_1_EN_ACT_part1_v13.pdf [hereinafter European Commission]. My

repeated some astute and valuable insights provided by others in their feedback to the Commission.

II. EU AND US PROPOSALS SUBSTITUTE RATE SETTING FOR NEGOTIATED RATES BASED ON COMPS.

A. *Established FRAND Licensing Practices*

If a prospective SEP licensor can demonstrate it owns infringed and valid patents it is entitled to a FRAND license. Where charges and other terms have been established in existing licenses, some of these might be comparable benchmarks for determining licensing charges in other agreements. In litigation, comparable licenses are generally considered to provide the very best benchmarks in determining royalty charges.[13] Some of these benchmarks become publicly available (e.g., in court decisions) so they are also used in separate licensing negotiations. The applicability and comparability of existing licenses depends upon the extent to which these are substantiated by licensed trade, the timing of that, and how similarly situated prospective licensees are (e.g., anticipated volumes and values in licensed trade). Absent these benchmarks or in addition to them, parties in patent licensing negotiations consider many other factors. These include the value standard-essential technologies bring to devices, size and quality of patent portfolios including product infringement and validity considerations. Focus is typically on "proud lists" of up to fifteen selected patents; patent litigation history is also pertinent.[14] FRAND licenses are typically negotiated in global agreements. When the courts are asked to adjudicate in disputes, courts may make FRAND rate determinations only for SEPs issued in their own countries, or for all SEPs worldwide that would likely be included if the license had been negotiated by the parties. This raises various inter-jurisdictional issues.

submissions include additional detailed analysis on some economic and commercial issues I have omitted here in the interests of brevity. For example, I explain the inapplicability of using patent pool royalty rates as aggregate royalty figures for apportionment in bilateral rate-setting.

[13] For example, in *Unwired Planet* decisions where the courts were able to review numerous confidential licensing agreements. [2020] UKSC 37.

[14] "[L]icense negotiations outside litigation tend to focus on a 'proud list' of patents, although licensees typically wish to extend the license to all potentially relevant patents in the licensor's portfolio and all of the licensee's potentially relevant products (or, at least, all those in a given category or field of use). Similarly, patent holders generally tend to not bring suit over every patent that they might assert against the defendant, but rather choose to sue over a relatively small group of patents (a 'proud list') that have the greatest likelihood of being seen as (i) valid, (ii) infringed by a significant portion of the prospective licensee's product and service offerings, and (iii) valuable (i.e., contribute significant additional profit to the sales of those products)." Michael P. Akemann, John Blair & David Teece, *Patent Enforcement in an Uncertain World: Widespread Infringement and the Paradox of Value for Patented Technologies* 7 (Tusher Ctr. for Mgmt. Intell. Cap., Working Paper No. 6, 2014), https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2845002.

While detailed analysis of these issues is beyond the scope of this article, concerns arising from these are significant factors in instigations such as that in the US to prohibit the recognition of FRAND rates set for US patents by foreign courts and to establish a US rate-setting tribunal for US patents. The proposed EU legislation would also result in rate setting. I consider the governance and organizational processes before focusing on the computational methodologies to be employed in such rate setting.

B. *New Regulation for SEPs in the EU*

1. Proposed legislation

The Commission has published its proposed legislation along with an impact assessment report.[15] The former states that "The overall objectives of [its] proposed initiative are to:

a) ensure that end users, including small businesses and EU consumers benefit from products based on the latest standardised technologies;
b) make the EU attractive for standards innovation; and
c) encourage both SEP holders and implementers to innovate in the EU, make and sell products in the EU and be competitive in non-EU markets."

Its initiative seeks to:

i. "make available detailed information on SEPs and existing FRAND terms and conditions to facilitate licensing negotiations;
ii. raise awareness of SEP licensing in the value chain and
iii. provide for an alternative dispute resolution mechanism for setting FRAND terms and conditions."[16]

The proposed regulation:

I. "requires the registration of all SEPs in force in EU Member States before the newly established Competence Centre at the EU

[15] *See* European Commission, *supra* note 12; European Commission, *Impact Assessment Report: Regulation of the European Parliament and of the Council,* (EU)2017/1001 (Apr. 2023), https://single-market-economy.ec.europa.eu/document/download/a009816a-3b24-46c8-9c3c-fd8bd89a1380_en?file-name=SWD_2023_124_1_EN_impact_assessment_part1_v4.pdf [hereinafter Impact Assessment Report].

[16] European Commission, *Explanatory Memorandum: Regulation of the European Parliament and of the Council,* (EU)2017/1001 (Apr. 2023), https://single-market-economy.ec.europa.eu/document/download/b7501cc3-febe-40ee-b4a0-6cd5a63a860c_en?file-name=COM_2023_232_1_EN_ACT_part1_v13.pdf.

Intellectual Property Office (EUIPO), as a pre-condition for litigation of SEPs in the EU;

II. provides for annual essentiality checks of registered SEPs;

III. introduces a system of notification of aggregate royalty rates for standards, and requires entering into mandatory FRAND determinations before initiating SEP litigation in the EU."[17]

The Commission indicates "uncertainty about the SEP royalty burden" and that "Stakeholders consider that the FRAND licensing concept could benefit greatly from some clarification, notably with regard to the determination of an aggregate royalty burden."[18] The proposed regulation also notes that "[i]n view of the global character of SEP licensing, references to aggregate royalty and FRAND determination may refer to global aggregate royalties and global FRAND determinations, or as otherwise agreed by the notifying stakeholders or the parties to the proceedings."[19] The proposed regulation and the above processes are evidently far from being fully defined, let alone planned out for execution. The Competence Centre needs to be set up from scratch. The EUIPO does not yet have any of the required expertise in SEPs, FRAND licensing, essentiality checking, aggregate rate setting, and individual royalty rate determination. It would be very enlightening if, instead of setting rates anew, a large and representative sample of implementers were to disclose how much they actually pay to individual licensors and in aggregate for various standards. Unhelpfully, such information is highly confidential. However, court decisions based on extensive review of executed licenses and associated licensed trade, including several in the UK, are already providing some indications of aggregate figures, notwithstanding redactions. If a trusted third party could confidentially collect such information more extensively it could calculate and publicly reveal various averages and ranges while preserving anonymity and not revealing individual royalty rates. Such ex-post figures could provide a most valuable indicator of aggregate royalties to be paid by others and such figures could be compared with the various ex-ante and other royalty rate figures disclosed (e.g., in licensors' rate cards). It is puzzling that the Commission has seemingly not sought to look into the horse's mouth in this way in its stated quest to increase transparency on aggregate royalty costs.

[17] Igor Nikolic, *Some Practical and Competition Concerns with the Proposed Regulation on Standard Essential Patents,* 4IP COUNCIL 1 (July 3, 2023), https://www.4ipcouncil.com/application/files/4616/8847/4214/2023.07.03_Proposed_Regulation_on_Standard_Essential_Patents__1.pdf.

[18] European Commission, *supra* note 12, at 8; Impact Assessment Report, *supra* note 15.

[19] European Commission, *supra* note 12, at 27.

2. Regulated Procedures

Despite the use of well-established licensing benchmarks and negotiating practices in determining royalty rates, there is significant dispute about how else, if at all, to value SEP portfolios and determine FRAND royalty charges for these. According to the impact assessment, "[a]lthough an impressive amount of scholarship has analyzed or interpreted the FRAND concept, this scholarship is characterized by persistent differences of opinion on key aspects of the FRAND concept such as royalty evaluation methods and obligations to license certain parts of the relevant industry."[20] My article critically analyses alternative valuation methods for aggregate and individual SEP owners' royalties. The Commission clearly expects aggregate royalties to be determined for some technology standards and that these in turn will be apportioned among SEP owners. As the impact assessment indicates from the results of its literature analysis: "An aggregate royalty for a standard is the royalty due for all SEPs on the standard. It is the starting point in a top-down determination of the royalty to be paid for a given portfolio."[21]

The Commission's proposals imply that SEP holders — including net licensors and net licensees — would voluntarily participate in negotiating aggregate royalties and proposing these to the EUIPO:

- "Holders of SEPs in force in one or more Member States for which FRAND commitments have been made may jointly notify the competence centre the aggregate royalty for the SEPs covering a standard." The notification shall contain information on "the estimated percentage of SEPs they own collectively from all SEPs for the standard."[22]
- "Holders of SEPs in force in one or more Member States representing at least 20 % of all SEPs of a standard may request the competence centre to appoint a conciliator from the roster of conciliators to mediate the discussions for a joint submission of an aggregate royalty."[23]
- "A SEP holder or an implementer may request the competence centre for a non-binding expert opinion on a global aggregate royalty" to be made within 150 days of publication of the relevant standard or new implementations being sold in the EU.[24]

---

20 Impact Assessment Report, *supra* note 15, at 86.
21 *Id.* at 118-119.
22 European Commission, *supra* note 12, at 35, Art. 15.
23 *Id.* at 36, Art. 17.1.
24 *Id.* at 37-38, Art. 18.

The Commission's willingness and intent to set global aggregate royalty rates is in conflict with its focus on SEPs in force in EU Member States.[25]

The aggregate rate notification deadline of 150 days from publication of the standard is unrealistic because this is insufficient time to know how a standard will be implemented. In a recent paper for 4iP Council, Dr. Igor Nikolic, Research Fellow at the European University Institute, notes that patent pool experience shows "it may take years for patent owners to agree on mutually acceptable and market-realistic rates."[26]

Notwithstanding collective public announcements in the 2000s on aggregate rate objectives for 3G and 4G, and with statements these were not caps, as discussed in Section III(E)(2)(d), major licensors have made no such announcements since then. Nevertheless, to the dismay of some SEP owners, some courts have regarded such statements as binding commitments to cap royalties while disagreeing with the major licensors making them about how those statements should be interpreted.[27]

The Commission also proposes essentiality checking by EUIPO assessors. SEP owners have shunned such a voluntary system in Japan.[28] There is no evidence that these European proposals will be any more welcome.

The Commission seeks that the essentiality of all patents or a representative random sample of them reading on standard are checked — not only small numbers of them (i.e., 50 or fewer) per patent owner's portfolio.[29] This stealthily implies that it wants patent counts to be used as measures of patent strength — as required in top-down approach FRAND rate setting — even though this widely contested apportionment method is not explicitly identified or advocated in the proposed legislation. I analyze such methods and their shortcomings in Sections III(C) and III(D).

The proposed EU legislation makes only one mention of comparable licenses — in passing when describing difficulties including transparency and complexity in making FRAND determinations.[30] The impact assessment includes references to comparable licenses to acknowledge that they are used

---

[25] "In view of the global character of SEP licensing, references to aggregate royalty and FRAND determination may refer to global aggregate royalties and global FRAND determinations, or as otherwise agreed by the notifying stakeholders or the parties to the proceedings." European Commission, *supra* note 12.

[26] Nikolic, *supra* note 17, at 1.

[27] For example, the *TCL* decision states, "the Court is unconvinced by [the Ericsson witness'] attempt disavow Ericsson's commitment to calculate royalties based on a proportional share of a total aggregate royalty capped at a modest single digit" and the decision mistakenly regards announced aggregate royalty goals as being multimode rather than single-mode rates. *See TCL*, 2017 U.S. Dist. LEXIS 214003 at *21; Mallinson August, *supra* note 12, at 14-15. While that decision was unanimously and entirely vacated on appeal, much of it is still relied on in expert FRAND licensing analysis including in litigation.

[28] European Commission Joint Research Centre, Rudi Bekkers et al., *Pilot Study for Essentiality Assessment of Standard Essential Patents*, EUR 30111 EN 51-54 (2020), https://publications.jrc.ec.europa.eu/repository/handle/JRC119894.

[29] The EU's proposed legislation includes the word "sample," "sampling," and "sampled" a total of 21 times. European Commission, *supra* note 12.

[30] European Commission, *supra* note 12, at 34-38.

and to indicate that some are dissatisfied with the extent of disclosure of existing licensing terms and licenses. It goes no further than stating under the Section IV heading "Qualitative royalty apportionment criteria" that "Other criteria that could be considered include comparable licenses, technical importance of the claimed subject matter to the product, technical contributions to the standard, technical contributions to key features of the standard."[31] Neither document finds that the established royalty charges in existing licenses are excessive or inapplicable FRAND licensing benchmarks. There is no discussion of how disclosures might be increased.

3. Antitrust Concerns

Dr. Igor Nikolic also indicates, in his recent paper, concern about possible buyers' cartel effects (i.e., monopsony rate-setting).[32] He states, "it is unclear from the text of the Draft SEP Regulation if implementers are allowed to coordinate their submissions to conciliators."[33] He is concerned that "implementers might use the process to exchange commercially sensitive information and agree on the maximum global aggregate royalties they would pay."[34]

He is also uneasy that the draft regulation does not include the "competition safeguards against the exchange of commercially sensitive information in the process of joint notification of aggregate royalty rates."[35] Patent pools "are expressly required by the Technology Transfer Guidelines to prevent the exchange of sensitive commercial information among their members."[36]

From an economic standpoint, price coordination (i.e., of royalty rates) among some SEP owners ought not to be problematic; but only so long as other licensors are not bound by such pricing. SEPs are necessarily complements — patented technologies are not in competition with each other once they have been selected for use in a standard and have become SEPs — and so the implementation and licensing of all of them is required.[37] Competition

---

31 Impact Assessment Report, *supra* note 15, at 215.

32 Nikolic, *supra* note 17.

33 *Id.* at 4.

34 *Id.*

35 *Id.*

36 *Guidelines on the Application of Article 101 of the Treaty on the Functioning of the European Union to Technology Transfer Agreements* 2014 O.J. (C 89/03) ¶¶ 259-61.

37 Cournot complements theory indicates that prices will be higher when complementary inputs are monopolized by different suppliers acting independently. *See generally* ANTOINE-AUGUSTIN COURNOT, RECHERCHES SUR LES PRINCIPES MATHÉMATIQUES DE LA THÉORIE DES RICHESSES (1838). There were some joint announcements including several SEP owners (that were also major device implementers at the time) that aimed to limit aggregate rates in 3G and 4G. However, others were not and should not be bound by such statements, and some have publicly rejected any suggestion they should be. As stated by Qualcomm in 2008, "Contrary to recent claims by a small number of manufacturers, FRAND does not,

authorities prevent anticompetitive effects by requiring that substitutes are not included in patent pools. This is one reason why essentiality checking is sometimes required by competition authorities to be undertaken by patent pools. Standard setting requires the selection of the best technology to perform a particular function so any alternative patented technology will not be included.

However, it should also be recognized that there is no clear line between implementers and SEP owners. Many SEP owners are also SEP implementers manufacturing or selling standard-based products. Submissions as SEP owners would thus likely also reflect some interests — possibly predominating interests — as implementers, and vice versa. Examination of patent pooling practices illustrates that rates for these agreed by the SEP owners are significantly affected by some of them also being major implementers that have more to gain through reduced outpayments at lower royalty rates than they would gain from higher in-payments if royalty rates were higher.[38] An extreme example of this phenomenon is in royalty-free pooling that dominates in the licensing of the widely adopted Bluetooth, USB, and DOCSIS standards. SEP owners forgo the possibility of any royalty income so they can implement standards without having to pay any royalties.

4. Governance, Process and Quality Control in Expert Determinations

Robust economic and statistical processes are required in rate setting. Scientific principles should be applied, including the need for reproducibility of results. The proposed legislation requires that "[t]he checks will be conducted based on methodology that ensures a fair and statistically valid selection capable of producing sufficiently accurate results about the percentage of truly essential patents among each SEP holder's registered SEPs."[39] The impact assessment also hopes that "if the register will be perceived by SEP holders as a means of indicating portfolio strength (and e.g. used in negotiations to determine the share of aggregate royalty applicable to them), they may increase the number of registered patents."[40]

It is unclear how the EUIPO will ensure that quality and consistency in rate setting and apportionment is achieved. Nevertheless, the aggregate rates set and their apportionments will seemingly be justified by the impressive

---

and never has, prescribed formulas for imposing cumulative royalty caps or proportional allocations of such royalty caps. Such formulas would arbitrarily limit the value of standards essential patents, discourage innovation, encourage the filing of marginal patents, complicate and delay the standardization process, and be impossible to implement in practice." Qualcomm, *LTE/WiMax Patent Licensing Statement* (Dec. 2008), www.qualcomm.com/content/dam/qcomm-martech/dm-assets/documents/lte-wimax-patent-licensing-statement_1.pdf.

38 Mallinson August, *supra* note 12, at 4-6.

39 European Commission, *supra* note 12 at 13, Art. 29.

40 Impact Assessment Report, *supra* note 15, at 37-41.

academic and other credentials of those who are chosen to make such determinations.

It is fanciful to believe that sub-contracting to a slew of economic, technical, and other kinds of experts to *make up* aggregate royalty values will produce better, fairer, or truer rates than those derived in market-based rates negotiated in bilateral licenses and offered by patent pools in competition to those on behalf of coordinated collections of SEP owners. Instead, processes will be susceptible to political capture, and rates derived will be significantly affected by interest group lobbying, self-interest, or conflicted interests of external experts as the proposed new Competence Centre is set up, governed, and operated.

Possibly even worse, absent adequate governance, leadership, and some standardization in the evaluation methods employed, results produced will be inconsistent and derivations will be opaque.

Expert opinions about aggregate and individual FRAND rates vary considerably. Empirical research also shows that different assessors tend to disagree with each other in around one quarter of their essentiality determinations.[41] That is worse than it might seem given that they would agree with each other in 50% of their determinations if one of the assessors was randomly making determinations based only on a coin flip.[42] High levels of disagreement on individual patents do not cancel out even when determinations are made on numerous patents. Shares of total patent counts determined essential for individual major SEP owners differ between assessors by double-digit multiples in some cases.[43]

Ericsson's feedback to the Commission on the proposed EU regulation identifies major concerns about how expertise will be applied including that:

> There also seems to be some confusion over the exact role the competence center will play going forward as many of the tasks the center is mandated with will be executed by external consultants. Indeed, the essentiality assessment, FRAND determination and aggregate royalty opinion will be done by external evaluators and conciliators. This was recently confirmed by the executive director of EUIPO Christian Archambeau indicating that "the EUIPO will not be an 'expert' as such in patent issues but will work as an administrative entity."
>
> Thus, there seems to exist no plan to build up patent (or standard) expertise within the competence center. This is worrying as it is unclear how the center will be able to evaluate the quality of the work performed by the external advisors, their independence or to

---

41 Bekkers et al., *supra* note 28; Rudi Bekkers et al., *Overcoming Inefficiencies in Patent Licensing: A method to assess patent essentiality for technical standards*, 51 RSCH. POL'Y 104590 (2022).

42 Keith Mallinson, *Essentiality Rate Inflation and Random Variability in SEP Counts with Sampling and Essentiality Checking for Top-Down FRAND Royalty Rate Setting* (Nov. 24, 2021), https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3933944 [hereinafter Mallinson, *Essentiality Rate Inflation*].

43 Keith Mallinson, *Do Not Count on Accuracy in Third-Party Patent-Essentiality Determinations*, IP FIN. (May 12, 2017), http://www.ip.finance/2017/05/do-not-count-on-accuracy-in-third-party.html.

> ensure consistency of the work performed by them. How will the center be able to train the external consultants if its tasks are of a purely administrative nature?[44]

Deriving applicable aggregate rates and determining patent essentiality, let alone portfolio value, are highly subjective processes that tend to produce widely varying results. Consistency of procedures and in outcomes is required.

Notwithstanding hazards such as the threat of political capture, and the need to safeguard against that, there needs to be some intellectual leadership on how to set aggregate rates, or select them from among proposals, and then apportion them among licensors. The EUIPO evidently lacks that competence. Governance, operational processes, and evaluation techniques are undefined but need to be sound to ensure consistent quality with economically optimal and fair rates. For example, the proposed regulation merely states that the "examination of essentiality shall be conducted following procedure that ensures sufficient time, rigorousness and high-quality."[45] Specifics have been left for an implementing act.[46]

Left to their own devices, aggregate rates set by different conciliators acting independently will likely come up with aggregate royalty figures that will be very disparate. As noted by Justus Baron, in 2023, "[o]verall, the process described in the proposed SEP regulation is likely to result in disparate and largely arbitrary opinions on aggregate royalties." Article 18 of the proposed EU legislation also provides for the Competence Centre to provide a "non-binding expert opinion on a global aggregate royalty." With the evident major differences of opinion among experts, outcomes will be a haphazard "luck of the draw" given that the determination will be made by a conciliator or majority voting in a panel of three conciliators.[47]

In the real world, so far, experts only provide input to final decisions made by others on royalty rates and other important terms in FRAND licensing. Experts advise their respective clients who make final agreements in bilateral negotiations. In litigation, determinations follow the courts' consideration of often widely differing expert testimony from opposing parties. In recent litigation, including *Unwired Planet v. Huawei*,[48] *TCL v. Ericsson*,[49]

---

[44] Telefonaktiebolaget LM Ericsson, Julia Brito, Comment Letter on Proposal for a Regulation of the European Parliament and of the Council on Standard Essential Patents and Amending Regulation 9 (Aug. 10, 2023), https://ec.europa.eu/info/law/better-regulation/have-your-say/initiatives/13109-Intellectual-property-new-framework-for-standard-essential-patents/F3434449_en (citation omitted).

[45] European Commission, *supra* note 12 at 45, Art. 31(1).

[46] *Id.* at 42-43, Arts. 26(5), Art. 29(1).

[47] Justus A. Baron, *The Commission's Draft SEP Regulation – Focus on Proposed Mechanisms for the Determination of Reasonable Aggregate Royalties* 13-16 (Aug. 14, 2023), https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4537591.

[48] *Unwired Planet*, [2020] UKSC 37.

[49] 943 F.3d 1360.

Interdigital v. Lenovo,[50] and *Optis v. Apple*,[51] metrics and amounts are finally determined by the courts in FRAND trial decisions or by the parties in settlement negotiations (e.g., following the unanimous and entire reversal on appeal of the *TCL v. Ericsson* decision).

As noted by Qualcomm in its feedback to the Commission on the proposed EU regulation:

> In two of the last three major FRAND determination decisions – *Unwired Planet v. Huawei* and *IDC v. Lenovo* – sophisticated parties with tens to hundreds of millions of dollars at stake, with access to the best experts and advocates in the world, in a forum with due process and procedural safeguards, and with a lot more time to develop a case than eight months were unable to provide reliable evidence from which the courts could find an aggregate royalty. In the other case, *Optis v. Apple*, the court rejected both parties' arguments that there was a single correct aggregate rate applicable to all, and instead calculated an aggregate applicable only to Apple and based on a subset of Apple's own license agreements. Thus, in each of these cases , [sic] despite the parties' best efforts, the courts found no reliable evidence from which they could derive a broadly applicable aggregate rate. There is no reason to believe – and much reason to doubt – that the abbreviated "opinion" proceeding of the Proposal would achieve a different result.[52]

The *Interdigital v. Lenovo* decision illustrates how disparate evaluations can typically be despite parties each spending millions of dollars in expert fees over a year or so. Parties differed by a factor of 4.2 in their last FRAND offers.[53] The court's $138.7 million lump sum award was much closer to Lenovo's $80 million offer than it was to Interdigital's most recent "5G Extended Offer" including a complex collection of terms that was translated into an equivalent lump sum figure of $337 million by Interdigital's accountancy expert. Interdigital agreed that the court should determine a lump sum.[54] The court shunned ad valorem rate comparisons and based its comparisons on monetary amounts per unit.[55] In the *Interdigital v. Lenovo* decision Justice Mellor was critical of Judge Selna stating in his *TCL v. Ericsson* decision that "'Ericsson's use of floors in its rates is itself discriminatory.'"[56] Despite extensive expert work, Justice Mellor rejected the top-down approach, even only as a cross-check, at least "as pleaded" in that case.[57]

---

50 Interdigital Technology Corporation & Ors. v Lenovo Group Ltd., [2023] EWHC 539 (Pat).

51 Optis v. Apple, [2023] EWHC 1095 (Ch).

52 Qualcomm Inc., Jillian Mertsch, Comment Letter on Proposal for a Regulation of the European Parliament and of the Council on Standard Essential Patents and Amending Regulation 10 (Aug. 10, 2023), https://ec.europa.eu/info/law/better-regulation/have-your-say/initiatives/13109-Intellectual-property-new-framework-for-standard-essential-patents/F3434463_en (citations omitted).

53 *Interdigital Technology*, [2023] EWHC 539, at [20]-[22], [26].

54 *Id.* at [20]-[22], [26], [944].

55 *Id.* at [22].

56 *Id.* at [268]-[269] (quoting TCL Commun. Tech. Holdings, Ltd. v. Telefonaktiebolaget LM Ericsson, No. CV 15-2370 JVS(DFMX), 2017 WL 6611635, at *57 (C.D. Cal. Dec. 21, 2017)).

57 *Interdigital Technology*, [2023] EWHC 539, at [881]

C. *Agency Policy and Proposed New Regulation in US With the SERA*

The US is eschewing FRAND rate-setting regulation and stipulation of rote valuation methods by withdrawing guidance from government agencies including the USPTO, NIST and DoJ, while proposed law-making has also been neutered in this aspect.

Following a couple of months' public consultation on a draft Policy Statement issued 6 December 2021,[58] on 8 June 2022 the USPTO, NIST and DoJ formally withdrew their joint 2019 Policy Statement while also indicating that the 2013 Policy Statement was not being reinstated.[59] These agencies decided that the courts were best placed in furthering "the interests of innovation and competition" in SEPs and FRAND licensing, "and as enforced by DoJ and other agencies," without any of these three policies.[60]

Calls for new legislation arise from concerns about foreign (particularly Chinese) anti-suit injunctions, and judicial determinations of global FRAND rates have prompted US federal legislators to propose regulation to reduce the effect of foreign proceedings on US patents. Two such bills were proposed to the Senate Judiciary Committee in 2022: the Defending American Courts Act (DACA)[61] and the Standard Essential Royalty Act (SERA).[62]

The proposed SERA legislation alleges — without support — that piecemeal adjudication of SEPs has resulted in inconsistent awards, in some cases an unreasonable cumulative rate and has denied American manufacturers licenses on reasonable terms. It also dubiously asserts "*in the absence of an efficient system in the United States for determining reasonable royalties for standard-essential patents*, some patent owners and manufacturers have

---

58 U.S. Dep't of Just., Draft Policy Statement on Licensing Negotiations and Negotiations and Remedies for Standards-Essential Patents Subject to Voluntary F/RAND Commitments (Dec. 6, 2021).

59 U.S. Dep't of Just., Withdrawal of 2019 Policy Statement on Remedies for Standards-Essential Patents Subject to Voluntary F/RAND Commitments (June 8, 2022).

60 *Id.* at 1-2.

61 S. 3772, 117th Cong. (2021). "If enacted, DACA would impose two types of penalties on a party that seeks to restrict an action for patent infringement before a U.S. court or the International Trade Commission (ITC) through the assertion of a foreign anti-suit injunction." Jorge L. Contreras, *A Statutory Anti-Anti-Suit Injunction for U.S. Patent Cases?*, 355 UTAH L. FAC. SCHOLARSHIP 4 (2022).

62 Standard Essential Royalties Act (Proposed Legislation), SENATE JUDICIARY COMMITTEE (2022), https://view.officeapps.live.com/op/view.aspx?src=https%3A%2F%2Fipwatchdog.com%2Fwp-content%2Fuploads%2F2022%2F11%2FSERA-text.docx&wdOrigin=BROWSELINK. ("Purpose: To provide a . . . system for adjudicating reasonable royalties for patents that are essential to the implementation of interoperable technical standards"); William New, *Draft US Bill Proposes Federal SEP Royalty Court with Global Impact*, IAM, (Oct. 18, 2022), https://www.iam-media.com/article/draft-us-bill-proposes-federal-sep-royalty-court-global-impact; Jorge Contreras, *National FRAND Rate-Setting Legislation: A Cure For International Jurisdictional Competition In Standards-Essential Patent Litigation?*, CPI ANTITRUST CHRON., at 6 n.41, 7 (July 13, 2022), https://www.pymnts.com/cpi_posts/national-frand-rate-setting-legislation-a-cure-for-international-jurisdictional-competition-in-standards-essential-patent-litigation/ (claiming the proposed SERA "embodies some of the recommendations contained in this essay," many of which are highly interventionist and "resemble rate-setting hearings that are currently conducted with respect to utility rates and various forms of copyright licensing.").

resorted to foreign courts to set royalties for patents issued by the United States" and that "a foreign court's compelled adjudication of royalties for a United States patent violates the sovereignty of the United States."[63] The SERA would create a US judicial tribunal, to be known as the "Standards Royalty Court," to determine FRAND rates for US SEPs, notwithstanding the findings of any foreign court. According to one public policy commentator, the SERA "proposes a new federal court to decide FRAND rates where there are inconsistencies across domestic rulings, or where foreign courts hand down verdicts that disadvantage American patent holders. The bill is clearly written with China's and Europe's standard essential patents regimes in mind."[64]

However, 18 months on the SERA seems to be going nowhere with rate-setting. Alternative suggestions being discussed privately among interested parties now are light touch and require balance in disclosure obligations, for example, with a registry of patents declared potentially standard essential by their owners and a registry of devices declared standard-compliant by their producers.

D. *Transatlantic Comparisons*

The US and Europe are heading in different directions on how SEP royalties are determined in FRAND licensing disputes. US authorities are increasingly hands off, while proposed EU legislation constrains SEP enforcement and prescribes a valuation methodology, which a Chinese court has recently used to drastically and defectively undercut established rates.[65]

The recent EU Proposal has some similarities with the initially proposed SERA, but also has notable differences. That proposed version of SERA also anticipates determination of "an overall reasonable royalty rate or rates for implementation of the technical standard" and "each plaintiff's entitlement to its appropriate portion of that royalty rate."[66] However, the US proposal is for a court (without jury) where "[c]ases and controversies shall be heard and determined by a panel of at least three judges" with ability to demand "production of information or evidence from persons who are not a party to the action" and that can make binding rulings, rather than an administrative

---

63 Standard Essential Royalties Act (Proposed Legislation), SENATE JUDICIARY COMMITTEE, §§ 2(9)-2(10) (2022) (emphasis added).

64 Marc L. Busch, *In the Latest 5G Fight, the US Should Support Market-based Patent Fees*, THE HILL (July 19, 2023), https://thehill.com/opinion/technology/4103521-in-the-latest-5g-fight-the-us-should-support-market-based-patent-fees/.

65 Keith Mallinson, *Race to the Bottom with Top-down Approach in FRAND Rate Setting for SEPs*, IAM (Jan. 23, 2024), https://www.wiseharbor.com/wp-content/uploads/2024/02/Race-to-the-bottom-with-top-down-approach-in-FRAND-rate-setting-for-SEPs-IAM.pdf [hereinafter Mallinson, *Race to the Bottom*].

66 Standard Essential Royalties Act (Proposed Legislation), SENATE JUDICIARY COMMITTEE, § 4.334(b) (2022).

agency with unclear governance and without such subpoena powers or legally binding authority.[67] Exhibit 1 summarizes some of the areas of commonality and divergence.

Exhibit 1: Comparison of Recent US and EU FRAND Tribunal Proposals[68]

	Proposed US Standard Essential Royalty Act (SERA) (June 2022)	**Proposed EU SEP Regulation (Mar. 2023)**[69]
Tribunal	A new federal court	EUIPO, an EU administrative agency
Authorization of collective negotiation of aggregate royalty burden	No	Yes
Binding effect	Binding in US	Non-binding
Effect on foreign FRAND determinations	Overrides foreign FRAND determinations for US patents	None
Confidentiality of decision	No	Yes
Creation of SEP registry	No	Yes
Essentiality testing	Possibly, though not required	Yes

Following a public consultation in 2022, the Commission sought written feedback on the new legislation it proposed in April 2023. In contrast to the extensive public debate and lobbying surrounding the proposed EU legislation since a draft version of it was leaked in March 2023, including 78

67 *Id.* at § 3.221.

68 Jorge Contreras, *The EU's Response to National Judicial Determinations of FRAND Royalty Rates,* PATENTLY O, (April 13, 2023), https://patentlyo.com/patent/2023/04/response-national-determinations.html.

69 This was a leaked draft version of the proposed regulation before its publication by the Commission. Nevertheless, depictions in this table remain consistent with final version of the proposed regulation that was published on April 27, 2023.

submissions to the Commission before its August 2023 deadline, there has been rather less public discourse on the proposed SERA legislation so far. Instead, in the US, various practitioners including licensees and licensors have worked together privately to seek common ground, compromise, and balance in improving disclosures for greater transparency and predictability in FRAND licensing.

The rest of this article focuses on practical matters in the setting of aggregate royalties for SEPs and in determining FRAND rates for individual licensors through apportionment of aggregate rates. Analysis is largely in consideration of the Commission's detailed proposals, but is also broadly applicable elsewhere, including in the US where the initially proposed SERA also requires aggregate royalty rate-setting and apportionment.

III. RATE SETTING AND APPORTIONMENT METHODS

A. *Definitions, Metrics, and Objectives*

According to the proposed EU legislation, "'aggregate royalty' means the maximum amount of royalty for all patents essential to a standard."[70] The Commission also indicates "uncertainty about the SEP royalty burden"[71] and that "[s]takeholders consider that the FRAND licensing concept could benefit greatly from some clarification, notably with regard to the determination of an aggregate royalty burden."[72]

Even the basis, as well as the level, of aggregate royalty rates in joint notifications will vary confusingly. For example, a group of SEP owners could announce an aggregate rate of $10 per end-product, another group announce a rate of 5% of the end-product price, while a third group would prefer a lower $1 per-product rate. And many licenses indicate lump sum payments. Translating between running-rate ad valorem and monetary amounts per unit, and between these and lump sum payment figures — in order to make comparisons — is always highly dependent on various subjective and often questionable assumptions. Aggregate royalty rates proposed to or set by the EUIPO's conciliators could be in quantification of the total payment burden or of the rate to be used in determining individual FRAND royalty rates with the *top-down approach* that apportions royalties among patent owners based on the relative strengths of their SEP portfolios.[73] The latter Aggregate Royalty Rate for Apportionment (ARRFA) should be a higher

70 European Commission, *supra* note 12, at 27, Art. 2(10).

71 *Id.* at 8.

72 Impact Assessment Report, *supra* note 15, at 21.

73 European Commission, *supra* note 12, at 27, Art. 18(1) (stating that "[a] SEP holder or an implementer may request the competence centre for a non-binding expert opinion on a global aggregate royalty").

figure than the former to allow for SEPs that remain unlicensed and for which there is no payment.

Any aggregate royalty rates set must be precisely defined, derived, and applied. Aggregate rate setting for standards, as proposed by the Commission, will enable proposed rates to be depicted and manipulated in ways which are anticompetitive, unfair, and will under-value patented standard-essential technologies.

B. *Royalty Burden*

Aggregate royalty figures might be gleaned or derived somehow from among various different formulations of aggregate rates reported. However, these reported rates vary enormously, for example, global rates from more than 35% to less than 5% of a smartphone's selling price. The correct ARRFA for a top-down approach FRAND determination and the rather lower maximum aggregate rate implementers will need to pay will fall well within those two extremes.

In FRAND determinations for bilateral licensing there is always a shortfall between the ARRFA and what is actually paid because the SEPs in any given standard are never fully licensed. The aggregate rates from which bilateral licensing rates are derived are never fully paid due to notional royalty allocations to patents that remain unlicensed. Any aggregate royalty setting must recognize this difference if such rates are to be used to determine FRAND rates using the top-down approach.

To mitigate shortcomings in rate setting, some guiding principles must be established on what the "SEP royalty burden" and ARRFA should include and exclude, as well as how and by whom such rates should be derived and applied. The interests of both SEP owners and implementers must be safeguarded while reflecting industry realities with the many factors that shape varied financial and other terms in established licenses. Application of economic theory must have full and proper regard for what royalty figures reported in the industry represent and how licensing actually gets done. FRAND licensing is about various terms, not just rates.

However, there is no consensus even on whether there should be some kind of aggregate royalty capping, let alone what figures these should be or which methods ought to be used to derive them.[74] For example, some patent owners publicly disagreed with setting aggregate royalty goals at all, as announced by some other patent owners and technology implementers for 3G

[74] Jorge Contreras, *Aggregated Royalties for Top-Down FRAND Determinations: Revisiting Joint Negotiation*, 65 UTAH LAW FACULTY SCHOLARSHIP (2017), https://dc.law.utah.edu/cgi/viewcontent.cgi?article=1064&context=scholarship.

and 4G in mobile phones in the 2000s.[75] Even those making the announcements did not regard aggregate figures as caps.

Even defining aggregate royalty is debatable: is this total a theoretical maximum that nobody would ever pay, a typical or average figure that would be or is actually paid after royalty-base caps (i.e., a different kind of cap than above) or sales volume discounts and with many patents remaining unlicensed? Or is it something in between? In my seminal research on aggregate royalty charges in 2015, I rebutted a common but speculative narrative based on misapplication of mid-19th Century economic theory regarding commodity complements — asserting that royalty charges could "stack" to as much as 30% of smartphone selling price — with my empirical proof that rates paid averaged no more than around 5%.[76] The difference is due to many factors including unlicensed patents, royalty-base caps, volume discounts, geographic discounts (e.g., for China), cross-licensing and pass-through rights bundled with chipset sales, as well as wishful thinking with the inflated expectations and demands of some patent owners.

Royalty charges — in lump sums, monetary figures per unit or ad valorem percentage rates, as parties agree — like most other negotiated prices, are usually established through consideration of market factors including value to customers, costs, and competition among various players.

C. *Top-Down Approach*

The top-down approach in deriving royalty charges for standard-essential patents requires the setting of aggregate royalties for specific standards and applications. These rates are then notionally apportioned among patent owners — typically including those that do not license and will never collect any royalties — based on a patent strength metric. Top-down approach rate determinations have been proposed to the courts by litigants in various SEP FRAND trials for more than a decade.[77] The top-down approach has several

75 For example, in 2008, Qualcomm stated "Contrary to recent claims by a small number of manufacturers, FRAND does not, and never has, prescribed formulas for imposing cumulative royalty caps or proportional allocations of such royalty caps. Such formulas would arbitrarily limit the value of standards essential patents, discourage innovation, encourage the filing of marginal patents, complicate and delay the standardization process, and be impossible to implement in practice." LTE/WiMax Patent Licensing Statement, QUALCOMM (Dec. 2008), www.qualcomm.com/content/dam/qcomm-martech/dm-assets/documents/lte-wimax-patent-licensing-statement_1.pdf.

76 Mallinson, *Cumulative*, *supra* note 4; *see also* Galetovic et al., *supra* note 4 (replicating, validating, and refining my analysis and findings in Mallinson, *Cumulative*, *supra* note 4); Sidak, *supra* note 4 (replicating, validating, and refining my analysis and findings in Mallinson, *Cumulative*, *supra* note 4).

77 *See generally* In re Innovatio IP Ventures, LLC Patent Litig., 2013 WL 5593609 (N.D. Ill 2013).

major shortcomings, as indicated below, and as I have explained previously elsewhere.[78]

Top-down apportionment is usually by some kind of patent count. Even top-down approach advocates have differing opinions about which patent strength metric to use — the number of declared-essential patents, number of independently-assessed-essential patents or number of contributions to the standard.

The top-down approach apportions an aggregate royalty figure to derive the different FRAND royalty rates for individual SEP owners. The top-down approach calculation is usually made to derive the royalty rate for a licensor using this apportionment formula:

Licensor's royalty rate (R) = aggregate rate (T) x licensor's share (S) of SEPs

R is the rate to be applied to actual sales prices or revenues.[79]

In the case of *Unwired Planet v. Huawei*, the court was unwilling to set a top-down rate due to the uncertainties in doing that.[80] Instead, the court used the same apportionment formula the other way around to *imply* an aggregate rate burden from comparable licenses (comps), as a cross check.

Implied aggregate rate (T') = Licensor's royalty rate implied from comps (R') ÷ S

This is a crucial difference in use of the same simple algebraic formula because T' is implied rather than set as it is in conventional use the formula. It would have been more apposite to call this a bottom-up method, but that term had already been bagged for another valuation method.

---

78 Keith Mallinson, *Unreasonably-low Royalties in Top-down FRAND-rate Determinations for* TCL v. Ericsson, IP FIN. (Apr. 30, 2018), http://www.ip.finance/2018/04/unreasonably-low-royalties-in-top-down.html [Hereinafter Mallinson, *Unreasonably-low*].

79 Much larger aggregates of headline maximum royalty rates before any discounting, as in licensors' individual rate card disclosures, than in the aggregates of rates actually paid after discounting is only to be expected. For example, if a licensee sells a handset for $400 where rates are subject to $200 cap, the royalty percentage rate actually paid will be only half as much as the headline royalty rate percentage.

80 Unwired Planet Int'l Ltd. v. Huawei Tech. Co. Ltd., [2017] EWHC 2988, at [268]-[269] (Pat) ("the main conceptual difficulty I have with the using a total stack in a top-down approach as opposed to using it as a cross-check is in the selection of the total royalty burden T to start with. In my judgment the statements set out above have little value in arriving at a benchmark rate today for a number of reasons. The claims are obviously self-serving. The statements about aggregate royalties in particular are statements about other people's money on the footing that the person making the statement says at the same time that the cake is quite small but they are entitled to a large piece of it").

This bottom-up use of the top-down formula is also how pejorative "royalty stacking" is sometimes alleged but cannot be proven. The Commission, among others, redefines royalty stacking as a counterfactual scenario.[81]

Absent evidence that anybody is actually paying aggregate rates as high as 20%, 30%, or even more, hypothetical assertions along the following lines are constructed. For example, *if* — these conjectures always start with this word — company A demands a 1% royalty while owning only 3% of the SEPs reading on a standard, then the aggregate royalty would be *1% ÷ 4% = 25%*.[82] However, as my subsequent analysis in this article shows, royalties paid are a long way below this hypothetical level for a variety of reasons. There is not, actually, any royalty stacking.

These aggregate royalty rates are absent cross-licensing effects that reduce net payments. All the above algebra is applied to one-way royalty rates (i.e., after any cross-license payment figures have been grossed-up in "unpacking"). Rates actually paid after any cross-licensing are lower than one-way rates.

Apportionment is based on the faulty premise that the relative value of different patent portfolios is directly proportional to the number of patents in each of these. On the contrary, there is abundant evidence that the value of patents, including SEPs, varies enormously. Some patented technologies are crucial in creation or improvement of standards; others, such as those reading on parts of the standard that are optional and are rarely or never implemented, are worth very little. The top-down approach ignores whether products actually infringe. Some SEPs read on optional parts of the standards that are not implemented in all products, and in some cases in none of them. Some SEPs relate to devices, and others relate to network equipment. The top-down approach ignores validity. Top-down only seeks to determine fair and reasonable royalties overall and on average for all licensees. It makes no attempt to determine non-discriminatory variations in rates among differently situated licensees.[83] For example, Small and Medium-sized Enterprises (SMEs) are

---

81 Baron, *supra* note 47, at 7-11.

82 *See, e.g.*, Microsoft Corp. v. Motorola Inc., No. C10-1823JLR, 2013 WL 2111217, at *73 (W.D. Wash. Apr. 25, 2013) ("*If* each of these 92 entities [owners of Wi-Fi SEPs] sought royalties similar to Motorola's request of 1.15 % to 1.73% of the end-product price, the aggregate royalty to implement the 802.11 Standard, which is only one feature of the Xbox product, would exceed the total product price. The court concludes that a royalty rate that implicates such clear stacking concerns cannot be a RAND royalty rate") (emphasis added). Elsewhere, evidence of actual stacking has been required by the court but has never been forthcoming. *See, e.g.*, Ericsson Inc. v. D-Link Sys. Inc., No. 6:10-CV-473, 2013 WL 4046225, at *18 (E.D. Tex. Aug. 6, 2013), (agreeing with Ericsson statements that the defendants' royalty stacking argument is theoretical, and that the defendants' expert failed "to present evidence of an *actual* stack on the 802.11n essential products") (emphasis in original).

83 This was evidently one of several reasons why Justice Mellor rejected the top-down approach in *Interdigital v. Lenovo* [2023] EWHC 539, [945] (Pat), in which discrimination in royalty charging through volume discounting was a most significant and contentious issue. *Id.* at [557]. The court recognized that while Judge Selna had used the top-down approach in *TCL v. Ericsson*, Judge Selna was also mindful of

markedly different companies from the few large SEP licensees such as OEMs Apple, Samsung, Sony, and Xiaomi that dominate smartphone product supply and can bargain for volume discounts from all their suppliers.

D. *Apportionment*

Top-down apportionment is usually by some kind of patent count. Even top-down approach advocates have differing opinions about which patent strength metric to use.[84]

1. Counting Declared-Essential Patents

Counting raw declared-essential patents that remain unchecked for essentiality by any third party is widely regarded as inaccurate and unreliable because there is no constraint on patent owners distorting this measure of their patent portfolios' strengths by making excessive declarations. These bloat the denominator in essentiality rate calculations and inflate the positions of patent owners that are most liberal and voluminous in their declarations. There is a conflict between the patent policies of Standard Setting Organizations (SSOs) that encourage liberal declaration of any patents owners believe might be or might become essential to ensure standards are not blocked, and the separate use of patent counts by other organizations as metrics of patent strength. The term "over-declaration" has been coined due to the distortions this causes in the latter. Over-declaration comes in two forms: declaring excessive numbers of patents and declaring individual patents excessively to multiple technical specifications within standards.[85]

---

this shortcoming and the superiority of comparable license benchmarks: "A top down method, however, cannot address discrimination as the Court interprets the term, and is not necessarily a substitute for a market-based approach that considers comparable licenses." TCL Commun. Tech. Holdings, Ltd. v. Telefonaktiebolaget LM Ericsson, No. CV 15-2370 JVS(DFMX), 2017 WL 6611635, at *9 (C.D. Cal. Dec. 21, 2017).

[84] For example, while Apple advocates simply counting patents declared essential by their owners, APPLE, *A STATEMENT ON FRAND LICENSING OF SEPS*, https://www.apple.com/uk/legal/intellectual-property/frand/ (last visited Jan. 20, 2024) ("A SEP licensor's *pro rata* share of declared SEPs is an objective reference point in a FRAND negotiation"), the European Commission and many others believe that independent essentiality checks are required for measurement of portfolio patent strength. European Commission, *supra* note 12, at 12.

[85] Keith Mallinson, *Gaming the System: A Scatter-Gun Approach to 5G Declarations*, IP FIN. (Dec. 5, 2022, 8:38 PM), http://www.ip.finance/2022/12/gaming-system-scatter-gun-approach-to.html.

2. Essentiality Checking and Random Sampling

While independent essentiality checking is widely demanded, this is not straightforward, and various mechanisms are proposed for this.[86] Many patent owners, implementers, and others prefer that patents are also checked for essentiality by someone other than the patent owner. With many tens of thousands of declared patents, that is very costly, and yet checking is inaccurate and subject to significant biases, with false positive essentiality determinations tending to exceed false negatives.[87] While sampling can significantly reduce the overall size of the task, random sampling errors, and non-random errors as well as random errors in essentiality determinations, must be considered in designing and evaluating patent counting studies.

Checking only samples of patents can significantly reduce costs, even if sampled patents are more thoroughly checked and even with the additional cost of claim charts. Nevertheless, sample sizes in the thousands per SEP licensor are likely to be required for adequate precision — particularly if true essentiality rates are low (e.g., at only around 12% for 4G and 8% for 5G, according to some experts).[88] This is because random sampling errors increase as a proportion of decreased essentiality rates.

Unfortunately, any use of sampling is problematic with determination errors. For example, if only one in ten patents is sampled, any determination errors and corrections after "re-checks" or appeals will have a 10-fold impact on total patent counts inferred by extrapolation. Allowing appeals on essentiality determinations of randomly sampled patents is likely to exacerbate rather than correct bias.[89] Appeals against determinations will inevitably not be random.

However, I also believe that parties must generally be able to challenge individual determinations or patent counts somehow. A right to appeal in case of error and inaccuracy is a basic right which must be preserved.

The Commission's impact assessment is confusing and misleading in its statement that "false positive and false negative random errors tend to

---

86 *See* Giuseppe Colangelo, *Finding an Efficiency-oriented Approach to Scrutinise the Essentiality of Potential SEPs: A survey,* 18 OXFORD ACAD. J. OF INTELL. PROP. L. & PRAC. 502, 505 (2023).

87 *See* Keith Mallinson, *Essentiality Checks Might Foster SEP Licensing, but Do Not Stop Over-Declarations from Inflating Patent Counts and Making Them Unreliable Measures*, WISEHARBOR (Nov. 16, 2022), https://www.wiseharbor.com/wp-content/uploads/2022/12/Mallinson-WiseHarbor-SEP-overdeclarations-2022.12.05.pdf [hereinafter Mallinson, *Essentiality Checks*]; Justus Baron & Tim Pohlman, *Precision and Bias in the Assessment of Essentiality Rates in Firms' Portfolios of Declared SEPs* (Nov. 2021), https://wwws.law.northwestern.edu/research-faculty/clbe/events/standardization/documents/baron_pohlmann_bias_and_precision_essentiality_rates.pdf.

88 Mallinson, *Essentiality Rate Inflation*, *supra* note 42.

89 Impact Assessment Report, *supra* note 15, at 101 n.294 (citing European Commission, Directorate-General for Internal Market, Industry, Entrepreneurship and SMEs, Justus Baron, *Essentiality Checks for Potential SEPs – Framework for assessing the impact of different policy options* (2023) https://data.europa.eu/doi/10.2873/002897).

cancel each other out."[90] The terms "false negative" and "false positive" in the context of essentiality checking and patent counting are usually understood to apply to individual essentiality determination errors rather than random errors in the totals of essential and not essential patents in an entire sample. It is true that random sampling errors do tend to cancel each other out (i.e., they may be substantial in any given sample, but at least they are unbiased from one sample to another). In contrast, false positive and false negative determination errors are perniciously not entirely random, do not tend to cancel out, and can result in significant bias.[91]

Consequently, checking can provide a false sense of security and precision. Over-declaration, by some patent owners, is only mitigated, not eliminated, by checking. The more a patent owner over-declares, the more inflated its patent counts and essentiality rates will be — even with checking.

Some interested parties prefer not to count patents at all and instead count the numbers of technical contributions that are approved by SSOs to be included in the standards. Among the advantages of this approach is its low cost in comparison to checking numerous patents for essentiality. Approved contributions are one of the metrics that is used by Avanci in its 4G automotive licensing programs and that is thus accepted as a valuation method by its 56 licensors and many automotive OEMs accounting for more than 80% of connected vehicle sales.

E. *Aggregate Royalty Valuation Measures*

Aggregate rate setting goes far beyond satisfying a requirement for transparency on royalty rates, which could generally be provided with disclosure of existing agreements, related ex-ante assumptions (e.g., volume and pricing forecasts in support of lump-sum figures) and royalty figures paid.

Fair and reasonable aggregate royalty figures ought to be based upon the value that the standardized technology confers. That value could be realized in higher product prices than those without the technology, increased demand volumes, or lower costs. The aggregate royalty rate-setting provisions in the proposed EU legislation must begin between 90 days and 150 days of either the publication of the standard or first sale of new implementations in the EU.

However, markets would not be sufficiently mature for such early determinations of aggregate royalties to be meaningfully estimated from figures in existing licenses or from product pricing. Alternatively, prices can be derived with linear regression in multi-factor hedonic pricing analysis,[92] but a

---

90 *Id.*

91 Mallinson, *Essentiality Checks*, *supra* note 88.

92 *See* Hamish Anderson, *Value of Nature Implicit in Property Prices – Hedonic Pricing Method (HPM) Methodology Note*, OFFICE FOR NAT. STAT. (July 12, 2018),

drawback with this technique is that value in use does not always align with pricing; for example, if pricing is solely based on manufacturing costs. Furthermore, explanatory variables are generally not entirely independent of each other. For example, one hedonic model included talk time and battery capacity variables in mobile phones.[93] It unsurprisingly found these two variables to have significant correlation with a coefficient of 0.71. This collinearity impairs the predictive power of the model. An alternative approach without all these constraints is conjoint analysis in which consumers are quizzed to determine their preferences and price sensitivities for various product capabilities.[94]

However, both methods derive a figure for total economic surplus — not only the proportion of it attributable to the SEP owners. How that surplus should be divided between OEMs and SEP owners overall to come up with an aggregate figure for apportionment among SEP licensors is also a major question. An expert for Interdigital in *Interdigital v. Lenovo* proposed a 50:50 division of the output from his hedonic model. Justice Mellor was having none of that simplistic split. He indicated there was insufficient substantiation to that and there were procedural deficiencies in submitting evidence on this.

There is clearly need for much improvement before any of these methodologies can be used to regulate aggregate royalties reliably.

As I have already indicated, the starting point aggregate figure is typically described as a maximum, but that is ambiguous. Is it supposed to be the maximum:

> i. That could ever potentially be paid on any individual device sold in the nation with strongest patent protection,

---

https://www.ons.gov.uk/economy/environmentalaccounts/methodologies/valueofnatureimplicitinpropertypriceshedonicpricingmethodhpmmethodologynote ("The Hedonic Pricing Method (HPM) relies on the assumption that a class of differentiated products can be broken down in to [sic] a number of characteristics. A combination of these characteristics and the external factors that affect the product determines the price. The most common example of this is property values, where the market price of a property is determined by a combination of structural characteristics (floor area, number of bedrooms, garden, garage and so on) and the socio-economic and environmental characteristics of the surrounding area (quality of schools, access to retail, transport, levels of water/air pollution, proximity to green space and so on)."). *See also, e.g.*, J. Gregory Sidak & Jeremy O. Skog, *Hedonic Prices for Multicomponent Products*, 4 CRITERION *J. INNOVATION* 301 (2019).

93 *See* Sidak & Skog, *supra*, note 93 at 305.

94 Tim Stobierski, *What is Conjoint Analysis & How Can You Use It*, HARV. BUS. SCH. ONLINE BUS. INSIGHTS BLOG (Dec. 18, 2020), https://online.hbs.edu/blog/post/what-is-conjoint-analysis#:~:text=Conjoint%20analysis%20is%20a%20form,of%20their%20products%20or%20services ("Conjoint analysis is a form of statistical analysis that firms use in market research to understand how customers value different components or features of their products or services. It's based on the principle that any product can be broken down into a set of attributes that ultimately impact users' perceived value of an item or service. Conjoint analysis is typically conducted via a specialized survey that asks consumers to rank the importance of the specific features in question. Analyzing the results allows the firm to then assign a value to each one").

ii. Of royalties averaged across all devices sold in that nation in a certain period, or

iii. Of royalties averaged across all devices sold in a certain period?

The devil is in the detail with any averaging versus the hypothetical corner case in (i). For example, in *TCL v. Ericsson,* R was further reduced for geographies where the licensor had fewer SEPs.[95]

The Court found:

> "Ericsson's 4G patent strength in China is 69.80% of its U.S. patent strength" and "that 0.45% is an appropriate FRAND for Ericsson's 4G SEP portfolio in the United States. This means that the FRAND rate for Ericsson's portfolio for the Rest of the World ("RoW") is 0.314%".[96]

The RoW rate is nearly a third less than the US rate. With most sales outside the US, the blended global average set by the court was rather closer to the RoW rate than the US rate.

The court also made a reduction for expired patents in its rate determinations. It included expired patents in the denominator while it excluded them from the numerator in calculating S. This also has a diluting effect on the royalty rate determined. In contrast, patent pools typically remove expired patents from their patent counts in both the numerator and denominator in calculating shares of fees for distribution.

On the other hand, patent portfolios tend to become enriched over the life of a standard or licensing agreement following additional patent applications, declarations, and as patents are granted. Standards are not static. For example, there were numerous improvements to 4G LTE over a decade in a succession of seven standard releases by 3GPP before the first standard release of 5G was completed in 2019.

Was it anticipated in existing licenses that royalty rate figures would reduce over time as patents expired? Alternatively, and more realistically, for example, are rates agreed for simplicity to remain at the same level for the

---

95 These adjustments, for example, as used in the *TCL v. Ericsson* Decision are contentious. *See* TCL Commc'n Tech. Holdings Ltd v. Telefonaktenbologet LM Ericsson, No. 8:14-cv-00341-JVS-DFM, 2018 WL 4488286 (C.D. Cal. Dec. 12, 2017). I first noted this in Mallinson, *Unreasonably-low*, *supra* note 79. However, the issue is not necessarily whether these reductions are made, but whether the aggregate royalty rate used as the top-down approach input correspondingly anticipates such adjustments. Some aggregate figures do, and others do not. If these reductions are taken, then the applicable input figure T needs to be higher than otherwise. For example, with regard to geography, is the aggregate the maximum to be paid where patent protection is strongest, or is it a globally a blended "maximum" across all licensed sales in a given period?

96 TCL, 2018 WL 4488286, at *50-51.

duration of the standard or licensing agreement regardless of expirations and new patent additions?

To be clear, I am not advocating application of adjustments to the royalty rate and apportionment factor as undertaken in *TCL v. Ericsson*, I am merely explaining what was done and stating that, if such an approach is taken, the ARRFA should be set higher, accordingly.

1. Ad Valorem, Fixed Monetary Figures Per Unit or Lump Sums

An aggregate royalty rate — like an individual royalty rate — can be an ad valorem percentage or a fixed monetary figure per unit of licensed product sales.[97] A fundamental question in any aggregate rate setting process is which to select. I am not prescribing or proscribing either. The most applicable and best to select depends on the application.

However, considering how SEP licensing has been agreed and how royalties have mostly been depicted, measured, and compared since the 1990s, I am focusing most of the following analysis in this article on ad valorem percentage royalties as applied to the royalty base of mobile phone selling prices. This is most illustrative because it enables me to draw upon many published aggregate royalty rate figures, which almost invariably until the late 2010s were and still mostly are also ad valorem percentages.

Ad valorem percentage royalty charging suited implementers as average selling prices for handsets reduced substantially throughout the 1990s and until the 2000s when the growth of smartphone sales started increasing overall average selling prices (ASPs). Since then, licensees have increasingly sought to cap the handset price used as the royalty base. On the other hand, with basic mobile phone prices as low as $20 since the mid-2000s, some licensors have also introduced floors to their licensing terms. When ASPs rise above a cap, or fall below a floor, royalty rates become fixed monetary amounts. In some cases, such as Nokia in 5G, its standard charge is a fixed monetary charge of €3.00 ($3.36) per unit. In IoT, where selling prices for licensed items vary enormously (e.g., from as little as $10 for a basic module to typically tens of thousands of dollars for a car), royalty rates as fixed monetary charges per unit tend to make best sense.

If aggregate rates are to be set at all — as they are for patent pools in their rate cards, but in the opinion of many is unnecessary and dysfunctional in bilateral licensing[98] — such rates must be derived in the applicable context.

---

[97] With ad valorem licensing, a royalty percentage rate is multiplied by the royalty base of the licensed product price, or price cap if the product price is higher than that, to derive the monetary figure for the royalty charge.

[98] Various court decisions have avoided or explicitly rejected aggregate rate setting. *See* Unwired Planet Int'l Ltd. v. Huawei Techs. Co. [2017] EWHC 711 (Pat); Interdigital v. Lenovo [2023] EWHC 539 (Pat). *Optis v. Apple* [2023] EWHC 1095 (Ch), also in the UK, also focuses on comparable licensing

Collective action in setting aggregate royalties — such as in patent pools where some major licensors are typically also major licensees — tends to imply individual rates that are lower than would be agreed bilaterally. Another crucial difference is that patent pool aggregate rates are the rates licensees actually pay.

Rates in apportionment calculations and in licensing agreement terms must also reflect whether they are single-mode rates or multimode rates. In cellular, for example, some devices are single-mode (e.g., 4G only and others are multimode (e.g., 2G, 3G, and 4G) with various different combinations of modes, each of which might command different FRAND rates.

2. Benchmarks

a. The Addition of Every SEP Owner's Maximum Wishes

Simply adding up all licensors' maximum royalty rates inevitably produces a hypothetical maximum aggregate royalty rate figure that is inflated far above what anyone would ever pay. For example, before the introduction of 4G LTE in 2009, industry association for mobile network operators NGMN appointed a Trusted-Third-Party (TTP) to collect publicly and privately indicated maximum royalty rates for licensing cellular standards from as many prospective licensees as it could and add up all those rates. In other words, it was attempting to measure a theoretical maximum "stack." Aggregate figures of around 30% for 4G LTE were derived. While this process was ostensibly to increase transparency on royalties, aggregate rate figures were only ever leaked and were never made public officially.

Licensing expert Eric Stasik published a widely-cited 2010 paper adding up the only nine publicly-announced 4G LTE royalty rates at that time for an aggregate royalty of 14.8% from a list of more than 30 firms with patents declared essential to the standard.[99]

No licensee ever paid anywhere near as much as the aggregate rates the TTP derived. Many of the figures in the summation resulted from wishful thinking by SEP owners. Maximum rates are very often reduced by selling price royalty base caps on ad valorem rates and many SEPs go totally unlicensed by any implementer. Fully licensed aggregate rates are thus not paid

benchmarks in its FRAND determinations. However, the very recent *Nokia v. Oppo* decision in China uses comparable license benchmarks and top-down determinations including the first judicially set aggregate royalty for 5G. *See Chongqing No. 1 Intermediate People's Court Sets Global FRAND Rate for 5G SEPs at $0.707/Unit in Nokia/OPPO Case*, CHINA IP L. UPDATE (Dec. 16, 2023). https://www.chinaiplawupdate.com/2023/12/chongqing-no-1-intermediate-peoples-court-sets-global-frand-rate-for-5g-seps-at-0-707-unit-in-nokia-oppo-case/; Mallinson, *Race to the Bottom*, *supra* note 65.

[99] Erik Stasik & David Cohen, *Royalty Rates and Licensing Strategies for Essential Patents on 5G Telecommunication Standards: What to Expect*, 3 LES NOUVELLES 176 (2020).

on a single device or model, let alone overall for any OEM when blended across all product sales in a nation or accounting period.

Also, according to Stasik's testimony in *Optis v. Apple* citing his same report:

> In 1998, ITSUG (an obscure organisation representing some operators and manufacturers) filed a complaint with the European Commission claiming that "when GSM mobile handsets first appeared on the market place cumulative royalties amounted to as much as 35 per cent to 40 per cent of ex-works selling price."
>
> In 2007, Lemley and Shapiro commentated that they had "seen estimates for [W-CDMA] as high as 30 per cent of the total price of each phone…based on summing royalty demands before any cross-licensing negotiations began."[100]

b. Academics' and Analysts' Published Estimates

Over the decades, academics and various industry and financial analyst firms have come up with widely differing estimated aggregate royalty rates, in some cases including some additional indication of what the figures represent. In addition to the above estimate of academics Lemley and Shapiro, estimates for WCDMA also included 25% to 30% by Dr. Bekkers in 2006,[101] 31.5% by ABI Research in 2008,[102] and 17.5% by ABI Research in 2011.[103] In 2005, investment bank Credit Suisse First Boston provided an estimate for cumulative WCDMA royalties at 17.3%, noting that rates "could be as high as 25-30%."[104] Industry expectation for aggregate royalties on the UMTS standard (which is effectively the same as WCDMA) were also reportedly up to 20% by Dr. Bekkers in 2009.[105]

---

[100] This hypothetical percentage is cited as evidence of alleged "royalty stacking" — based on the Cournot complements theory described *supra* in footnote 33 — with bilateral negotiations between individual SEP owners and implementers supposedly leading to excessive aggregate royalties. *See* Mark Lemley & Carl Shapiro, *Patent Holdup and Royalty Stacking*, 85 TEX. L. REV. 1991 (2007).

[101] Rudi Bekkers, *The Rules, Norms, and Standards on Knowledge Exchange* (DIME, Working Paper No. 9, 2006). https://rbekkers.ieis.tue.nl/Bekkers%20West%20(2006)%20DIME%20IPR%20working%20paper%209%20.pdf.

[102] Stuart Carlaw & Clint Wheelock, *Mobile Device Royalties: Intellectual Property Rates for GSM, WCDMA, and LTE,* ABI RESEARCH (2008) (table 1.2 indicates royalty stacks of 31.5% for 3G likely for licensees without patent strength).

[103] Phil Solis & Stuart Carlaw, *Mobile Device Royalties: GSM, WCDMA, and LTE,* ABI RESEARCH, 31-33 (Dec. 20, 2011) (royalty rate for licensees without patent strength is 17.5% for GSM/WCDMA).

[104] Credit Suisse First Boston, *3G Economics* (Sept. 6, 2005).

[105] Rudi Bekkers & Joel West, *The Limits to IPR Standardization Policies as Evidenced by Strategic Patenting in UMTS,* 33 TELECOMMS. POL'Y 80, 92 (2009), (total royalties of up to 20% for UMTS).

Estimates for 4G LTE have also varied, with rates including 23.6% by ABI Research in 2008[106] and 35.4% by ABI Research in 2011.[107]

c. The Overall Royalty Yield in All Potentially Licensable Sales

My seminal empirical research on aggregate royalty rates in 2015 indicated that the overall aggregate royalty paid as a percentage of total phone sales revenues for all standards and including all cellular handset vendors was no more than around 5%.[108]

This article is where I coined the term "royalty yield" for that kind of aggregate rate.[109] The term was subsequently adopted by others in their published literature where they validated my methodology and derived even lower rates than mine.[110] Such labeling, and that of ARRFA, are required in FRAND licensing royalty assessments to distinguish between the different complexions of aggregate rate with significant differences among them in what various figures presented are actually depicting.

The huge differences between aggregate figures in Section III(E)(2)(a) *The Addition of Every SEP Owner's Maximum Wishes* and Section III(E)(2)(b) *Academics' and Analysts' Published Estimates* — versus Section III(E)(2)(c) *The Overall Royalty Yield in All Potentially Licensable Sales* of only around 5% or even less result from many omissions and reductions. Licensors' aggregate royalty yields — after royalty caps, volume and geographic discounts, discounts to get deals done, discounts on prior sales, cross-licensing, and patents that remain unlicensed — tend to be a lot lower. The headline maximum rates and "program rates"[111] disclosed by many licensors are much higher than the individual royalty yields paid by licensees after all those exclusions and reductions.

For example, royalty caps can result in dramatically lower royalty yields than headline rates. Interdigital's web site rate card indicates a 0.5% headline

---

106 Carlaw & Wheelock, *supra* note 103 (table 1.2 indicates a royalty stack of 23.6% for single-mode LTE is likely for licensees without patent strength).

107 Solis & Carlaw, *supra* note 104 (Table 10 indicates royalty rate for licensees without patent strength is 35.4% for LTE multimode devices).

108 Mallinson, *Cumulative*, *supra* note 4.

109 The royalty yield for a licensee, licensor, or an entire standard is defined as royalties paid by licensee to licensor, divided by corresponding handset revenues. It can be considered the effective royalty rate achieved across all licensed and unlicensed phone sales after all omissions and adjustments including caps, discounts (e.g., for volume and geography) and net of cross-licensing. The sum of yields for all licensors, all licensees, and in a standard, is the same.

110 Galetovic et al., *supra* note 4.

111 Program rates are also referred to as rate card rates. Absent clear or universally accepted definitions, I am distinguishing between undiscounted headline maximum rates indicated on rate cards and the lower rates that are actually applied with any discounts including those due to handset selling price caps that are also made explicit on those rate cards.

maximum royalty rate with a $200 royalty cap on handset price (i.e., $1.00 maximum royalty) for 4G.[112] The corresponding royalty yield on a $1,000 phone is, therefore, only 0.1%.

d. Publicly-Stated Aggregate Royalty Goals by Some Companies

The first collective attempts to agree aggregate rates "enabl[ing] the cumulative royalty rate for W-CDMA to be at a modest single digit level"[113] and for a "single-digit percentage of the sales price"[114] for 4G LTE were around when the standards were first introduced in the early 2000s and late 2000s, respectively. A key objective in setting these single-mode aggregate rate goals was to encourage adoption of these standards in competition to 3G CDMA2000 and 4G WiMAX, respectively.[115] Public announcements in press releases were made by various SEP owners including Alcatel-Lucent, Ericsson, Nokia, and Siemens. All of these also had predominant interests — then, but no longer today — as net payers of royalties on handset sales, as did other OEMs and network operators making these announcements. For example, Nokia's global handset market share was in excess of 40% for much of the 2000s. Nokia and all the other European companies named above had exited the handset market by 2014.[116]

These announcements by only a handful of companies faced plenty of opposition from others. While the former companies have maintained that they were seeking broader support, they have also argued that was not obtained and the goals were not achieved (i.e., aggregate rates paid ended up being higher than goals). Other announcements by some of the same

---

112 *Rate Disclosure,* INTERDIGITAL, https://www.interdigital.com/rate-disclosure (last visited Jan. 20, 2024).

113 Press Release, NTT DoCoMo, Ericsson, Nokia, Siemens, Industry Leaders NTT DoCoMo, Ericsson, Nokia and Siemens, and Japanese Manufacturers Reach a Mutual Understanding to Support Modest Royalty Rates for the W-CDMA Technology Worldwide, (Nov. 6, 2002), https://www.sec.gov/Archives/edgar/data/924613/000110465902006769/j6199_6k.htm.

114 Press Release, Nokia, Wireless Industry Leaders Commit to Framework for LTE Technology IPR Licensing, (April 14, 2008), https://www.sec.gov/Archives/edgar/data/924613/000110465908029241/a08-13064_16k.htm.

115 An additional objective was to reallocate shares of royalties among SEP owners, versus some existing licensing, with "licensing arrangements whereby essential patents for W-CDMA are licensed at rates that are proportional to the number of essential patents owned by each company[,]" Press Release, NTT DoCoMo, Ericsson, Nokia, Siemens, Industry Leaders NTT DoCoMo, Ericsson, Nokia and Siemens, and Japanese Manufacturers Reach a Mutual Understanding to Support Modest Royalty Rates for the W-CDMA Technology Worldwide, (Nov. 6, 2002), and for LTE "according to the licensors' proportional share of all standard essential IPR for the relevant product category[,]" Press Release, Nokia, Wireless Industry Leaders Commit to Framework for LTE Technology IPR Licensing, (April 14, 2008).

116 Keith Mallinson, *How Europe can Build on Strengths in SEPs to Reclaim Leadership in Cellular with 5G and 6G*, 4IP COUNCIL (Apr. 28, 2022), https://www.4ipcouncil.com/features/how-europe-can-build-strengths-seps-reclaim-leadership-cellular-5g-and-6g.

licensors indicated that aggregate figures should not be regarded as royalty caps and licenses should be negotiated bilaterally not simplistically apportioned based on patent counts.[117] "The Minimum Change Optimum Impact (MCOI)" proposal, issued jointly by Ericsson, Motorola, and Nokia in 2006, sought to codify the twin principles of aggregated reasonable terms (ART) and proportionality into the FRAND definition.[118] Two years later, Tim Frain, Nokia's Director of IPR Regulatory Affairs, gave a public address at a European Commission workshop stating that "ART is not any kind of royalty cap. . . . It is no more than an individual patent owner's own understanding or articulation of what a reasonable cumulative royalty would be given all the market conditions. Also, Proportionality is not simply about patent counting . . . Actual royalties remain to be negotiated bilaterally in the normal way." [119]

As these announcements were targets for aggregate rates actually paid, these are also effectively target royalty yields, rather than input figures to be used as ARRFAs, which would necessarily need to be higher figures given that standards are never fully licensed, and some apportionments would not result in any payments.

Such figures have created self-reinforcing "anchoring."[120] Despite all the above, the figures in these announcements are still commonly cited, for example, in FRAND licensing litigation (e.g., *Unwired Planet*), and are proposed as prospective benchmarks for use in making FRAND rate determinations.

e. Other Estimates of Hypothetical and Actual Rates Paid

Cases in litigation include consideration of various estimates for aggregate royalties. Little or no weight is given to the hypothetical maximum aggregate rates in Section III(E)(2)(a) *The Addition of Every SEP Owner's Maximum Wishes* that nobody would actually ever pay because these ignore discounting and unlicensed SEPs. At the other end of the scale, consideration is given to royalty yields derived bottom-up from royalties paid using the top-down approach formula, but these are typically higher than in Section III(E)(2)(c) *The Overall Royalty Yield in All Potentially Licensable Sales*

---

117 Tim Frain, Director, IPR Regul. Affs., Nokia Corp., Address at European Commission Workshop on IPR in ICT Standardisation: FRAND Best Practice, 3 (Nov. 19, 2008); Ericsson, Motorola, and Nokia, Joint Proposal to ETSI, Minimum Change Optimum Impact (MCOI) (2006).

118 WIPO, Tim Frain, *Patents in Standards & Interoperability*, at 7-8 (Nov. 29, 2006) (explaining the MCOI approach laid out in Ericsson, Motorola, and Nokia, Joint Proposal, *supra* note 119.

119 Frain, *supra* note 119, at 3.

120 In their research about the anchoring effect, psychologists Daniel Kahneman and Amos Tversky showed that when we're asked to make a judgment in the face of uncertainty, we are easily swayed by the first figure that's introduced into the conversation, however irrelevant, outrageous, or insulting it may seem. *See generally*, Daniel Kahneman & Amos Tversky, *Judgment Under Uncertainty: Heuristics and Biases*, 185 SCI. 1124 (1974).

because the denominators in those royalty yields focus on sales of phones conforming to specific standards such as 4G or 5G, albeit in multimode devices, and because rates considered are typically higher one-way rates after unpacking cross-licenses.

In *Unwired Planet v. Huawei,* the court derived an "implied total burden" of 8.8% for multimode 4G from the comparable licenses and the share of Unwired Planet's relevant SEPs.[121] The court also noted that "[o]n Huawei's figures the implied total aggregate [4G] royalty burden T would be 13.3% while for Unwired Planet it would be 10.4%."[122] The rates derived from unpacking comparable licenses are based on amounts that would be paid, but for cross-licensing. Consequently, aggregate rates implied from these with use of the top-down formula are theoretical. They are adjusted royalty yields, before cross-licensing reductions and are elevated by including notional royalties (i.e., royalties not paid) for unlicensed SEPs that are counted in the denominator for the derivation of S (licensor's share of SEPs).

In the UK's *Optis v. Apple* FRAND trial, expert witness Eric Stasik, with many years' experience in licensing negotiations was asked by Optis' solicitors to give [his] view as to whether it would be reasonable, assessed as of today, for implementers to be expected to bear a theoretical notional aggregate royalty burden for 4G multimode handsets in the range of around 8% to 15% (i.e., a total royalty burden in respect of all relevant (i.e., handset) SEPs in the 2G, 3G and 4G "universe."

In response he testified:

> [I]n the (hypothetical) scenario where implementers do all behave as willing licensees and all in fact therefore pay truly "FRAND rates" for the whole stack, a range of 8% to 15% is appropriate ["in respect of all relevant (i.e. handset) SEPs in the 2G, 3G and 4G 'universe'"].[123]

Stasik also noted that "[i]n practice, implementers do not pay the theoretical total aggregate royalty burden for a 4GMM handset because implementers in my experience are never fully licensed under all SEPs in the 4G, 3G and 2G universe." His description is therefore, seemingly of more than a royalty yield — by pretending unwilling licensees are willing, licensed, and paying royalties. While I presume cross-licensing did not feature much in that particular case because Optis is not an implementer, it is unclear whether the rate at the lower end of that range is supposed to be net of cross-licensing reductions.

The wide percentage range — with the top figure nearly double the bottom figure — seemingly reflects the variability in amounts paid — largely to major licensors. Major licensees such as OEMs Apple, Samsung, Sony, and Xiaomi with relatively large sales and ability to pay large lump sum fees up

121 Unwired Planet v. Huawei, [2017] EWHC 711, at [478] (Pat).
122 *Id.* at [261].
123 Optis v. Apple, [2023] EWHC 1095, at [400] (Ch).

front might be able to obtain significant further discounts to headline rates offered in rate cards and as are also disclosed on licensors' web sites. In contrast, payments made by small licensees with little or no negotiating power will be much closer to rates indicated initially in rate cards.[124] Various aggregate rate figures have also been presented to government agencies including competition authorities. Where such figures are reported, in some cases confidentially, it is not always clear how terms such as "typical" aggregate rate are defined — if at all — or what exactly they depict.

F. *Comparing and Setting Aggregates*

The aggregate royalty rate selected as the starting point *input* for apportionment among licensors in top-down approach determinations of FRAND royalties for SEPs (i.e., the ARRFA) must reflect the actuality that the *output* aggregate rate paid in cash or in kind by licensees will generally be lower. Some SEP royalty pie is left uneaten when it is shared in top-down approach apportionments.

It would be inapplicable to use the maximum stack of single-mode or multimode program rates in Section III(E)(2)(a) *The Addition of Every SEP Owner's Maximum Wishes* as the ARRFA because the inflated claims of some owners would over-value the entire pie, and in turn, also the apportionments.

However, apportioning only the aggregate royalty rate figures in Section III(E)(2)(c) *The Overall Royalty Yield in All Potentially Licensable Sales* or in Section III(E)(2)(d) *Publicly-Stated Aggregate Royalty Goals by Some Companies* will in turn result in sub-FRAND rate determinations for individual licensors and licensees, and yet lower aggregate royalty rate payments. If this approach caught on, there would be a vicious cycle of rates spiraling lower and lower as sub-FRAND rates are used to set the next aggregate rate for apportionment, and so on ad infinitum. The total of all licensors' R figures would fall short of T. Aggregate royalty yields in Section III(E)(2)(c) *The Overall Royalty Yield in All Potentially Licensable Sales* are inapplicable as the input for apportionment because the top-down approach allocates royalties that generate no royalty payments. Unpaid royalty allocations to unlicensed SEPs and to SEPs that are cross-licensed without payment are not royalty costs. There is no direct or variable cost in cross-licensing to reduce

[124] As I pointed out in my previous feedback to the Commission's proposed legislation, the top-down approach makes no attempt to determine non-discriminatory variations in rates among differently situated licensees. Mallinson June, *supra* note 12. It is beyond the scope of this paper to consider whether or how to adjust aggregate rates for apportionment to deal with this major issue in FRAND licensing.

royalty net payments. The R&D costs in developing patents for cross-licensing are sunk fixed costs.[125]

Similarly, target maximum payments in Section III(E)(2)(d) *Publicly-Stated Aggregate Royalty Goals by Some Companies* also appear to be something like royalty yields — derived from what is actually paid, or would actually be paid in accordance with those announcements — not based on what should be available for payment in the hypothetical and unrealistic circumstance of full licensing. All those paid rates, or to be paid rates, would need to be grossed-up by various factors before being used as the top-down input ARRFA.

Implied total burden figures such as those derived in *Unwired Planet* appear to be more appropriately formulated to be used as ARRFAs because they account for unlicensed SEPs. However, the precision and reliability of such figures is highly questionable — particularly as an ARRFA, rather than as an implied figure for cross checking, as was the sole intention of the judge. The court noted in that case that for 4G from the comparable licenses its "[8.8%] is lower than the aggregate implied by either party's case (Huawei's 13% and Unwired Planet's 10.4%)."[126] Implied aggregate rates are proportionate to rates derived from unpacking and inversely proportional to shares of total SEPs. An aggregate is implied by dividing an SEP owner's unpacked rate by its respective estimated share of all SEPs in the applicable standard.

The cost to the licensee is what it actually pays, not what it avoids paying when it should pay, or the discount it receives for geography or patent expirations, or for any other notional charges that it has not and will not be asked for. Unpaid liabilities might eventually be paid, but back royalties are often only paid as deeply discounted release payments when new licenses are negotiated and agreed.

While the formulation in Section III(E)(2)(a) *The Addition of Every SEP Owner's Maximum Wishes* depicts rates that are too high, even as the starting point input for apportionment, let alone an indication of what one would have to pay, the royalty yield formulations in Section III(E)(2)(c) *The Overall Royalty Yield in All Potentially Licensable Sales* and Section III(E)(2)(d) *Publicly-Stated Aggregate Royalty Goals by Some Companies* indicate rates that are too low to be the ARRFA. In between, such formulations and figures, with suitable adjustments, and some formulations in Section III(E)(2)(b) *Academics' and Analysts' Published Estimates* and Section III(E)(2)(e) *Other Estimates of Hypothetical and Actual Rates Paid*, might well be suitable for that purpose, subject to applicability of the timing and verified accuracy of such estimates.

While the following pie chart in Exhibit 2 is not to scale it is intended to include everything that might be depicted in various aggregate rate figures.

---

125 I agree with Alexander Galetovic, Stephen H. Haber, and Lew Zaretski about how to deal with cross-licensing in deriving aggregate royalty costs. Mallinson, *Cumulative*, *supra* note 4; Galetovic et al., *supra* note 4; Sidak, *supra* note 4.

126 *Unwired Planet*, [2017] EWHC 711, at [476].

Some slices might be very small or non-existent under certain circumstances. It also shows how pieces of aggregate royalty pie will be left uneaten (e.g., unlicensed SEPs). A proportion of the value ascribed from any aggregate rate figure other than the royalty yield is not paid for in cash. Instead, some payments are made in kind, as in aforementioned cross-licensing. Whether these should be regarded as royalty charges — from an economic, management accounting, or financial accounting point of view — depends on what is provided in kind and how that is costed. For example, product supply in kind is likely to require significant variable cost.

Exhibit 2: Aggregate pie gets left on the table in top-down apportionments among licensors (not to scale)

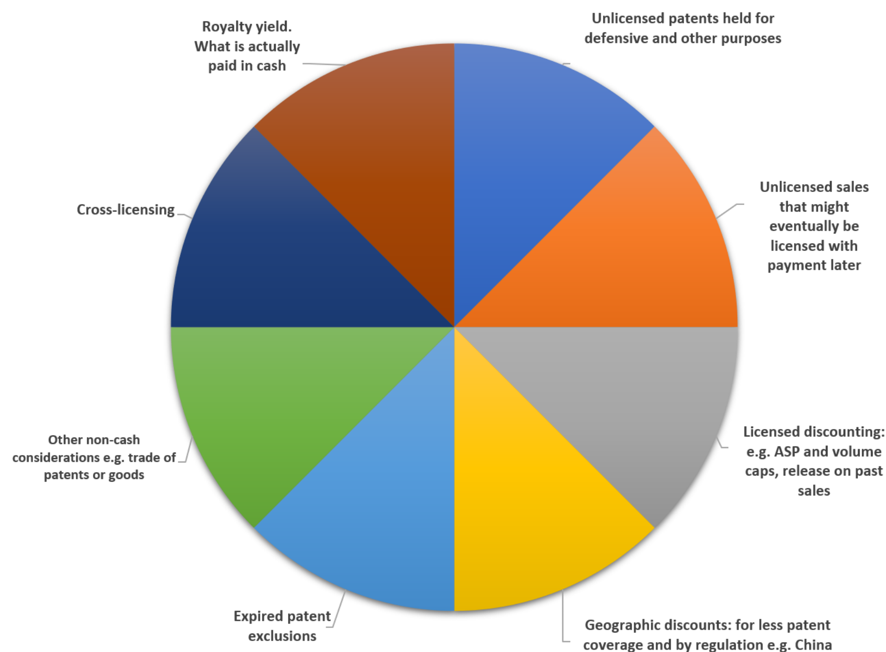


Exhibit 2 aims to include all hypothetically possible charges, including the maximum rates for all SEP owners, as do some of the highest among aggregate rates presented at around 30% for 4G LTE in Section III(E)(2)(a) *The Addition of Every SEP Owner's Maximum Wishes*. However, only the royalty yield slice is actually monetized in cash payments to licensors. It corresponds to the lowest among aggregate rates, such as only around 5% or even less including all standards, as indicated in Section III(E)(2)(c) *The Overall Royalty Yield in All Potentially Licensable Sales*.

G. *How Much More Than Royalties Paid Would Aggregates be with Full Licensing?*

The "fully licensed" aggregate rate is the applicable ARRFA.[127] The aggregate royalty allocated needs to include all the SEPs counted in the denominator of the apportionment calculation of S as if all SEPs are fully licensed for FRAND royalty payments. In contrast, in the special case of patent pools, there are no allocations for patents outside the pool because patent pools do not count SEPs that are outside the pool, even though some of them might be licensed bilaterally or by another pool. In the top-down approach, the count of all SEPs in a standard are included in the denominator calculating S whether or not they are licensed. Consequently, the ARRFA must be increased above the aggregate royalty yield figures, as if those additional SEPs are under license and paid for at FRAND rates.

The same goes for geographic reductions. If the overall royalty rates being determined are attenuated due to geographies where patent protection is relatively weak, as I illustrate above with the example in the *TCL v. Ericsson* decision, then the ARRFA needs to be increased correspondingly. Such increases will be taken back out to the extent applicable on case-by-case basis in specific FRAND determinations. In practice, for example, handset OEMs almost invariably sell in in multiple jurisdictions, with higher rates paid in some than others, and so overall royalty rates paid will average out.

There also needs to be an upward adjustment if expired patents are excluded from the numerator while being retained in the denominator in calculating the rate of apportionment S. Alternatively, as patents expire, they should be removed from both the numerator and denominator, as they typically are by patent pools. Similarly, new SEPs should be added to both numerator and denominator. Fully licensed royalties should be derived entirely from the non-expired patents in the standard, as numbers of these fluctuate.

It is also necessary to gross-up for cross-licensing. Imagine a world where the aggregate royalty yield was zero due to completely balanced cross-licensing. While net royalty rates are zero there, one-way rates could still be substantial. Top-down apportionments derive one-way rates. These can then be netted off to determine how much should be paid in cash and to whom.

However, there should be no upward adjustment for licensors' discounting against their maximum headline rates or for rates agreed below the indicated discounts offered in rate cards. This is on the assumption that their SEPs are being fully monetized by receiving FRAND royalties overall at the discounted rates they have bilaterally agreed through negotiation and that they receive in payments. There should be upward adjustments for notional charges that are unilaterally not sought (i.e., no licensing offered) or not paid (i.e., unlicensed hold-out by unwilling licensees).

[127] "Fully licensed" is a term that was used with this meaning by Eric Stasik in the *Otis v. Apple* decision. *See* Otis v. Apple, [2023] EWHC 1095 (Ch).

CONCLUSION

The US and Europe are heading in different directions on how to determine FRAND charges and other licensing terms for SEPs. While the US has shunned rate-setting regulation by withdrawing guidance from government agencies including the USPTO, NIST and DoJ and is diminishing proposed lawmaking, the European Commission's advocated legislation requires mandatory — albeit non-binding — patent registration, essentiality checking, aggregate royalty setting and rate apportionment among licensors.

There is no evidence of market failure in market-based pricing of SEP royalty rates. To the contrary, established licensing incentivizes innovation and has brought success throughout the ecosystem including implementers and consumers.[128] Disrupting this would harm US and European licensors including Qualcomm, Interdigital, Ericsson, and Nokia among others. The result would be a massive transfer of wealth, principally to Asian implementers and would be a substantial setback for future innovation including upcoming standards such as 6G in the emerging IoT.

Setting aggregate rates and apportioning them among patent owners, centrally by the EUIPO or its subcontracted conciliators — even on a non-binding basis — will unnecessarily distort the free market processes in standards development and FRAND patent licensing compensation. This has been effective in enabling the world's fastest growing and largest ever technology ecosystem serving more than five billion people and 16 billion connections with cellular worldwide. Parties in licensing disputes will feel obliged in the proposed mandatory — but notionally non-binding — conciliation process to give significant weight to the EUIPO's determinations, as will the courts. However, there is no basis whatsoever, let alone supporting evidence, to infer there is harm to be fixed, or that established benchmarks for royalty charges need to be replaced.

Limited checking to ensure that licensors have at least some SEPs would show that they can legitimately demand licensing and royalties. Many patent owners are already able to do this with their proud lists of patents that have been scrutinized by experts and, in some cases, verified by the courts. The proposed processes at the EUIPO, including submission and checking of patents and some claim charts, as well as conciliators setting royalty rates, is fraught with all kinds of issues that will lend to manipulation, favoritism, or bias and also subject checks or patents to subsequent legal challenges. SEP owners have shunned voluntary essentiality checking by an official body in Japan. There is no evidence that these European proposals will be any more welcome or widely adopted.

---

128 *See also* Mallinson, *Don't Fix What Isn't Broken*, *supra* note 6.

Until 2014 we were still being told by some that aggregate royalty rates paid on smartphones could be as much as 30%.[129] In 2015 I showed that rates paid were only around 5%.[130] While both percentages are aggregate rates, they are depicting very different phenomena. An appropriate percentage to be used as the ARRFA in FRAND rate determinations for smartphone licensing will surely fall well between those two extremes and will be higher than any of the royalty yield figures derived. The recent *Optis v. Apple* decision included expert testimony that an aggregate rate range from 8% to 15% would be applicable for multimode 4G, while also indicating that those rates are what would be paid if SEPs were, hypothetically, fully licensed, which is never the case in practice.

Parties in negotiation may agree to use whatever methods they wish to value patents and determine royalties, and courts also decide what to use case-by-case in litigation where they have often rejected top-down rate setting. Rote, formulaic methods for setting and allocating royalties by a central government bureaucracy are unnecessary and will harm a vibrant and well-functioning ecosystem in standards-based technology innovation and development. Better to obtain and reveal more information about existing licensing charges and other terms in many existing licenses than to make up alternatives.

ARRFA figures need to be net of licensors' rate reductions, such as royalty base price caps and other discounts agreed bilaterally between licensor and licensee. However, figures such as royalty yields should be grossed-up for what is unilaterally missing from aggregate payments received from all licensors. These unpaid royalties are due to SEPs being unlicensed, for example, where licenses are not offered and the SEPs are held only for defensive purposes, and where implementers are unwilling licensees and are not paying. Upward adjustments to royalty yield figures are also needed to adjust for the effects of cross-licensing in existing licenses.

We are still in the process of properly identifying and describing all the factors that should be incorporated or excluded in setting aggregate rates for apportionment, and building rigorously-reasoned consensus on what the figures should be with coherent methods for their apportionment.

If we are going to do top-down apportionment properly and with precision, we must develop well-defined ARRFAs, as distinct from and among other kinds of aggregate rates. For example, some will need to be fixed monetary figures per unit rather than percentages, depending on application (e.g., fixed monetary figures in IoT). This article contributes to the ongoing debate about the need for such figures, what exactly they should include and exclude, how to apportion them, if at all, where to find the benchmark royalty

---

[129] *See, e.g.*, Ann Armstrong, Joseph J. Mueller, & Timothy D. Syrett, *The Smartphone Royalty Stack: Surveying Royalty Demands for the Components Within Modern Smartphones* (June 1, 2014), https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2443848.

[130] Mallinson, *Cumulative*, *supra* note 4.

data, and what other valuation methods can be used in determining those rates.

NEW MONEY, OLD STATUTES: INFLATION AND STATUTORY DRIFT

Patrick Sullivan[1]

"The value of money may not only alter but the State of Society may alter. In this event the same quantity of [goods], the same value would not be the same compensation . . . [Amounts] must always be regulated by the manners & the style of living in a Country"

– Governour Morris to the Constitutional Convention[2]

I. INTRODUCTION

The United States Code contains uncounted monetary values, from the penalties that give the law its force to the thresholds that define its very limits. When stated nominally these provisions erode as the value of money changes. As they erode their impacts drift away from what legislators and citizens expect, sometimes in serious ways such as adding years to criminal sentences or shifting tax burdens.

This inflationary drift[3] presents a separation of powers quandary. Most scholars agree that only Congress can update such unambiguous statutory provisions.[4] But Congress has often failed to address even large inflationary drifts. The amount in controversy floor for diversity jurisdiction was last set

---

[1] J.D., Harvard Law School. An earlier version of this article won Harvard Law School's Irving Oberman Memorial Prize. I appreciate the prize fund's assistance with publication costs. Many thanks to Jacob Gersen, Rosa Vargas, and Amanda Madigan for their invaluable insights and support during the writing process.

[2] The Records of the Federal Convention of 1787, at 45 (Max Farrand ed., 1911) (opposing indexation of judicial salaries to the price of grain).

[3] Inflationary drift is used throughout as shorthand for shifts in expected impacts of statutes driven by the changing value of money, both in inflationary and deflationary directions.

[4] *See* Jim Chen, *The Price of Macroeconomic Imprecision*, 54 HASTINGS L.J. 1375, 1378 (2003) ("Whatever power courts may have in other settings to forestall statutory obsolescence through dynamic interpretation, judges are mostly impotent to adjust numbers or quantitative formulas engraved directly into a statute") (citations omitted); Note, *Desuetude*, 119 HARV. L. REV. 2209, 2220 (2006) (identifying longstanding and broad consensus that legislative monopoly on lawmaking prevents other entities from revising obsolete statutes); *see also* JOHN F. MANNING & MATTHEW C. STEPHENSON, LEGISLATION & REGULATION 79 (2017) (describing widespread acceptance of the primacy of clear text); *cf.* William Eskridge, Jr., *Dynamic Statutory Interpretation*, 135 UNIV. PA. L. REV. 1479, 1494–95 (1987) (arguing against "conventional" view that clear text forecloses room for an interpretation more consistent with current context).

at $75,000 in 1996.[5] It has declined 40% in real terms since.[6] From 1864 until 1988 the fee payable to an attorney in a veteran's benefits dispute was capped at $10,[7] by the time it was finally changed that cap was 13% of its original value.[8] The threshold for construction contracts covered by the Davis Bacon Act's wage regulations was set at $2,000 in 1931 and as of 2022 had yet to be updated.[9] The same threshold would be $38,000 today.[10] Legislative failure coupled with administrative and judicial impotence has led to many such cases of statutory disrepair.

The problem of obsolete statutory text has inspired several eminent critiques of strictly separated lawmaking powers.[11] Inflationary drift is a less-discussed subcategory of that problem. Further exploration is warranted because of its novel aspects. First, inflationary drift lacks the interpretive safety valve of other forms of obsolescence. Money values are too clear and often too central to be reinterpreted by courts and agencies.[12] Second, the effects of statutory aging are measurable—via long-spanning price indices—to a degree unheard of with other forms of obsolescence.[13] Third, the pace of obsolescence has shifted over time. The 19th century saw short term inflation and

---

5 28 U.S.C. § 1332(a), *last amended by* Federal Courts Improvements Act of 1996, Pub. L. 104-317 § 204, 110 Stat. 3847, 3850 (1996).

6 Steven Gensler & Roger Michalski, *The Million Dollar Diversity Docket*, 47 BYU L. REV. 1653, 1714 (2022).

7 Charles L. Craigin, *The Impact of Judicial Review on the Department of Veterans Affairs' Claims Adjudication Process: The Changing Role of the Board of Veterans' Appeals*, 46 MAINE L. REV. 23, 26 (1994).

8 Calculation based on FED. RES. BANK OF MINN., *Consumer Price Index 1800-* [hereinafter *Historical CPI-U data*].

9 Notice of Proposed Rulemaking: Updating the Davis-Bacon and Related Acts Regulations, 87 FED. REG. 15698, 15700 (proposed Mar. 18, 2022).

10 Calculation based on *Historical CPI-U data*, *supra* note 8.

11 *See, e.g.*, Henry J. Friendly, *The Gap in Lawmaking—Judges Who Can't and Legislators Who Won't*, 63 COLUM. L. REV. 787, 797 (1963) (discussing as an example the obsolescence problems created by nationwide circulation of media for defamation suits); GUIDO CALABRESI, A COMMON LAW FOR THE AGE OF STATUTES *passim* (1982).

12 Compare Richard A. Merrill, *FDA's Implementation of the Delaney Clause: Repudiation of Congressional Choice or Reasoned Adaptation to Scientific Progress?*, 5 YALE J. REG. 1, 21–41 (1988) (describing FDA's interpretive "escape" from strict text of non-monetary Delaney Clause), *and* Cass R. Sunstein, *Interpreting Statutes in the Regulatory State*, 103 HARV. L. REV. 405, 493–96 (1989) (describing common interpretive responses to obsolete non-monetary text), *with* Randall v. Sorrell, 548 U.S. 230, 262 (2006) (finding judicial adjustment of political donation limits to account for inflation outside of the judiciary's interpretive authority).

13 *Cf.* PATRICK HANKS, LEXICAL ANALYSIS: NORMS AND EXPLOITATIONS 145 (2013) (describing difficulties in measuring shifts in linguistic meaning). Even our most eloquent jurists have struggled to precisely express the magnitude of the drift of statutes, often resorting to intuition and example. *See, e.g.*, Benjamin Cardozo, *A Ministry of Justice*, 35 HARV. L. REV. 113, 117 (1921) ("I have spoken in generalities, but instances will leap to view. There are fields, known to us all, where the workers in the law are hampered by rules that are outworn and unjust.").

deflation but relative stability in the long-term value of the dollar.[14] Since roughly 1930 long-run inflation has been the norm, with decade-over-decade increases in the value of money.[15]

These characteristics provide a valuable historical experiment. The clarity of money values controls for the confounding potential of updating disguised as statutory interpretation—this allows for observation of how a strictly formal separation of powers approach functions in practice. Price indices help measure the magnitude of obsolescence. And the change from a long-term stable to long-term dynamic monetary environment presents an exogenous shock to the legal process. Collectively, they provide insight into how well the original conception of separation of powers has adapted to a more dynamic context. They help answer the pressing question: can just one helmsman keep the ship of state from drifting amidst increasingly strong socioeconomic tides?

This article evaluates the results of that historical experiment. It argues that traditionally strict separation of powers has not adapted well to the changing monetary environment. It does so by describing the evolution of the monetary context from 1789 to today and evaluating five efforts by Congress and its agents to address the statutory drift caused by that shift. It demonstrates that a highly formalist approach has proved ill-suited to the mix of technical and normative updating problems inflationary drift presents.

The article proceeds in three parts. Part II describes the problem of inflationary drift and why separation of powers doctrine hinders efforts to address it. It supplements and synthesizes the work of prior scholars who described particular aspects of the problem but did not address the broad and evolving scope of this statutory pathology.[16] Part III interprets the results of the historical experiment created by the changed monetary context. It describes the major shift from long run price stability to long run inflation that occurred beginning in the 1930s and then analyzes five efforts to address the consequences of that shift for nominally worded statutes. These efforts began in the 1970s as the reality of the new monetary environment sank in and continue into the current decade. Part IV concludes by arguing that a strict separation of powers approach is outdated and ill-suited to a society of rapid and lasting change. A more collaborative approach to lawmaking is needed.

---

14 Section III, *infra*.

15 *Id.*

16 *See, e.g.*, KEITH S. ROSENN, LAW AND INFLATION (1982) (discussing inflationary erosion across private law issues); Chen, *supra* note 4, at 1384–1402 (reviewing several areas where inflation impacts public law as prelude to discussing tradeoffs in inflation indexes); KENT R. WEAVER, AUTOMATIC GOVERNMENT: THE POLITICS OF INDEXATION (2010) (analyzing indexation of mostly entitlement programs from a political science perspective).

II. INFLATIONARY DRIFT CAUSES WIDESPREAD HARMFUL EFFECTS WHICH HYPER FORMAL SEPARATION OF POWERS STRUGGLES TO ADDRESS

Congress enacts three types of monetary values—penalties, payouts, and thresholds. Penalties include criminal fines and civil penalties. Payouts are the amounts received under various welfare or stimulus programs such as Social Security or corporate tax credits.[17] They also include statutorily-defined transaction values such as public employee salaries and caps on federal contracts. Thresholds include provisions that define the boundaries of the law like the amount in controversy precondition for diversity jurisdiction. They also include provisions which trigger different treatment like income tax brackets. Unlike penalties and payouts, thresholds do not always have direct fiscal impacts.

When stated nominally all these are affected by monetary fluctuations. The quality of the change depends on the structure of the provision and interests of the affected parties. For example, an increase in nominal wages driven by broad-based inflation can move a taxpayer into a higher tax bracket even though their real wealth has not changed.[18] At the same time it makes the nominal penalty for a willful failure to pay tax, set at $10,000,[19] less burdensome. Deflation would work in opposite directions. Impacts also depend on the duration of the inflation. Sustained year over year price shifts lead to the most dramatic shifts as the gap between the nominal amount in the statute and real values is compounded.[20] Short run fluctuations can also change how the law is experienced but in a less universal way. A civil penalty assessed during a six-month period of deflation will feel more burdensome than one assessed when the value of a dollar is worth less, but only citizens who pay during the deflationary period will experience this deviation. Long running appreciation in the value of the dollar would subject more people to the more punitive experience.

The consequences of these drifts profoundly affect legal outcomes. Prior to income tax bracket indexation, Milton Friedman estimated that every 10% increase in prices led to a 15% increase in personal tax rates as inflation moved taxpayers into higher brackets.[21] According to one analysis the failure to index the current child tax credit to inflation will leave roughly 10% more

---

[17] Some payouts are not stated as an explicit dollar value but are based on a percentage of historical income. These also drift because the base is nominally stated.

[18] Reed Shuldiner, *Indexing the Tax Code*, 48 TAX L. REV. 537, 541–42 (1993).

[19] 26 U.S.C. § 7202.

[20] *See* Alan Reynolds, *The Mystifying Arithmetic of Year-to-Year Inflation Estimates*, CATO AT LIBERTY (July 29, 2021).

[21] Milton Friedman, *Inflation, Taxation, Indexation*, *in* INFLATION: CAUSES, CONSEQUENCES, CURES 14 IEA READINGS 71 (1974).

children in poverty in 2032.[22] The inflation-driven depreciation of offense severity thresholds in federal sentencing guidelines contributed to thousands of years of additional prison time.[23] The income threshold to be considered an accredited investor has not been updated since 1982, bringing millions more households into less regulated private securities markets.[24]

This Part argues that these shifts are harmful because they frustrate citizen expectations and legislative plans. It also describes how separation of powers formalism limits efforts to address these harms.

A. *Drift causes laws to deviate from citizens' and lawmakers' expectations*

Inflationary drift is primarily a problem of divergence from expectations. Economic theory suggests that the bulk of inflation's welfare costs arise when citizens and policymakers cannot plan for changing prices.[25] In the legal context, two types of analogous planning errors arise. One can misestimate the rate of inflation. One can also mistime the occurrence and duration of inflation. Both are harmful. A citizen whose nominal income grew faster than expected might find herself unprepared for her higher tax bracket. A legislator proposing an unindexed tax credit to save money will see her expectations frustrated if a deflationary period arrives earlier than expected. Because inflation is difficult to predict (in rate and duration) both errors are probably common.[26]

In fact, instances of drift regularly cause expectation mismatches. Citizens are often surprised and burdened by shifts in the real values of their obligations to and payouts from the government.[27] Public officials regularly

---

22 Sophie Collyer, Christopher Wimer & David Harris, *Keeping Up with Inflation: How Policy Indexation Can Enhance Poverty Reduction*, THE CENTURY FOUND. (Aug. 25, 2022), https://tcf.org/content/report/keeping-up-with-inflation-how-policy-indexation-can-enhance-poverty-reduction/.

23 *See* Sentencing Guidelines for United States Courts, 80 Fed. Reg. 25782-01, 25789–90 (May 5, 2015) (discussing how updating thresholds for inflation would free up 956 prison beds per year by the fifth year of implementation).

24 *See* Michael L. Monson, *The Evolution and Future of the Accredited Investor Standard for Individuals*, 23-Dec UTAH B.J. 36, 37 (2010).

25 *See* N. GREGORY MANKIW, MACROECONOMICS 119 (8th ed. 2006).

26 *See* Tim Sablik, *Forecasting Inflation: For policymakers and market participants inflation can be challenging to predict*, FED. RES. BANK OF RICHMOND ECON FOCUS (2021); Jeanna Simalek, *Inflation Forecasts Were Wrong Last Year. Should We Believe Them Now?*, N.Y. TIMES (Dec. 12, 2022).

27 *See, e.g.*, Kate Dore, *How soaring inflation may deliver a higher tax bill — especially for retirees, homeowners and high earners*, CNBC (July 18, 2022) (discussing "surprise" tax burdens), https://www.cnbc.com/2022/07/18/how-soaring-inflation-may-deliver-a-higher-tax-bill-for-retirees.html; Martha C. White, *Higher food costs make the math even harder for Americans on food stamps*, NBC NEWS (Nov. 23, 2021), https://www.nbcnews.com/business/consumer/higher-food-costs-make-math-even-harder-americans-food-stamps-rcna6446; Gabriella Cruz-Martinez, *Child Tax Credit: Parents miss the money for their children as inflation rises*, YAHOO MONEY (July 5, 2022); Jason Delisle, *What*

grapple with unforeseen impacts on their programs. Recent inflation has upended many established federal contracts.[28] The SEC has acknowledged that the inflation-driven expansion of the definition of accredited investor was an unanticipated departure from the limited scope expected in 1982.[29] Members of the U.S. Sentencing Commission acknowledged in 2015 that the drift of offense severity thresholds was unplanned.[30] Congress can also be caught off guard. Efforts to update inflation-eroded provisions have been justified as correcting unpredicted deviations from the plan of the enacting congress.[31]

Even predictable drift may have subtle costs. Affected parties will incur mental costs from regularly recalculating the real values of legal thresholds.[32] To the extent the effects of drift are unevenly distributed they may result in social division.[33] And the need to regularly update laws can undermine public confidence and increase opportunities for pork legislation.[34]

These costs add up. Some instances of drift have grave individual consequences, such as longer sentences.[35] And instances that seem merely inconvenient individually can have concerning aggregate effects. While empirical evidence is scarce in the legal context, economic work indirectly illuminates the seriousness of the problem. Economic models predict that rapid unexpected price movements, such as the fluctuations experienced during the Great Depression and post WWII, can cause nontrivial drops in GDP.[36] And even more predictable price movements can harm those who have not fully planned for them.[37] Menu costs—the costs to firms of changing posted

---

Better Data Reveal about Pell Grants and College Prices, URBAN INST. (Aug. 18, 2021) (describing how Pell grants have imperfectly kept pace with inflation of education costs).

28 *See, e.g.*, DEPT. OF DEFENSE, Memorandum on Managing the Effects of Inflation with Existing Contracts (Sept. 9, 2022) (describing impacts on fixed-price contracts), https://www.acq.osd.mil/dpap/policy/policyvault/USA001773-22-DPC.pdf; NAID, Can an institution apply an inflation rate to its budget on competing grant applications?, NAID FUNDING NEWS (Jun. 16, 2021) (describing prohibition on automatic inflation adjustments in NAID grants).

29 *See* SEC. & EXCH. COMM'N, Revisions of Ltd. Offering Exemptions in Regul. D, 72 Fed. Reg. 45117 n.53, 45119 (Aug. 3, 2007) (to be codified at 7 C.F.R. pts. 200, 230, 239).

30 Section 0, *infra.*

31 *See, e.g.*, 136 CONG REC. 1493–95 (daily ed. Feb. 12, 1990). (remarks of Sen. Lautenberg on OSHA penalties) ("If it were presented for a vote, would the Senate approve a two-thirds cut in OSHA penalties, when workplace hazards persist? . . . The answer, I think, is no. Yet inaction gives us the same result.").

32 Ruchir Agarwal & Miles Kimball, *How Costly is Inflation?*, INT'L MONETARY FUND: FINANCE & DEVELOPMENT (Apr. 7, 2022), https://www.imf.org/en/Publications/fandd/issues/2022/03/Future-of-inflation-partII-Agarwal-kimball.

33 *Cf.* Mankiw, *supra* note 25, at 119.

34 *See* Section III.C.1, *infra*.

35 *See* Section 0, *infra*.

36 Miquel Faig & Zhe Li, *The Welfare Costs of Unexpected Inflation*, 56 J. MONETARY ECON. 1004, 1012 (2009). The authors also find that more stable monetary cycles have smaller welfare effects, but that even these periods can erode cash balances in a significantly harmful way. *Id.*

37 *Id.* (discussing costs of even regular inflation to individuals who, perhaps irrationally, hold significant cash balances).

prices—are a useful private sector analog to drift. The available data shows that firms expend significant resources to avoid incurring these costs.[38]

The history of political reactions to erosive inflation also hints at the scale of the problem. Populist unrest in the 1880s and 1890s was largely driven by the uneven impacts of deflation on the agricultural class's debts and taxes.[39] Major inflationary and deflationary periods in Chile, Russia, and Germany famously toppled governments and have been sources of instability elsewhere.[40]

Of course, some instances of inflationary drift are intentional. Forgoing tax credit indexation is often a strategy to reduce a bill's fiscal footprint.[41] And Judge Calabresi has posited that a failure to index could represent an indirect effort at sunset legislation or transition smoothing.[42] He also suggests the lack of indexation could reflect some sort of compromise-enhancing ambiguity, though he acknowledges the difficulty in identifying confirmed examples of this behavior.[43]

But these theories of drift by design are either implausible or only explain a limited subset of the instances of inflationary drift. The transition smoothing theory is implausible because of the difficulty in forecasting the direction and magnitude of price changes. Suppose Congress wanted more individuals to be defined as accredited investors over time and thus have access to private securities offerings, but it wanted the adjustment to occur gradually to allow the SEC time to implement necessary safeguards. It could accomplish this by allowing the threshold income qualifications to be nominally stated or it could spell out a schedule of gradual real decreases. The costs of either drafting method are similar, but their likely effectiveness diverges dramatically. Using drift by design assumes nominal incomes will rise predictably. That is an extremely risky bet on the quality of inflation

---

[38] Daniel Levy et al., *The Magnitude of Menu Costs: Direct Evidence from Large U. S. Supermarket Chains*, 112 Q. J. ECON. 791, 815–18 (1997).

[39] Katherine Unterman, *1896: A Populist Insurgency in America's First Gilded Age*, 34 S. CENT. REV. 26, 27 (2017).

[40] *See* Dave Blanchard & Kenny Malone, *When Bricks Were Rubles*, NPR: PLANET MONEY (Apr. 1, 2022), https://www.npr.org/2022/04/01/1090312774/when-bricks-were-rubles; c*f.* Harvey D. Palmer & Guy D. Whitten, *The Electoral Impact of Unexpected Inflation and Economic Growth*, 29 BRITISH J. POL. SCI. 623, 631–36 (1999) (describing significant connection between failure to keep inflation within expectations and decline in incumbent votes across elections in over 100 countries); Lewis E. Hill et al, *Inflation and the Destruction of Democracy: The Case of the Weimar Republic*, 11 J. ECON. ISSUES 299 (1977); Israel Shenker, *Power Eluded Allende, Then Slipped From His Grasp,* N.Y. TIMES, Sept. 12, 1973, at 16.

[41] Alexis Leondis, *Why Are Only Some Tax Breaks Adjusted for Inflation?*, WASH. POST (Aug. 16, 2022); MARGOT L. CRANDALL-HOLLICK, CONG. RSCH. SERV., R45124, THE CHILD TAX CREDIT: LEGISLATIVE HISTORY (2021).

[42] CALABRESI, *supra* note 11, at 66.

[43] *Id.* at 67 n.25.

forecasting, which suffers from a number of well-known limitations.[44] If inflation moves unexpectedly quickly the SEC might have too little time implement safeguards. If deflation occurs the exact opposite effect than intended—fewer people accessing private markets—will occur. Given the salience of inflation as a political issue, lawmakers are surely aware of these prediction difficulties.[45] And legislators do not appear to have a meaningful advantage in forecasting economic trends.[46] It seems far easier for Congress to simply spell out the path it wishes the threshold to take, avoiding the possibility that prices could move faster or in the opposite direction than anticipated. Why rely on such a risky drafting strategy when a low-cost alternative is available?

The theory that the ambiguity of nominally drafted statutes is somehow compromise-enhancing has not been empirically confirmed. Moreover, the same forecasting uncertainty that undermines the transition smoothing theory makes compromise hard to explain. Perhaps lawmakers could have divergent expectations for inflation and thus drafting with nominal values would allow for votes consistent with both sets of expectations. For example, a nominally drafted threshold for accredited investors would attract the votes of representatives who wanted the standard to relax over time and expected inflation and votes from those who wanted the standard to be more stringent and expected deflation. But the idea that a bill's drafter would know these expectations and respond to them tactically seems farfetched. It also seems unlikely that lawmakers possess sufficient confidence in their own predictions of inflation to make them decisive factors in voting decisions over more traditional drivers like currying favor with key constituents.

A more plausible motivation for nominal drafting is securing fiscal savings. Many payouts have been restated in nominal terms to reduce their fiscal footprint.[47] In these cases CBO scoring rules may artificially reduce the forecasting problems described above. CBO reports provide point estimates of the revenue and expenditure effects of a bill.[48] These estimates incorporate

---

44 *See* Julie Bennett & Michael T. Owyang, *On the Relative Performance of Inflation Forecasts*, 104 FED. RES. BANK ST. LOUIS REV. 131, 132 (2022) (finding a tendency to overestimate inflation).

45 *Cf., e.g.*, Victoria Guida & Kate Davidson, *'Deeply troubled': Lawmakers Challenge Fed's Inflation War*, POLITICO (Nov. 7, 2022) (describing examples of political rhetoric around inflation); REPUBLICAN POLICY COMM., *Rising Prices Hit American Families* (July 13, 2021).

46 *See* William Belmont et. al., *Do senators and house members beat the stock market? Evidence from the STOCK Act*, 207 J. PUB. ECON. 104602, 104607 (2022) (finding public equities held by members of Congress do not outperform the market).

47 *See* Leondis, *supra* note 41.

48 CONG. BUDGET OFF., HOW CBO PREPARES COST ESTIMATES, at *8 (2018).

specified inflation forecasts,[49] which appear to be relied on by legislators.[50] Within this artificial information environment, non-indexation becomes a less risky drafting strategy.

However, that artificial certainty is only available when drift is used as a tool to secure fiscal impacts. The CBO does not publish regular estimates of how nominally stated thresholds might shift with inflation.[51] For provisions where inflationary drift would have non-fiscal impacts no additional certainty is created. Moreover, fiscally-motivated use of nominal values only explains legislative intent. To the extent that the relied-on prediction of inflation turns out to be incorrect, the provision will still create expectation mismatches. Thus legislators still must contend with the risk that the actual impact of the nominally drafted bill will diverge from expectations.

In sum, inflationary drift of statutes causes all sorts of laws to diverge harmfully from expectations. And even predictable drifts create meaningful costs, such as undermining confidence in government's effectiveness. These deviations seem unintentional in the majority of cases.

B. *Drift is under-addressed due to separation of powers formalism*

Inflationary drift is significantly under-addressed. Despite some recent efforts, many instances of historical drift have yet to be corrected and important statutory provisions remain vulnerable to erosion.[52]

Most striking is the lack of coherence as to which provisions are resiliently drafted and which remain subject to drift. Difficult to explain variations exist in many policy areas. The definition of accredited investor under the Securities Exchange Act is not indexed, instead updates are left to the

---

[49] Nathaniel Frentz et. al., *A Simplified Model of How Macroeconomic Changes Affect the Federal Budget*, CBO Working Paper 2020-01 at *32 (2020) (describing statutorily required inflation measure for discretionary spending); *id.* at *10 (describing role of CBO price forecast in estimating tax credit impacts); *id.* at *24 (explaining reliance on CPI-W in scoring benefit programs).

[50] *See, e.g.*, Philip Rocco, *Congress is waiting on the CBO for its Build Back Better report – but how did fiscal scorekeepers come to be so powerful in politics?*, THE CONVERSATION (Nov. 16, 2021), https://theconversation.com/congress-is-waiting-on-the-cbo-for-its-build-back-better-report-but-how-did-fiscal-scorekeepers-come-to-be-so-powerful-in-politics-171642; Jason Dick, *CBO Score Will Ring in Another Round of House Fight*, ROLL CALL (Mar. 13, 2017), https://rollcall.com/2017/03/13/cbo-score-will-ring-in-another-round-of-house-fight/.

[51] *See* CONG. BUDGET OFF., HOW CBO PREPARES COST ESTIMATES, at *8 (2018); CONG. BUDGET OFF., ESTIMATING THE COST OF ONE-SIDED BETS: HOW CBO ANALYZES THE EFFECTS OF SPENDING TRIGGERS, at *3 (2020) (explaining that for fiscal estimates contingent on certain one-sided thresholds, such as a nominal price, being passed the CBO publishes averages of fiscal impacts based on probability distributions of the relevant threshold but does not regularly publish point estimates of that threshold).

[52] *See* KENT R. WEAVER, AUTOMATIC GOVERNMENT: THE POLITICS OF INDEXATION 240–41 (2010); *cf.* Suzanne Mettler, *The Policyscape and the Challenges of Contemporary Politics to Policy Maintenance*, 14 PERSP. ON POL. 369, 379–82 (2016) (describing lower frequency of efforts to revisit old statutes).

discretion of the SEC.[53] But the SEC is required to update the threshold for Emerging Growth Companies every five years pursuant to an explicit indexing procedure in the Securities Act.[54] The limits on campaign contributions to Congressional candidates have been indexed for inflation since the 1970s, but the levels of improper contributions that trigger criminal penalties are nominally stated.[55] Thresholds for certain federal procurement supervision policies must be updated using the urban CPI every five years.[56] But many of the appropriations funding those contracts remain vulnerable to inflation.[57]

The tax code is especially messy. The base income above which social security benefits are taxable has not been adjusted since the provision was introduced in 1983.[58] But the contribution limits for tax-deferred retirement plans are required to be updated annually by a set cost of living adjustment.[59] The Child Tax Credit and the income thresholds at which it phases out have oscillated between indexation and being allowed to drift since being introduced in 1997.[60] The sustainable fuel credits contained in the recent Inflation Reduction Act are, perhaps optimistically, not indexed.[61]

Most of the aforementioned statutes are of relatively recent vintage. Nevertheless, significant erosion can occur over just a few decades. The $25,000 nominal income threshold that triggers taxation of social security benefits would need to double to keep pace with inflation that has occurred since it was enacted in 1993.[62] Older statutes have eroded still further.

To be sure, Congress has made some efforts to clean up statutes that have significantly drifted. In Dodd-Frank, for example, Congress finally indexed the upper bound for consumer leases to be exempt from requirements of the Truth in Lending Act after letting it drift from 1968 to 2011.[63] And Part III.C outlines steps it has taken to make certain provisions more resilient. But as a general matter, the law remains a crazy quilt of nominally stated and inflation-resilient provisions. And cleanup efforts have yet to reach many

---

53 Monson, *supra* note 24, at 38.

54 15 U.S.C. § 77(b).

55 *Compare* 52 U.S.C. § 30116(c) (contribution limits indexed annually), *with* 52 U.S.C. § 415(d) (not indexing penalty threshold).

56 41 U.S.C. § 1908. *See also* 75 FED. REG. 53129 (2010) (NASA updates); 85 FED. REG. 62485 (2020) (Dept. of Labor updates).

57 *See, e.g.*, Andrew Duehren, *Inflation threatens to erode impact of 1 trillion infrastructure law*, WALL ST. J. (Feb. 24, 2022), https://www.wsj.com/articles/inflation-threatens-to-erode-impact-of-1-trillion-infrastructure-law-11645698601.

58 26 U.S.C § 86; PL 98–21, April 20, 1983, 97 Stat 65 at § 121 (showing original text).

59 26 U.S.C. § 415(d). *See also* Kelly Tyko, *IRS increases 401(k), IRA contribution limits for inflation*, AXIOS (Oct. 21, 2022), https://www.axios.com/2022/10/21/401k-contribution-irs-limits-retirement-tax-benefit.

60 MARGOT L. CRANDALL-HOLLICK, CONG. RSCH. SERV., R45124, THE CHILD TAX CREDIT: LEGISLATIVE HISTORY 2-3 (2021).

61 *See* Inflation Reduction Act § 40B, Pub. L. No. 117-169 (Aug. 16, 2022).

62 Calculation based on *Historical CPI-U data*, *supra* note 8.

63 15 U.S.C. § 1603 (enacted 2011) (updating to $50,000 and indexing); *see also* Pub.L. 90-321, Title I, § 104 (1968) (showing original threshold of $25,000).

adrift provisions. Given current legislative productivity,[64] a fully resilient code seems far away indeed.

Many factors contribute to this status quo, but the inability of judicial and administrative actors to respond is critical. Their impotence arises from two limits rooted in separation of powers—interpretive methodology and constraints on delegation. The following subsections analyze these limits. The analysis is organized by the type of limit, rather than actor (courts versus agencies) or problem (designing resilient new laws versus correcting old ones), because the limitations apply across actors and across problems, albeit with different force.

1. Limits imposed by interpretive methodology: nominalism & plain meaning

One strategy for keeping law current is to rely on courts and agencies to interpret enacted money values in real terms. Under this approach a value of $100 enacted in 1980 would simply be read at its current value. This would be analogous to the common law's reliance on judicial reinterpretation of precedents to fit new contexts.[65] Leaving aside the administrative challenges, two interrelated legal barriers foreclose this strategy: nominalism and the plain meaning rule.

a. The role of nominalism

Nominalism refers to the tendency to treat all dollars as having the same value regardless of purchasing power. Dollars today are equivalent to dollars tomorrow and paper dollars are equivalent to coined dollars. To the extent nominalism is incorporated into a legal system it constrains judges and officials by limiting a money value to a single meaning: they must read $100 to mean $100 in current money. This is not the inevitable approach. At various times and places a purchasing-power definition has been employed.[66] But nominalism defined money at common law in the 17th and 18th centuries and is a core assumption of the legal systems of most economically important nations today.[67]

In the United States the nominalist approach prevails, though its source is not entirely clear. One possible source is § 20 of the Coinage Act of 1792 which states "the money of account of the United States shall be expressed in dollars or units, dimes or tenths . . . and that all accounts in the public

64 *Cf.* GOVTRACK, Statistics and Historical Comparison: Bills by Final Status (Apr. 5, 2023, 5:44 PM), https://www.govtrack.us/congress/bills/statistics.

65 *See* CALABRESI, *supra* note 11, at 4.

66 Rosenn, *supra* note 16, at 38.

67 *Id.*

offices, and all proceedings in the courts of the United States, shall be kept and had in conformity to these regulations."[68] While the requirement of nominalism is not explicit in the text, the Supreme Court clarified its meaning in 1868 in *Bronson v. Rodes*. In Bronson, the Court was asked to decide whether Civil War-era paper money could satisfy obligations under preexisting contracts. It held that the general wording of § 20 required all lawful money to use the same units.[69] Thus, by virtue of their shared units, paper dollars and gold dollars had the same legal value despite their different purchasing powers.[70] However, because *Bronson* involved a contract which specified payment in gold it did not provide the Court with an opportunity to squarely hold that all dollars were equivalent. That came two years later in *Knox v. Lee*, one of the *Legal Tender Cases*.[71] The *Knox* court held that Congress' declaration of the paper greenback as lawful money made a greenback dollar equivalent to a gold dollar.[72] Thus forcing individuals to accept paper money with less purchasing power than gold-backed money did not violate the Contracts Clause.[73] The *Knox* majority did not explicitly discuss the 1792 Act but its reasoning is quite similar to that in *Bronson*, arguing that nominal equivalence existed even prior to the advent of paper money.[74] This was so "not because of the intrinsic value of the coin, but because of its legal value."[75]

The *Bronson* and *Knox* decisions were immediately controversial. Justice Clifford authored a fifty three page dissent in *Knox* arguing in part that § 20 of the 1792 Coinage Act should be read exactly opposite to the majority's dicta in *Bronson.*[76] He thought that other provisions of the act assigning weights to silver and gold coins treated "unit", "dollar", and "coined dollar" as synonyms.[77] Thus, in his view, the dollar was defined according to a certain real value in gold.[78] Other commentators advanced similar arguments.[79]

Despite this controversy, nominal equivalence was never seriously challenged after *Knox.* By the next century it had transcended its origins as a debatable construction and was considered a canonical legal rule. The most cited articulation comes from Justice Holmes in 1926 in *Die Deutsch Bank*

---

68 1 Stat. 246, § 20 (1792).

69 Bronson v. Rodes, 74 U.S. 229, 254 (1868).

70 *See id.*

71 Knox v. Lee, 79 U.S. 457 (1870).

72 *Id.*

73 *Id.* at 550–51.

74 *Id.* at 548–49.

75 *Id.* at 549.

76 *Knox*, 79 U.S. at 593–94.

77 *Id.* at 594.

78 *Id.*

79 *See* Kenneth W. Dam, *The Legal Tender Cases*, 1981 SUP. CT. REV. 367, 382 (1981) (describing reaction).

Filiale Nurnberg v. Humphrey.[80] The case involved the calculation of damages from a breach of a contract by Deutsche Bank.[81] The contract was denominated in German Marks, and the liability was incurred in 1915 but the suit to collect was not brought until 1921.[82] During the intervening years the Mark depreciated, thus Humphrey, the injured party, wanted damages based on an earlier value of the mark while Deutsche Bank preferred a more recent valuation. Justice Holmes, writing for the majority, rejected Humphrey's argument that damages should be valued in real terms and held that "an obligation in terms of the currency of a country takes the risk of currency fluctuations . . . [o]bviously in fact a dollar or a mark may have different values at different times but to the law that establishes it[,] it is always the same."[83] The four dissenters agreed on this point.[84] Notably, neither the majority nor dissent referenced the 1792 Act or *Bronson* for this proposition. Instead they portrayed nominal equivalence as a settled principle shared across legal systems.[85] Indeed, Justice Holmes considered this approach so uncontroversial that he "refrain[ed] from citing the many cases that have touched upon it and content[ed himself] with stating . . . the proper rule."[86]

Since then courts have consistently interpreted money values in contracts with a nominal approach.[87] However, they have provided little guidance on how the doctrine extends to statutory interpretation and questions remain as to the source and strength of nominalism. When confronted with issues of nominalism in statutory interpretation the Supreme Court has generally ignored the issue or adopted the principle without explanation. In 1985 in *Walters v. National Ass'n of Radiation Survivors* the Court confronted a Due Process challenge to the cap on attorney's fees related to veteran's benefit disputes, which was set at $10 in 1862.[88] The majority opinion found the cap justified by the enacting legislature's interest in making dispute proceedings non-adversarial,[89] but it *made no mention* of the fact that in 1862 the $10 cap was equivalent to $580 in 1985 dollars.[90] As the dissent pointed out, this threshold would have enabled limited rather than zero legal assistance in the

---

80 272 U.S. 517 (1926); *see also* Shaw, Sahill, Albion & Co. v. The Fredericksburg, 189 F.2d 952, 955 (2d Cir. 1951) (citing *Humphrey*).

81 272 U.S. at 518.

82 *Id.*

83 *Id.* a 519.

84 *Id.* at 522.

85 *Id.* at 519 (Homes, J.) ("We may assume . . . [the] liability . . . by the German law . . . was fixed in [M]arks only, not at the extrinsic value those marks then had."); *id.* at 521 (Sutherland, J., dissenting) ("[That the] liability was fixed by German law at a certain number of German marks . . . and was open to satisfaction in that number of marks . . . however much the mark might have fallen in value . . . of course is true if the payment be made in Germany, where marks remain legal tender at all times.").

86 *Id.* at 520.

87 *See* Rosenn, *supra* note 16, at 59.

88 473 U.S. 305, 307–08 (1985)

89 *Id.* at 333.

90 *Id.* at 361 (Stevens, J., dissenting).

1860s.[91] In *Schweiker v. Gray Panthers* the Court rejected a challenge to Medicaid regulations that required a certain amount of spousal income above a threshold to be considered available to support the other spouse when calculating Medicaid benefits.[92] The plaintiffs argued those thresholds were too low to assure spouses had adequate funds to live on and thus were outside the agency's authority which only allowed it to include "available" spousal funds.[93] The Court explicitly denied the relevance of inflation, stating that even though inflationary erosion was the real cause of the hardships experienced by the challengers—because states had failed to update the nominally-stated exclusions for spousal income—such considerations were not relevant to interpreting whether the regulation was valid.[94] It did not explain why dollar values had to be interpreted this way. These cases make it clear that nominalism also prevails in statutory interpretation, but they say little about the basis for the rule. One could take their complete lack of discussion of the 1792 Coinage Act or the contract precedents as indicating that nominalism is a background principle of interpretation rather than the product of legislative guidance. But the Court has not explicitly adopted this view.

Recent cases involving contract construction point towards a statutory source. These cases deal with the aftermath of the sole amendment to section 20 of the 1792 Act. In 1982, Congress engaged in a project to recodify Chapter 31 of the US Code.[95] During the recodification, certain supposedly non-substantive changes were made, including to section 20 (previously codified in 31 U.S.C. § 371, now recodified in 31 U.S.C. § 5101).[96] First, the phrase "money of account" in the original was adjusted to "money."[97] According to a House of Representatives report, this was done to eliminate unnecessary words.[98] The second change removed the "or units" language next to "dollar" as redundant, though the "or tenths" and "or hundredths" next to dimes and cents were kept. [99] The third change removed the final clause requiring that "all accounts in the public offices and all proceedings in the courts shall be kept and had in conformity to this regulation" as surplus.[100]

Several courts and authorities have found the repeal of the final clause of § 20 consequential. In *Competex, S.A. v. Labow*, the Second Circuit

---

91 *Id.*

92 Schweiker v. Gray Panthers, 453 U.S. 34, 38–42 (1981).

93 *Id.*

94 *Id.* at 49 n.19.

95 Act of Sep. 13, 1982, Pub. L. No. 97-258 (to revise, codify, and enact without substantive change certain general and permanent laws, related to money and finance, as title 31, United States Code, "Money and Finance").

96 H. REP. NO. 97-251, at 146–47 (1982).

97 *Id.*

98 *Id.* at 146.

99 H. REP. NO. 97-251, at 147 (1982).

100 *Id.*

considered whether a judgment by a United States court had to be in dollars or could be issued in foreign currency unaffected by exchange rate fluctuations.[101] This involved the same fundamental question of whether § 20 prohibited a court from considering real changes in the dollar's value, only in this case the variance was cross-jurisdiction as well as over time. The court observed that the assumption "that American judgments must be entered in dollars . . . probably deserves reexamination in light of the repeal of [the last clause of] section 20."[102] It also discussed the unsettled foundation of the doctrine, having been derived alternatively from common law notions of sovereignty and from the 1792 Act.[103] Subsequent opinions and authorities have also found the repeal consequential.[104]

This winding history of nominalism has several implications for the inflationary drift problem. Regardless of its source, nominalism is clearly the law of the land in contractual and statutory interpretation. Dollar amounts cannot be read in real terms. However, the unsettled source of the doctrine raises questions about the options available for addressing inflationary drift. A statutory source for nominalism provides room only for Congress to institute alternate assumptions.

b. The Plain Meaning Rule

Nominalism integrates with the second limit on interpretive solutions to drift, the plain meaning rule. This rule requires courts and agencies to always "give effect to the unambiguously expressed intent of Congress."[105] If "dollar" clearly means current dollars because of nominalism then courts and agencies are not permitted to override that meaning. This faithfulness to written law has been acknowledged since the beginning of the republic.[106] Today it is nothing short of canonical, a cornerstone of separation of powers.[107]

---

101 Competex, S.A. v. Labow, 783 F.2d 333, 337 (2d Cir. 1986).

102 *Id.* at 337.

103 *Id.*

104 *See* Leidos, Inc. v. Hellenic Republic, 881 F.3d 213, 218 n.5 (D.C. Cir. 2018); In re Oil Spill by Amoco Cadiz, 954 F.2d 1279, 1328 (7th Cir. 1992); Mitsui & Co. v. Oceantrawl Corp., 906 F. Supp. 202, 203 (S.D.N.Y. 1995). *See also* RESTATEMENT (THIRD) FOREIGN REL.'S LAW U.S. § 823 cmt. a (Am. L. Inst. 1987) (finding dollar judgment requirement abrogated by repeal of § 20).

105 Chevron U.S.A. v. Natural Resources Defense Council, 46 U.S. 837, 843 (1984).

106 Marbury v. Madison, 5 U.S. 137, 176 (1803) ("To what purpose are powers limited, and to what purpose is that limitation committed to writing, if these limits may at any time be passed by those intended to be restrained?").

107 *See* Cass R. Sunstein, *Nondelegation Canons*, 67 UNIV. CHI. L. REV. 315, 329 (2000) ("When Congress has spoken clearly, everyone agrees that agencies are bound by what Congress has said."); John F. Manning, *What Divides Textualists and Purposivists*, 106 COLUM. L. REV. 70, 87 (2006); Justice Elena Kagan, The Scalia Lecture: A Dialogue with Justice Kagan on the Reading of Statutes, at 8:28 (Nov. 17, 2015) ("I think we're all textualists now . . . ").

Perhaps because the principle is so widely accepted, courts have had little chance to confront the issue of inflationary drift. But the cases that do address the issue make clear that non-legislative actors cannot disregard clearly enacted monetary values, no matter how much they have been or might be eroded. The most revealing strand concerns challenges to campaign contribution limits because they have found drift to be a cause of statutory invalidity and still refused to revise the plain text. In 2006 in *Randall v. Sorrell,* the Supreme Court considered a First Amendment challenge to a Vermont election law which capped individual contributions at a nominal $200 for a state representative as of 1997.[108] This was the lowest contribution limit in the country and well below the lowest limit the Court had previously upheld ($1,275) which was, as the court stressed, indexed for inflation.[109] A splintered court ultimately struck down the contribution limit as "too restrictive" and thus overly limiting of the ability to conduct an effective campaign, especially by challengers.[110] The failure to index for inflation was one of the factors which Justice Breyer's plurality opinion found, taken together, made the limit unconstitutional.[111] In particular, the plurality was concerned that the legislature might not "diligently police" the erosion of levels in the future.[112]

Despite misgivings about the legislature's ability to cope with inflationary erosion, the Court was unequivocal in rejecting any invitation to make its own indexing amendment, characterizing such a move as "writing words into the statute" in a manner clearly beyond its authority.[113] In just three sentences Justice Breyer's opinion resolved that the proper remedy was to completely invalidate the contribution limits and leave the legislature "free to rewrite those provisions."[114] Justice Alito and the Chief Justice joined this portion of Justice Breyer's opinion and none of the concurring or dissenting opinions expressed disagreement with the remedy.

In 2019, the Supreme Court confronted a similar challenge to Alaska's unindexed contribution limit, set at $500 twenty-three years prior.[115] In a per curiam opinion, the Court again found the provision's lack of indexation was one of several "danger signs" that might require special justification.[116] On remand, the Ninth Circuit held that, in part because of the lack of indexation,

---

108 Randall v. Sorrell, 548 U.S. 230, 253 (2006).

109 *Id.* at 251.

110 *Id.* at 253–54.

111 *Id.* at 261.

112 *Id.* Importantly, the Court only found the statute invalid due to the combination of factors; failure to index was not individually sufficient to create a constitutional violation. Since *Randall*, lower courts have not regarded failure to index as an independent ground for invalidation. *See* Anh Cao v. Federal Election Comm'n, 688 F. Supp. 2d 498, 548 (E.D. La. 2010).

113 *Id.* at 262.

114 *Id.* at 262.

115 Thompson v. Hebdon, 140 S. Ct. 348, 350–52 (2019).

116 *Id.* at 352.

the low contribution limit was not justified.[117] But, consistent with *Randall,* both the Court and the Ninth Circuit declined to remedy the infirmity by writing indexation into the law.[118]

Even when provisions do not contain explicit money values—thus arguably opening space for finding ambiguity—courts and agencies have been reluctant to read such statutes in real terms. The most notable instance involves the calculation of capital gains, discussed in more detail in Part 0. The appreciation (gain) on capital assets is taxed preferentially to other income and is calculated as the price upon a realization event (typically a sale) minus the basis (defined in statute as the "cost of such property").[119] Inflationary drift enters the equation through the basis because cost is measured nominally. Arguably this consequence could be avoided by defining "cost" in the definition of basis in real terms, but such arguments have been rejected since they were first raised.[120] It is conceivable that a similar non-numeric provision relying on more ambiguous terminology could be unilaterally indexed, but such provisions are far from the norm.[121]

Outside the monetary context, courts have been even clearer in their disapproval of judicial remedies to existing obsolescence. For example, the doctrine of desuetude—the abrogation of outdated and unenforced criminal statutes by courts—has been repeatedly and near-unanimously rejected.[122] The principal reason given is that such statutory revision of the plain words of a statute is clearly antithetical to the fundamental precepts of separation of powers.[123] And administrative efforts to update outdated statutes via interpretation, while occasionally successful, have been rare.[124]

2. Delegation Barriers

The second non-legislative strategy for addressing drift would be to delegate the task of updating monetary values to administrative or judicial

---

117 Thompson v. Hebdon, 7 F.4th 811, 821–22, 827 (9th Cir. 2021).

118 *See id.* at 827; Thomas R. Lucas, Alaska Public Offices Comm., AO 21-09-CD, Contribution limits in light of Ninth Circuit Court of Appeals ruling, at *2 (Nov. 3, 2021).

119 I.R.C. §§ 1001(a), 1011(a), 1012(a).

120 *See* Jeff Strnad, *Deflation and the Income Tax*, 59 TAX L. REV. 243, 244 (2006); Bruce Bartlett, *Indexing Capital Gains by Fiat*, 135 TAX NOTES 883, 884 (2012). *See also* Bates v. United States, 108 F.2d 407, 408 (1939), *cert. denied*, 309 U.S. 666 (1940) (rejecting real interpretation of cost); Legal Authority of the Department of the Treasury to Issue Regulations Indexing Capital Gains for Inflation, 16 Op. O.L.C. 136, 146–51 (1992) (same).

121 *Cf.* Verizon Communications, Inc. v. F.C.C., 535 U.S. 467, 500 (2002) (finding the use of "cost" in statutory guidance on rate setting ambiguous enough to encompass real or historical cost).

122 Note, Desuetude, 119 HARV. L. REV. 2209, 2209 (2006).

123 *Id.* at 2213.

124 *Cf.* Jody Freeman & David B. Spence, *Old Statutes, New Problems*, 163 UNIV. PENN. L. REV. 1, 19 (2014) (describing two unusual efforts in the environmental and energy contexts).

actors. To a limited extent, this already occurs. Over forty agencies update civil penalties annually based on a prescribed formula.[125] The U.S. Sentencing Commission can suggest inflation adjustments subject to congressional override.[126] And certain provisions of the tax code receive specified yearly cost of living adjustments.[127] These delegations mostly involve applying a pre-determined index formula or are subject to Congressional veto, but a few more discretionary delegations exist. The SEC, for example, is actually required to review the accredited investor definition at least every four years and may modify it as needed based on, among other things, the "state of the economy."[128]

Any viable delegation solution will need to extend well beyond the scope of these existing delegations. Mechanical and veto procedures have serious practical shortcomings, discussed in Part 0 below. The SEC model is closer to the flexibility a delegation approach would need to achieve, but it is already an outlier. The SEC has enjoyed greater administrative discretion than many agencies for historical reasons,[129] but that latitude is increasingly in question.[130] Extending a similar approach to other areas would be a significant departure from tradition. Moreover, none of the existing modes address how delegates can update statutes for which no adjustment instructions are likely to be provided.

The main constraint facing such a solution will be the nondelegation doctrine. The doctrine is important but murky because of the tension between its strong theoretical formalism and the weakness with which it has been customarily applied. The basic theory of nondelegation is that agencies may not engage in lawmaking that usurps the legislative power granted to Congress.[131] There are good reasons for this. Certain decisions may be so important they

---

125 *See* Section 4. Civil Penalties round two: flexible formulas but oversight challenges 0, *infra*.

126 *See* Section 0, *infra*.

127 *See* Section 0, *infra*.

128 Dodd–Frank Wall Street Reform and Consumer Protection Act, Pub. L. No. 111-203, § 413, 124 Stat. 1376 (2010).

129 *Cf.* Matthew C. Stephenson, *Legislative Allocation of Delegated Power: Uncertainty, Risk and the Choice Between Agencies and* Courts, 119 HARV. L. REV. 1035, 1040 n.12 (2006) (describing unusually large grant of discretion to SEC in 1934 as reflecting Congressional desire to enable regulator-driven development of detailed scheme).

130 *Cf.* Jarkesy v. Sec. & Exch. Comm'n, 34 F.4th 446, 461 (5th Cir. 2022) (narrowing traditional enforcement power); Lucia v. Sec. & Exch. Comm'n, 138 S. Ct. 2044, 2051 (2018) (invalidating traditional ALJ appointment procedure).

131 *See* Martin H. Redish., *Pragmatic Formalism, Separation of Powers, and the Need to Revisit the Nondelegation Doctrine*, 51 LOY. UNIV. CHI. L. J. 363, 383 (2019). Nondelegation is sometimes framed as focusing on what Congress cannot grant and other times on what agencies cannot do. *Compare* Gundy v. United States, 139 S. Ct. 2116, 2121 (2019) (Kagan, J.) ("The nondelegation doctrine bars Congress from transferring its power."), *with* Cass R. Sunstein, *The American Nondelegation Doctrine*, 86 GEO. WASH. L. REV. 1181, 1186–89 (2018) ("[The doctrine] does not forbid Congress from doing anything at all. Instead it forbids agencies from acting in particular ways."). There are important differences between these approaches, but they agree that agencies and Congress have some nonoverlapping roles.

should only be made by a highly democratically responsive legislature.[132] The framers' carefully calibrated system of ambition checking ambition may fall apart if the proper roles of the branches are not respected.[133] And, to the extent the Constitutional text prescribes limits on non-legislative lawmaking, those limits should be respected.[134] While nondelegation is traditionally thought of as a limit on administrative activity, most of its concerns are just as applicable to judicial updating of statutes.[135]

The nondelegation doctrine has been administered by requiring Congress to provide an intelligible principle to guide agency action, the idea being that such guidance confines an agency merely to implementing Congress' intent rather than making its own judgments into law.[136] This has been described as a sliding scale.[137] The more important the area of discretion the more guidance Congress must provide.[138] Alternate methods of distinguishing between those activities reserved to Congress and those permitted to other branches—such as asking whether agencies are merely filling up the details of a statutory scheme, or whether such a delegation accords with historically accepted practices—have also gained traction among some jurists.[139] But the underlying objective of all approaches is the same: to keep the most sensitive and open-ended judgments within the legislature while allowing enough agency discretion for government to function.

In practice, nondelegation is honored in the breach. Congress instructs other branches to make all sorts of decisions that implicate similar policy choices to those associated with lawmaking.[140] Because of the ubiquity and necessity of those delegations, courts virtually never reject such schemes outright.[141] Consequently, some commentators have declared nondelegation dead.[142] Some historians argue it was never really alive to begin with—just a Frankenstein doctrine trotted out as a rear guard action against political and

---

132 Redish, *supra* note 131, at 384.

133 *Id.* at 384–85.

134 *Id.*

135 *See* Margaret H. Lemos, *The Other Delegate: Judicially Administered Statutes and the Nondelegation Doctrine*, 81 S. CALIF. L. REV. 405, 428 (2008).

136 *Id.* at 416–18.

137 *See* Sean P. Sullivan, *Powers, But How Much Power? Game Theory and the Nondelegation Principle*, 104 VA. L. REV. 1229, 1234 (2018).

138 *Id.*

139 *See* Gundy v. United States, 139 S. Ct 2116, 2136 (2019) (Gorsuch, J., dissenting) (fill up details approach); *id.* at 2129–30 (Kagan, J.) (looking to prior grants of discretion).

140 *See* David Schoenbrod, Power Without Responsibility: How Congress Abuses the People Through Delegation, 58 (1993).

141 Redish, *supra* note 131, at 378.

142 *See* Eric A. Posner & Adrian Vermeule, *Interring the Nondelegation Doctrine*, 69 U. CHI. L. REV. 1721, 1723 (2002).

governmental change.[143] This may be true in certain areas, such as antitrust, where doctrinal evolution has long been entrusted to administrative and judicial actors.[144] But the doctrine retains significant rhetorical and legal force. It is frequently cited and several members of the Supreme Court have expressed a desire for more rigorous policing of administrative delegations.[145] Moreover, formalist separation of powers concerns that motivate the doctrine clearly still have teeth, especially in the presence of unambiguous statutes.[146] Courts regularly implement related tools such as the Major Questions doctrine and interpretive canons to blunt the most extreme uses of delegated authority.[147] It would be foolish to think a meaningful expansion of non-congressional inflation updating would escape nondelegation scrutiny.

The question then is what that scrutiny will look like for different updating approaches. The most extreme option would be to update statutes without any instructions from Congress. This scenario is plausible because updating instructions require legislation. Congress lacks the time to do this for the multitude of vulnerable provisions and may be disinclined to invite the political battles that arise when well-established laws are revisited. Unfortunately, completely uninstructed updating is not viable. The one thing nearly everyone agrees on with regard to nondelegation is that Congress, not agencies or courts, must construct the principle that guides non-legislative discretion.[148] In fact, such unilateral updating does not appear to have ever been attempted. The one case in which the possibility was significantly debated involved the aforementioned effort to modify the capital gains formula to account for inflation.[149] The White House, Treasury, and Justice Department all resoundingly rejected this proposal.[150] Their main objection was that the proposal skipped over *Chevron*'s first step.[151] When Congress "writes legislation in specific terms . . . [that do] not leave policy choices to be resolved by an administrative agency, then Congress's decision binds both the executive branch and the judiciary."[152] This finding was reaffirmed by academic

---

143 *See* Keith E. Whittington & Jason Iuliano, *The Myth of the Nondelegation Doctrine*, 165 UNIV. PENN. L. REV. 379, 380 (2017); *cf.* Nikolas Bowie & Daphna Renan, *The Separation of Powers Counterrevolution*, 131 YALE L. J. 2020, 2024 (2022).

144 *Cf.* Lemos, *supra* note 135, at 461 (describing incompatibility of large-scale delegation in antirust with nondelegation).

145 *See, e.g.*, Gutierrez-Brizuela v. Lynch, 834 F.3d 1142, 1152 (10th Cir. 2016) (Gorsuch, J., concurring) (describing *Chevron* as an "abdication of judicial duty"); Gundy v. United States, 139 S. Ct. 2116, 2130–31 (2019) (Alito, J., concurring in judgment); Gillian E. Metzger, *The Roberts Court and Administrative Law*, 2019 SUP. CT. REV. 1 (2019) (describing incremental narrowing of administrative latitude).

146 Sunstein, *supra* note 131, *passim*.

147 *Id.*

148 *See* Redish, *supra* note 131, at 375.

149 *See* Section 0, *infra*.

150 *See* Section 0, *infra*.

151 *See* Legal Authority of the Department of the Treasury to Issue Regulations Indexing Capital Gains for Inflation, 16 Op. O.L.C. 136, 146–51 (1992).

152 *Id.* at 139.

commentators when the same proposal was revived during the Trump Administration.[153]

A more limited form of unilateral updating might be accomplished via re-interpretation of an agency's existing grants of discretion. The viability of this approach will largely turn on the nature of the discretion-granting provision and the regulatory context. Some cases will fit well with the task of updating statutory text for inflation. The SEC is permitted under § 413 of the Dodd-Frank Act to modify the threshold for the accredited investor definition "for the protection of investors, in the public interest, and in light of the economy."[154] This authority is easily broad enough to permit inflation updates.[155] On the other hand, the Department of Labor might be more hamstrung in overriding the $2,000 value threshold for federal contracts to be subject to the Davis Bacon Act's minimum wage requirements. The Department has been given authority to administer the act in certain ways, such as setting minimum wages at levels "the Secretary of Labor determines to be prevailing."[156] But only the most strained reading could extend that authority to modifying the coverage threshold. From the outset then, one can see this approach may only be selectively applicable.

Attempts to stretch existing authority also suffer from a number of practical limitations. It may be difficult to sustain this approach year over year thus some of the expectation-mismatch problems of drift may persist. The SEC's authority only allows it to "review" the threshold periodically not necessarily to prescribe a predictable adjustment procedure for the medium and long term.[157] Any rule adjusting a monetary threshold, penalty, or payout will also likely have to go through notice and comment,[158] possibly delaying updates and making them less predictable. And such updating will introduce inconsistency within regulatory schemes where regulatory discretion only extends to certain provisions.

---

153 *See* Daniel Hemel & David Kamin, *The False Promise of Presidential Indexation*, 36 YALE J. ON REG. 693, 706–15 (2019).

154 Dodd-Frank Wall Street Reform and Consumer Protection Act, Pub. L. 111-203, § 413, 124 Stat. 1376 (2016).

155 *Cf.* Revisions of Ltd. Offering Exemptions in Regul. D, Release No. 8828, 72 FED. REG. 45115, 45126 (Aug. 10, 2007) (implying existence of pre Dodd-Frank authority to make one-off inflation adjustment but declining to use).

156 40 U.S.C. § 3142(b). *See also* 87 FED. REG. 15698, 15702 (Mar. 18, 2022).

157 *Cf. id.* ("not less frequently than once every four years . . . the Commission shall undertake a review . . . to determine whether such requirements should be adjusted").

158 *See* Community Nutrition Inst. v. Young, 818 F.2d 943, 945–48 (D.C. Cir. 1987) (holding that FDA rule setting quantitative limits for certain food contaminants was a binding legislative rule and thus required to go through notice and comment). *Cf.* Michael A. Rosenhouse, Annotation, *Construction and Application of Good Cause Exception to Notice and Comment Rulemaking*, 26 A.L.R. Fed. 2d 97, at § 2 (2008) (summarizing cases on applicability of § 553 good cause exception to price setting regulations and finding exception only found applicable when "dislocations likely to be caused by advanced notice . . . but not otherwise.").

Agencies with less clear modification authority than the SEC might also run afoul of the Major Questions doctrine—a special application of the non-delegation approach.[159] The doctrine comes into play when an agency interprets its power to promulgate regulations in an extremely broad or novel way.[160] If it is not clear that Congress intended the agency to have such rule-making authority the Court will reject the interpretation.[161] The Court has typically found regulations involving issues of great economic or political significance, such as public health or the environment, to be beyond agency authority absent a very clear commitment of regulatory power in the relevant area.[162] But the doctrine potentially encompasses far more than just attempts to address front-page problems. More specialized assertions of authority can also be struck down for being too broad. *MCI Telecommunications Corp. v. American Telephone & Telegraph Co.*, while not explicitly an invocation of the doctrine, is closely on point.[163] In *MCI* the FCC attempted to interpret its statutory authority to "for good cause, modify any requirement" of the rate filing provisions of the Communications Act of 1934 as allowing it to make the filing requirement optional for long-distance carriers.[164] The Court rejected this construction, finding it "highly unlikely" that Congress would assign the determination of whether an entire industry or part of an industry, will be rate regulated to an agency in such a subtle way.[165] Recent Supreme Court opinions, such as *West Virginia v. EPA*, have also found departures from settled understandings of authority relevant to the major questions analysis.[166]

These recent decisions indicate that merely demonstrating a delegation of regulatory discretion may not be enough for a novel regulatory move to survive major questions review. Agencies attempting to update inflation-eroded provisions may have to prove that the issue at hand is minor across several dimensions and that their interpretation is consistent with prior ones.

---

159 *See* Gundy v. United States, 139 S. Ct 2116, 2142 (2019) (Gorsuch, J., dissenting) ("Although it is nominally a canon of statutory construction, we apply the major questions doctrine in service of the constitutional rule that Congress may not divest itself of its legislative power by transferring that power to an executive agency.").

160 *See* Nat'l Fed'n of Indep. Bus. v. Dep't. of Labor, 142 S. Ct. 661, 663 (2022) (rejecting OSHA attempt to interpret to promulgate standards "reasonably necessary . . . to provide safe or healthful employment" as authorizing *nationwide* vaccine mandate); West Virginia v. Env't Prot. Agency, 142 S. Ct 2587, 2601, (2022) (rejecting attempt to interpret § 111 of Clean Air Act—permitting agency to promulgate "federal standards of performance" reflecting the "best system of emission reduction" for certain pollution sources—to allow issuance of standards incentivizing shifts to cleaner fuels *in contravention of historical practice*).

161 West Virginia v. Env't Prot. Agency, 142 S. Ct. 2587, 2607–08 (2022).

162 CONG. RSCH. SERV., IFI2077, THE MAJOR QUESTIONS DOCTRINE 1 (2022).

163 512 U.S. 218 (1994)

164 *Id.* at 225.

165 *Id.* at 231.

166 *See West Virginia*, *supra* note 161, at 2595; Food and Drug Admin. v. Brown & Williamson Tobacco Corp., 529 U.S. 120, 154–55 (2000).

This would be quite difficult. There are strong arguments, especially when one considers the nominalism precedents discussed above, that Congress has kept the discretion to adapt the law to macroeconomic shocks to itself. Indeed, one is hard-pressed to think of an issue of more far-reaching economic and political significance than inflation. And consistency with prior interpretations would be nearly impossible to show. No agency has attempted to exercise such authority previously.

The final option is the promulgation of a clear updating directive from Congress. This presents a number of practical challenges which are discussed at length in Part 0, the most important being the vast number of provisions needing attention. This strategy is also not immune to delegation challenges. The difficulty lies in balancing flexibility with constraint in the instructions. Congress could provide an updating procedure with zero room for deviation. This is essentially what it has done with tax bracket indexation: the IRS is required to update income tax brackets each year by a completely predetermined formula based on the CPI-U.[167] Such an approach would present no delegation issues. However, the provisions vulnerable to inflationary drift are diverse, complex, and touch on many sensitive policy areas. This raises the question of how much flexibility Congress can provide before the agency begins to legislate thresholds, penalties, or payouts. The law is unsettled on this. Existing updating delegations indicate Congress likely has room to maneuver so long as the general package adds up to a certain level of restraint. The Sentencing Commission is utterly free to choose when and by what procedure it updates its guidelines, but Congress has a direct veto over these.[168] Meanwhile, agencies can make civil penalty updates without the need for explicit approval from Congress, but only according to a prescribed formula and with a narrow economic harm exception for initial adjustments that must be approved by the OMB.[169] To be sure, the propriety of these updating mechanisms has not been litigated and not all possible delegation packages would be allowed. Congress could probably not permit an agency to unilaterally decide when and if an inflation adjustment could be applied.[170] It may also be unable to give an agency discretion to design an updating process that is triggered by especially open-ended judgments such as a finding of economic hardship, though this may be more permissible in conjunction with a prescribed methodology.[171] What is clear is that any approach beyond a preset

---

[167] 26 U.S.C. § 1(f).

[168] Section 0, *infra.*

[169] Section 0, *infra*.

[170] *See* Gundy v. United States, 139 S. Ct 2116, 2123 (2019) (A provision allowing executive officer to "change her policy for any reason and at any time . . . would face a nondelegation question.").

[171] *Compare* A. L. A. Schechter Poultry Corp. v. United States, 295 U.S. 495, 552–553 (1935) (Cardozo, J., concurring) (criticizing a hypothetical delegation that would allow president to mandate competitive practices thought "desirable"), *with* Yakus v. United States, 321 U.S. 414, 424–27 (1944) (finding power to set maximum prices contingent on administrator's factual determination was permissible delegation because, unlike the delegation in *Schechter*, the action was constrained by instructions that administrator consider specified factors, including historically prevailing prices, in setting price).

formula or a direct approval mechanism goes beyond traditional delegations. More flexible delegations may be permissible but only up to an as-yet indeterminate point.

* * * *

As the forgoing discussion demonstrates, the inflationary drift problem is serious and widespread. Addressing it is made harder by doctrinal obstacles that require most solutions to run through Congress. The first Congress imported nominalism into the law, foreclosing findings of ambiguity with respect to monetary values. Legislation is arguably required to adjust that assumption. Only Congress can revise unambiguous statutory text. And delegated updating requires, at minimum, legislative development of updating procedures for hundreds of provisions, procedures which will have to be somewhat inflexible to satisfy intelligible principle requirements.

These limits can be traced to a key source: an understanding of separation of powers that requires careful segregation of the power to create and revise clearly written laws. That highly formal understanding is not universally applied, but it is far from a dead letter, especially when it comes to revision of unambiguous, specific statutory provisions. There may be good reasons to construe separation of powers so strictly. But the approach is risky. It depends on a single actor to avoid the expectation mismatch challenges of inflationary drift. The next section evaluates whether Congress has been successful in that effort.

III. THE CONSTITUTIONAL ALLOCATION OF LAWMAKING POWER HAS NOT KEPT PACE WITH MONETARY CHANGE

Inflationary drift is not new. The issue was discussed at the Constitutional Convention with respect to judicial salaries. The framers escaped the problem by leaving the task of periodically setting judges' compensation to Congress. That approach—relying on legislators to regularly update money values in statutes—would predominate for the next century.[172] It appears to have worked reasonably well. The limited monetary values enshrined in statutory law, such as tariffs, were often revisited and the economic dynamics were such that drifts were somewhat self-correcting.

This is no longer true. In the first part of the 20th century two shifts began to place significant strain on the traditional approach. The volume of statutory law exploded with the birth of the administrative state and, roughly contemporaneously, the country experienced a macroeconomic revolution as it began the bumpy transition from a specie-backed currency to a floating one, overseen by a central bank. This transition reduced the dramatic short-term

---

172 *Cf.* COMMITTEE ON GOVERNMENT OPERATIONS, STUDY ON FEDERAL REGULATION 1 (Feb. 1977) ("For close to 100 years Congress chose to exercise the commerce power directly.").

monetary corrections the country had experienced in the 19th century but introduced a trend of long-term inflation that made legislative action the only way to avoid statutory erosion.

The arrival of the new monetary environment was partly obscured by the Depression and war-related price shocks. But eventually, through the painful process of attempting to keep welfare benefits current through ad hoc legislation, Congress realized laws would need to accommodate this new environment. By the 1970s and early 1980s Congress began to act, starting with indexing some of the most obviously vulnerable benefits and tax provisions. The next few sections explore that response. Fifty years later it is apparent that Congress has come up short. The full legislative calendar and special interests have kept it from revisiting key provisions and led to inconsistent treatment within the same statutory schemes. It has made staggeringly counterproductive errors in drafting indexation formulas and generally failed to adapt its broad-brush formulas to the complexity of the U.S. Code. Most critically, it has been unable to strike the right balance between administrative flexibility and Congressional oversight of key policy decisions.

A. *The traditional approach to drift emerged during a period of long-run price stability and greater legislative agility*

Monetary fluctuation has been a fact of life since before the founding of the United States.[173] The motivation for the Constitution was in large part driven by the unstable monetary environment in the United States in the 1780s.[174] And in almost every decade since the country experienced some sort of monetary price shock.[175] The impact of these price shifts was never lost on Congress. Some of the most heated political battles of the 19th century—the debate over the Second Bank of the United States, the financing of the Civil War, and the battles over free silver—were about the money supply.[176]

---

173 *See, e.g.*, *Letters of Delegates to Congress: Volume 7 May 1, 1777 - September 18, 1777, Henry Marchant to Nicholas Cooke* (Aug. 18, 1777) (discussing impact of inflation on efforts to pay for revolution); Charles W. Calomiris, *Institutional Failure, Monetary Scarcity, and the Depreciation of the Continental*, 48 J. ECON. HIST. 47, 54–60 (1988).

174 Farley Grubb, *The US Constitution and monetary powers: an analysis of the 1787 constitutional convention and the constitutional transformation of the US monetary system*, 13 FIN. HIST. REV. 43, 50 (2006).

175 *See* Michael D. Bordo, *The Classical Gold Standard: Some Lessons for Today*, FED. RES. BANK OF ST. LOUIS REV. 11–13 (1981).

176 *See* RICHARD BENSEL, THE POLITICAL ECONOMY OF AMERICAN INDUSTRIALIZATION: 1877-1900 394 (1949) (discussing free silver debates); Leon M. Schur, *The Second Bank of the United States and the Inflation after the War of 1812*, 68 J. POL. ECON. 118, 119 (1960); Julliard v. Greenman, 110 U.S. 421, 422–24 (1884) (involving controversy around monetary impacts of greenbacks, the primary method of Civil War finance).

However, the historical experience of price shocks was different from our modern experience in a crucial way: until the 1930s price shocks were largely short term dynamics.[177] Sharp deflationary and inflationary periods accompanied wars and panics, but these did not dramatically move long term prices.[178] According to the best historical data, $100 in 1840 would still be worth roughly $97 in 1880.[179] The average annual rate of change during that period was roughly 0.15%.[180] Comparatively, $100 dollars in 1940 would be worth $588 by 1980, an annual rate of change of roughly 15%.[181] Figure 1, from an analysis by Carmen Reinhart and Kenneth Rogoff, demonstrates the remarkably different character of changes in the price level after 1930.

---

177 Stephen B. Reed, *One hundred years of price change: the Consumer Price Index and the American inflation experience*, BUREAU LAB. STAT.: MONTHLY LABOR REV. (2014) ("Most living Americans have essentially known nothing but inflation. . . . However, before World War II the experience of price change was very different. Prices zigged and zagged rather than following a consistent upward course."). *See also* Williams v. United States, 122 S. Ct. 1221, 1227 (2002) (mem.) (Breyer, J., dissenting from denial of certiorari) (describing the inflationary erosion of judicial salaries as "a phenomenon familiar to the Nation's founders, but absent during much of the 19th century").

178 Allan H. Meltzer & Saranna Robinson, *Stability Under the Gold Standard in Practice*, *in.* MONEY, HISTORY, AND INTERNATIONAL FINANCE: ESSAYS IN HONOR OF ANNA J. SCHWARTZ 163, 164 (Michael D. Bordo ed., 1987).

179 *Historical CPI-U Data*, *supra* note 8.

180 Michael D. Bordo, *The Classical Gold Standard: Some Lessons for Today*, FED. RES. BANK OF ST. LOUIS REV., 8–10 (1981) (finding prices declining on average only 0.14% annually between 1834 and 1913).

181 *Historical CPI-U Data*, *supra* note 8.

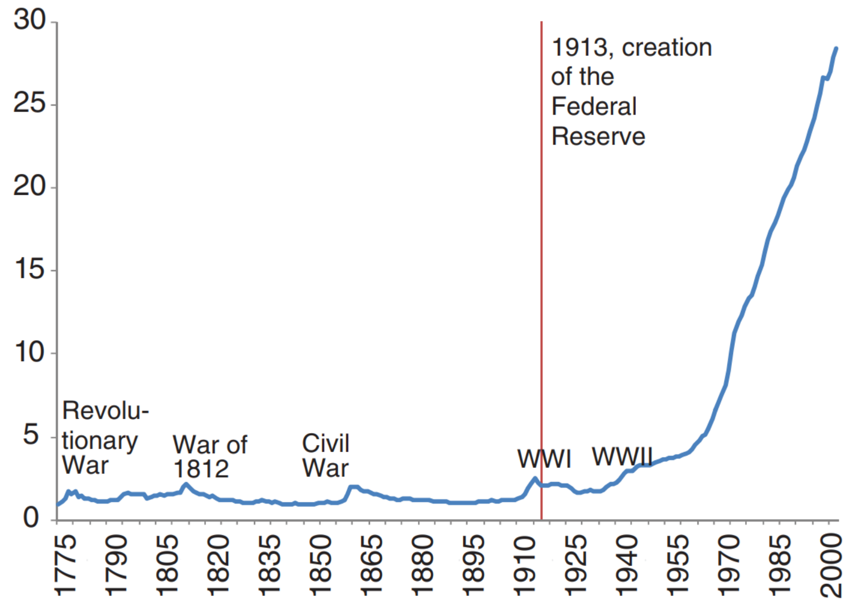


Figure 1: Consumer Price Index 1775-2012[182]

This period of stability was largely due to the makeup of the money supply. Prior to 1933 the United States mostly used specie-backed money, except for brief hiatuses in wartime.[183] This kept prices stable in the long run, with intermittent periods of rapid inflation associated with war or the suspension of convertibility of paper money into coin followed by deflation as ordinary life resumed.[184] Because prices during this period tended to revert to the mean, inflationary drift was somewhat self-correcting.[185]

To be sure, short-term fluctuations in the monetary and economic environment still created mismatches between expected and actual impacts of nominally drafted statutes. Given the limited nature of statutory law during this period, the effect is mostly observable through taxes. The excise tax

---

182 Carmen M. Reinhart & Kenneth S. Rogoff, *Shifting Mandates: The Federal Reserve's First Centennial*, 103 AM. ECON. REV. 48, 48 (2013).

183 Michael D. Bordo et al., *Aggregate Price Shocks and Financial Instability: An Historical Analysis* 8 (Nat'l Bureau of Econ. Rsch., Working Paper No. 7652, 2000); Fernando M. Martin, *A Short History of Prices, Inflation since the Founding of the U.S.*, FED. RES. BANK ST. LOUIS (July 25, 2017), https://www.stlouisfed.org/publications/regional-economist/second-quarter-2017/a-short-history-of-prices-inflation-since-founding-of-us.

184 Fernando M. Martin, A Short History of Prices, Inflation since the Founding of the U.S., FED. RES. BANK ST. LOUIS (July 25, 2017), https://www.stlouisfed.org/publications/regional-economist/second-quarter-2017/a-short-history-of-prices-inflation-since-founding-of-us; *see also* Michael D. Bordo, *The Classical Gold Standard: Some Lessons for Today*, FED. RES. BANK OF ST. LOUIS REV. 2, 8–10 (1981).

185 *See* Meltzer & Robinson, *supra* note 178, at 164–67.

giving rise to the Whiskey Rebellion was stated nominally.[186] Thus its already unpopular burden was worsened by the scarcity-driven deflation concurrent with its enactment.[187] Import duties were stated in both nominal and percentage terms under various tariff acts throughout the period.[188] To the extent commodity prices fluctuated during the time duties were nominally stated, their real impact also shifted.[189] The nominally-fixed taxes of farmers generated enormous hardship during the deflation of the 1890s.[190]

These shifts, however, would not have induced lasting expectation mismatches, nor do they seem to have evinced much concern among legislators for more resilient drafting mechanisms. The political discussion of inflation and deflation, while frequent,[191] was focused on fiscal and monetary policies to combat the effects of short-run volatility.[192] Discussion of long-term price movements, to the extent it shows up at all in economic reform debates, is mostly centered on the hypothesized effects on trade and financial competitiveness.[193] The lack of concern with long-term drift of statutes shows up most starkly when one considers how Congress dealt with member salaries. There were no adjustments made to member compensation between 1818 and 1856.[194] Salaries were raised after the Civil War to $7,500 in 1871, reduced to $5,000 annually in 1874, and then left alone until they returned to $7,500 in 1907.[195] In comparison, Congress passed compensation bills multiple times

---

186 Act of March 3, 1791, 1 STAT. 199 at § 1 (1791) (setting duties at nominal amounts between 20 and 40 cents).

187 *See* Jeffrey J. Crow, *The Whiskey Rebellion in North Carolina*, 66 N.C. HIST REV. 1, 13–16 (1989) (describing pre-enactment concerns about the exacerbating effect of money scarcity).

188 *See* F.W. Taussig, *The Tariff, 1830-60*, 2 Q. J. ECON. 314, 324–25 (1888) (describing duties on iron).

189 *Cf. id.* (describing occurrence of converse case of shifts in nominal values of the tariff when percentage rates were applied to changing prices).

190 Rory Kreitner, Money in the 1890s: The Circulation of Politics, Economics, and Law, 1 UC IRVINE L. REV. 975, 978, 986 (2011).

191 *See* RICHARD BENSEL, THE POLITICAL ECONOMY OF AMERICAN INDUSTRIALIZATION: 1877-1900, 355 (1949).

192 *See, e.g.*, James Madison, "Vices of the Political System of the United States" (April 1787) *in* 9 *The Papers of James Madison*, William T. Hutchinson et. al., eds. 349–50 (1975) (discussing uniform federal money system as response to economic crisis of 1780s); Andrew Jackson, Eighth Annual Message to Congress (Dec. 5, 1836) (describing specie circular as solution to excessive issues of state banks); Nicolas Barreyre, *The Politics of Economic Crises: The Panic of 1873, the End of Reconstruction, and the Realignment of America*, 10 J. GILDED AGE & PROGRESSIVE ERA 403, 409–16 (2011) (discussing congressional proposals to combat the effects of the Panic of 1873 with inflationary fiscal tools).

193 *Cf.* Jeffrey A. Frieden, *Monetary Populism in Nineteenth Century America: An Open Economy Interpretation*, 57 J. ECON. HIST. 367, 388 (1997) (characterizing "principal" arguments on both sides of free silver debates as centered on trade).

194 IDA A. BRUDNICK, CONG. RSCH. SERV., RL97-1011, CONGRESSIONAL SALARIES AND ALLOWANCES: IN BRIEF 18 (Dec. 14, 2023).

195 *Id.*

per decade from the 1930s to 1980s, even when such raises might have been unpopular (such as the three raises made between 1933 and 1935).[196]

The agility with which 19th century legislatures enacted and revisited major statutes likely also made resilience measures an afterthought. While there was little need to respond to long-term monetary shifts, Congress regularly passed legislation to adapt to other economic changes. New tariff legislation adjusted import duties in response to shifts in trade conditions at an average pace of once every seven years until the Civil War (in 1789, 1816, 1818, 1820, 1824, 1828, 1832, 1833, and 1846).[197] Congress also possessed a robust capacity for dealing with exigent circumstances through what has come to be known as "disaster legislation."[198] From 1800 to 1900 Congress passed at least forty bills providing funding for victims of unexpected disaster.[199] While it largely declined to use this capacity in response to economic calamities for ideological reasons, the frequency of the practice demonstrates a nimble legislature.[200] The massive petition system, which produced hundreds of private bills annually throughout the 19th century, also reveals a Congress ready to address problems of fit between the law and individual experience.[201]

B. *The 20th century produced long-run inflation that exacerbated statutory erosion*

The last nine decades have been different. The demise of long-run price stability and the explosion of federal statutory law during the 20th century transformed inflationary drift from a short-term issue into a long-term challenge. The first major shift concerned the structure of government. The Progressive and New Deal eras produced a massive increase in regulations and the birth of a number of new complex statutory schemes creating the agencies and codes needed for the modern state.[202] The additional statutes created more

---

196 *Id.*

197 *See* Summary of the History of the United States Tariff Legislation and Trade Agreements Procedure Before the S. Comm. on Finance, 80th Cong. 1–5 (1947) (statement of Oscar B. Ryder, Chairman United States Tariff Commission).

198 *See generally* MICHELLE LANDIS DAUBER, THE SYMPATHETIC STATE: DISASTER RELIEF AND THE ORIGINS OF THE AMERICAN WELFARE STATE (2013).

199 *Id.* at 46.

200 *Cf. id.* at 48–51 (describing shift towards inclination to use relief powers for economic crises beginning with the panic of 1893).

201 *See* Maggie L. McKinley (now Blackhawk), *Petitioning and the Making of the Administrative State*, 127 YALE L. J. 1538, 1573 (2008); Naomi R. Lamoreaux & John Joseph Wallis, *Economic Crisis, General Laws, and the Mid-nineteenth Century Transformation of American Political Economy* 5–6 (Nat'l Bureau of Econ. Rsch., Working Paper No. 27400, 2020).

202 *See* Susan E. Dudley, *Milestones in the Evolution of the Administrative State*, 150 DAEDALUS 33, 34 (2021); *see, e.g.,* Interstate Commerce Act, 24 Stat. 379 (Feb. 4, 1887), Federal Trade Commission Act, 38 Stat. 717 (Sep. 26, 1914), Securities Act of 1933, 48 Stat. 74 (May 27, 1933), Securities Exchange Act of 1934, 48 Stat. 74 (Jun. 6, 1934).

points of drift-vulnerability because their largely economic focus included numerous money values.

The second shift was economic. When the United States partially suspended the gold standard during World War I, it began a process that would jettison the anchor that made inflationary drift self-correcting.[203] The standard was fully suspended by 1933, and the country never truly returned to it.[204] As a result, the CPI increased 2,978% from 1913 to 2022.[205] This sustained and enormous increase created the potential for unprecedented erosion of dollar amounts in statutes.

The full ramifications of these developments were not immediately apparent.[206] There was some legislative discussion of adapting laws for price fluctuations during the 1930s, but it was ad hoc and focused on dealing with past price changes. The first law to use a cost of living indexing procedure of which the author is aware is the Economy Act of 1933, which instructed the President to reduce the compensation of federal employees based on decreases in the cost of living from 1928 to 1932.[207] The law included a mechanism for future updates but only if the cost of living continued to fall, indicating it was focused on ensuring public employees did not benefit from the Depression and not on accommodating future trends.[208] And even Social Security, one of the most important new programs of the period and which was intended to be long-lasting, was drafted in nominal terms.[209]

It was not until after World War II that the reality of secular price increases began to cement itself in the minds of policymakers.[210] Congressional debates reflect this dawning realization. In the 1950s it began to make regular ad hoc updates to benefit programs to deal with real declines in benefit

---

203 *See* Leland Crabbe, *The International Gold Standard and U.S. Monetary Policy from World War I to the New Deal*, 75 FED. RES. BULLETIN 423, 426 (1989).

204 Kenneth W. Dam, *From the Gold Clause Cases to the Gold Commission: A Half Century of American Monetary Law*, 50 U. CHI. L. REV. 504, 504-505 (1983); *see also* Richard N. Cooper et al., *The Gold Standard: Historical Facts and Future Prospects*, 1982 BROOKINGS PAPERS ECON. ACTIVITY 1, 3-4 (1982). The United States would return to a modified version from 1946 through 1979 but with nothing like the stability of the historical policy. *Id.* at 7. Milton Friedman described this as a pseudo-gold standard. *See* Dam at 526 (1983).

205 *Historical CPI-U Data*, *supra* note 8.

206 *See* Missouri ex rel. Sw. Bell Tel. Co. v. Pub. Serv. Comm'n, 262 U.S. 276, 303 n.16 (1923) (Brandeis, J., concurring) (expecting, mistakenly, that prices following WWI would revert to historical patterns: "peak price levels were practically the same during the War of 1812, the Civil War, and the World War, and it shows that practically continuous declines, for about 30 years, followed the first two wars. The experience after the third may be similar.").

207 Pub. L. 73-2 II. §§ 2–3 (1933).

208 *See id.* at § 3(b).

209 *See infra* Section 0.

210 *See* Weaver, *supra* note 52, at 55 (describing 1949 as the beginning of the first wave of commodity indexation efforts).

values.[211] And by 1958 it started to include future resiliency measures, adding the first automatic cost of living adjustment to veterans' retirement benefits.[212] In 1962 it authorized indexing of civil service retirement benefits as well.[213] The inflationary crises of the 1970s only heightened awareness of the negative consequences of price shifts.[214] At that time Congress began to focus on updating penalties as well as benefit programs.[215] Since the mid-1980s, legislative enactments have seen increasing use of indexation provisions and authorization of agencies to update for inflation.[216]

Two reasons likely explain the three decades between the shift to a floating currency and the beginnings of Congressional efforts to address the statutory effects. First, the economic confusion accompanying two world wars and the Great Depression during the first portion of this new inflationary period obscured the underlying economic trends that made indexation necessary. To observers in the first half of the 20th century, the combination of major inflationary and deflationary periods would have been difficult to distinguish from prior short-run price fluctuations during economic crises and wartime.[217] Surely no one thought the price conditions surrounding the Great Depression and both world wars were business as usual. It was only after a decade or so of regular peacetime inflation that the new reality would have been apparent. A second reason is the recent vintage of reliable price indicators. The Bureau of Labor Statistics was created in 1884, but it did not begin publishing a rudimentary cost of living index until 1905.[218] The original Consumer Price Index was launched in 1913, and time series data were not published until 1921.[219] Thus, legislators likely lacked familiarity with the tools needed to contextualize and address drift until mid-century.[220]

Nevertheless, as price increases began to span decades rather than years, it became clear that something had to be done. The next subsection explores efforts to develop a workable response.

---

211 *See* Section 0 (discussing Social Security).

212 *See* STAFF DATA WITH RESPECT TO H.R. 17550, *infra* note 232, at 14.

213 *Id.*; *see also* CONG. RSCH. SERV., INDEXATION OF FEDERAL PROGRAMS 7 (1981).

214 *See* Adam Clymer, *40% in Survey Say Inflation is Major Issue for 1980 Race*, N.Y. TIMES (Oct. 19, 1979).

215 *See infra* Section 0.

216 *Id.*

217 *See* Stephen B. Reed, *One hundred years of price change: the Consumer Price Index and the American inflation experience*, 137 MONTHLY LAB. REV., 1, 16 (2014) (describing 1950s as "turning point" in American inflation experience); Reinhart & Rogoff, *supra* note 182, at 48 ("It is probable that in 1913, while financial panics were not uncommon, high inflation was still largely seen by the founders of the Fed as a relatively rare phenomenon associated with wars and their aftermath.").

218 Darren Rippy, *The first hundred years of the Consumer Price Index: a methodological and political* history, 137 MONTHLY LAB. REV., 1, 1-2 (2014).

219 *Id.*

220 *See* Hugh Rockoff, *On the Controversies Behind the Federal Origins of Economic Statistics*, 33 J. ECON. PERSPECTIVES 147, 152 (2019).

C. *Congress began to address increasing drift in the 1970s but with only limited success*

1. Social security: ad hoc updating proves unworkable

The first major effort to address the inflationary erosion of a statutory scheme involved Social Security.[221] Old age benefits were described in essentially nominal terms in the Social Security Act of 1935. Benefits were set at certain percentages of eligible wages earned after 1936 and capped at $85 a month;[222] thus as prices rose during and following World War II, the real value of the historical wages from which benefits were calculated fell. The real value of the cap also declined. In 1950 Congress acknowledged that benefit levels had become inadequate due to rising prices, and it enacted old age benefit increases ranging from 40-50%.[223] Of course this one-off adjustment was soon eroded, and Congress had to make another increase in 1952.[224] That increase also quickly proved insufficient, and further adjustments were need in 1954, 1959, 1965, 1968, 1970, and 1971.[225] Each ad hoc adjustment was a major political battle where a seemingly technical correction became a vehicle for battles over the underlying program. Fiscal hawks attempted to ensure the updates lagged inflation or resisted them entirely. [226] While advocates of broader social support attempted to increase real benefits under the guise of updating for inflation. [227]

Starting in the late 1960s both political parties began to grow weary of this political combat masquerading as updating.[228] Substantively, both sides recognized the process had veered uncomfortably far from its stated purpose, with increases well in excess of changes in cost of living occurring in

221 There were earlier efforts to adjust federal wages for past price changes. *See supra* note 207 and accompanying text. And Congress had instituted efforts to keep certain agricultural commodity prices at real parity in the 1930s. Chen, *supra* note 4, at 1399–41. But these earlier efforts were attempts to address contract terms or past erosion.

222 Social Security Act of 1935 § 202(a)-(b).

223 S. Rep. No. 1669, at 20 (1950) ("There are compelling social and economic reasons for liberalizing benefits . . . [beneficiaries] need benefits which are revised to take into account that the 1939 benefit formula proved to be inadequate soon after its enactment and that prices have risen since then.").

224 Wilbur J. Cohen, *Social Security Act Amendments of 1952*, SOC. SEC. BULLETIN 3 (1952) ("The rapid rise in wages and prices during the last few years makes immediate benefit adjustments imperative.").

225 PAUL S. DAVIES & TAMAR B. BRESLAUER, CONG. RSCH. SERV., R94803, SOCIAL SECURITY: COST-OF-LIVING ADJUSTMENTS 9 (2023).

226 Nancy Altman & Ted Marmor, *Social Security from the Great Society to 1980: Further Expansion and Rekindled Controversy*, *in* CONSERVATISM AND AMERICAN POLITICAL DEVELOPMENT 162 (Brian J. Glenn et al. eds., 2011).

227 *Id.* at 163; *see also* Interview with Robert Ball, former Commissioner of Social Security (May 1, 2001) https://www.ssa.gov/history/orals/ball4.html.

228 Altman & Marmor, *supra* note 226, at 162.

democratic and republican administrations.[229] And procedurally the updates were deeply flawed. There was growing concern that updates inevitably became "Christmas tree bills", with gifts to special interests riding alongside substantive provisions.[230] Even when cost of living updates were passed individually without pork, they were generally attached to veto-proof tax and debt bills, providing little in the way of democratic approval.[231]

Several proposals were made to institute an automatic update to keep pace with inflation and take politics out of updating.[232] Finally, in 1972 Senator Frank Church introduced a rider to a debt extension bill that created the modern CPI-based cost of living adjustment mechanism for old age benefits.[233] According to Senator Church, the intent of this provision was to ensure benefits kept pace with inflation without the accompanying political battles.[234] President Nixon agreed on the purpose in his remarks upon signing the bill.[235] The procedure adopted was simple and inflexible. The annual update is calculated mechanically based on the CPI-U (or, in special circumstances an alternate but no less mechanically calculated wage index).[236] The only exceptions for when an update will not be made are if a legislative increase was made the prior year or the update would decrease benefit values.[237] The administrator has no discretion to avoid an update and there is no special legislative override other than new legislation.

This shift from ad hoc to automatic adjustments implies a growing belief in Congress that long-run inflation was a reality it needed to face. It also demonstrates the practical difficulties of ad hoc updating. The need to pass a cost of living update every few years taxed even the series of congresses in the 1950s and 60s known for passing complex social legislation. Most importantly, it illustrates the susceptibility of ad hoc updating to political capture. Opening up any provision of a statute to amendment, even a technical one, invites more controversial changes. And the tactics needed to avoid

---

229 MARTHA DERTHICK, POLICYMAKING FOR SOCIAL SECURITY 280, 343–46 (1979).

230 *Id.* at 41–42.

231 *Id.* at 346.

232 *See, e.g.*, Staff of S. Comm. on Finance, 91st Cong., Rep. on Data with Respect to H.R. 17550 Social Security Amendments of 1970 2 (1970) (discussing these motivations for a 1970 bill).

233 SOC. SEC. ADMIN., *1972 Cola Adjustments*, https://www.ssa.gov/history/tally1972b.html.

234 Frank Church, Letter to the Editor, N.Y. TIMES (Aug. 17, 1972) ("My proposal would . . . authorize cost-of-living adjustments to protect low-income older Americans from the cruel effects of inflation.").

235 Richard M. Nixon, Statement on Signing a Bill Extending Temporary Ceiling on National Debt and Increasing Social Security Benefits (July 1, 1972) ("This action constitutes a major break-through for older Americans, for it says at last that inflation-proof social security benefits are theirs as a matter of right, and not as something which must be temporarily won over and over again from each succeeding Congress.").

236 *See* 42 U.S.C. § 415(i).

237 *Id.*

opening this can of worms, such as attaching updates to unrelated omnibus bills, are undemocratic.

Congress absorbed many of these lessons and shifted to an automatic indexing strategy for many benefit programs in the 1970s.[238] The results have been imperfect. On the one hand, indexation may have been the only way for benefit programs to survive. The repeated ad hoc updating of the 1950s and 1960s was already breaking down in 1972, even before the great inflation at the end of the decade. On the other hand, the formulaic approach has drawbacks. Most obviously, it has locked in an escalating fiscal impact.[239] Costs have ballooned in a way that may not be consistent with democratic preferences, but indexing is an update first and ask hard policy questions later procedure.[240] Only Congress can apply the brakes. Another drawback is that the CPI-U—which was chosen with limited reflection in 1972—may be introducing error into the updates.[241] Congress has acknowledged that the relatively narrow basket of goods used to construct the CPI-U may not reflect the basket of goods that seniors who receive old age benefits typically buy.[242] But given the rigid adjustment procedure, no remedial action can be taken without legislation.

2. Taxes: tradeoffs between legislative control and coherence in updating complex schemes

The next major battle over updating involved the tax code. As with social security, Congress would attempt an indexing strategy, but factors not present in the benefit context complicated matters. The tax code was older than social welfare legislation and more complex. It relied on a mess of interconnected provisions, with many different nominal thresholds, to calculate a single person's tax liability. Welfare benefits generally operate more independently from each other. The politics were also different. Welfare indexation had a large and vocal base of supporters whose crucial benefits were universally threatened by inflation. This allowed for more sustained action even in the face of political opposition. Conversely, public opinion split on tax indexation as taxpayers staying within a bracket saw their taxes become less burdensome under inflation. Consequently, the indexation of taxes was piecemeal and never fully achieved. In one sense this incompleteness is

---

238 *See* DAWN NUSCHLER, CONG. RSCH. SERV., R42000, INFLATION-INDEXING ELEMENTS IN FEDERAL ENTITLEMENT PROGRAMS 6 (2013).

239 *Id.* at *1.

240 *See* B. Guy Peters & Donald J. Savoie, *Managing Incoherence: the Coordination and Empowerment Conundrum*, 56 PUB. ADMIN. REV. 281, 286 (1996) (Rigid social security indexation formulas "provide simple solutions for complex problems, and substitute algorithms for thought."); Daniel Hemel, *Indexing, Unchained*, 83 L. & CONTEMP. PROBLEMS 83, 85 (2020) ("How Social Security benefits ought to change year to year . . . are not questions of measurement. They are, instead, value judgments.").

241 Jim Chen, *The Price of Macroeconomic Imprecision*, *supra* note 4, at 1405.

242 *Id.* at 1406.

desirable. It represents the political process controlling indexation, avoiding the lock-in problem encountered in social security updating. But it also introduced an incoherently divergent approach to income in the tax system that creates distortions no reasonable drafter would have wanted.

When the modern federal income tax was introduced in 1913, items were dealt with in nominal terms.[243] While tax rates were described as percentages the thresholds which divided the various income brackets were set nominally.[244] This caused bracket creep, a process whereby taxpayers near the top of tax brackets were forced to pay higher rates during periods of inflation as they hurdled bracket thresholds despite no change in real incomes.[245] When the standard deduction was introduced it was similarly stated in nominal terms.[246] The alternative minimum tax was also nominally stated at inception.[247] Over time, tax brackets (1981), the personal exemption (1981), the standard deduction (1986), and the alternative minimum tax (2013) have been indexed to inflation.[248] However, important portions of the code remain susceptible, such as the basis for capital gains and the interest deduction.[249]

Lawmakers first began to consider doing something about the drifting tax code during the inflation of the late 1970s.[250] Increasing nominal incomes meant some families were pushed into higher brackets without any uptick in real income.[251] The last half of the decade saw repeated failures to index

243 *See* Strnad, *supra* note 120, at 245 ("Since its inception, U.S. tax law has measured "taxable income" and other tax accounting quantities on a nominal basis"). It is worth noting that the concept of income itself was not defined as nominal or real from the inception of the tax code. Charlotte Crane, *The Income Tax and the Burden of Perfection*, 100 NW. L. REV. 171, 179–82 (2006). Thus a nominally-based tax was not inevitable given the text of the statute.

244 Tracey M. Roberts, *Brackets: A Historical Perspective*, 108 NW. U. L. REV. 925, 929 (2014).

245 *See* Burkhard Heer and Bernd Süssmuth, *Tax Bracket Creep and its Effects on Income Distribution*, 38 J. MACROECONOMICS 393, 393 (2013) (defining the phenomenon).

246 Strnad, *supra* note 120, at 240–42.

247 *See* Richard Sousa, *Bracket Creep*, HOOVER DIGEST (Oct. 19, 2007) (criticizing nominal AMT).

248 *See* Reed Shuldiner, *Indexing the Tax Code*, 48 TAX L. REV. 537, 548–50 (1993); American Taxpayer Relief Act of 2012, § 104, Pub. L. 112-240, 126 Stat. 2313 (2012) (amending AMT to include inflation adjustment); TAX POLICY CENTER, *What is the AMT?*, https://www.taxpolicycenter.org/briefing-book/what-amt.

249 *See* Daniel Hemel & David Kamin, *The False Promise of Presidential Indexation*, 36 YALE J. ON REG. 693, 694 (2019).

250 *See* Weaver, *supra* note 52, at 195–96. The earliest indexed provisions were the limits on qualified retirement plans enacted in 1974. Richard J. Kovach, *Technical and Policy Standards for Inflation Adjustments Under the Internal Revenue Code*, 33 OKLA. CITY U. L. REV. 603, 605–06 (2008). However, these were enacted under ERISA, not specific tax legislation. *Id.* Thus the income bracket indexation is the first example of tax-focused indexation.

251 *See* S. Rep. 97-144, at 11–12 (1981); Karen W. Arenson, *Tax Rate Cuts vs. Inflation*, N.Y. TIMES (Sept. 16, 1981); Emmanuel Saez, *The effect of marginal tax rates on income: A panel study of 'bracket creep'*, 3 n. 1 (Nat'l Bureau of Econ. Rsch., Working Paper No.7367, 1999) (describing "substantial" 1979-81 increase as leading to tax revolt). The effect was not universal. For those families that stayed within the same bracket their real tax incidence would have decreased. *Id.* at 1.

brackets in 1975, 1976, 1977, 1978, and 1979.[252] Finally, Republicans mustered a majority in 1981 to enact the Economic Recovery Tax Act (ERTA), which added an indexing provision to annually update income tax bracket thresholds for increases in the CPI-U over the preceding year.[253] It also included a measure to index the personal exemption deduction.[254] According to the Senate committee report, the use of indexing was motivated by the hardships imposed by bracket creep and recognition that ad hoc adjustments had failed to keep pace with inflation.[255]

ERTA passed both the House and Senate by healthy margins,[256] but it was not without controversy. While many supporters saw the measure as one of simple fairness and technocratic governance, others saw profound policy choices at stake.[257] A key criticism concerned interactions with recent and proposed tax cuts also designed to account for inflation.[258] Opponents were concerned that overlapping tax breaks and indexing would lead to repeated overcompensation for inflation, especially in periods of stress where Congress might feel pressure to show action even when automatic adjustments were already working in the background.[259] Critics also worried that revenue losses from indexation might limit other tax credits and programs, though some conservative supporters saw the same revenue declines as means to engender fiscal discipline.[260] This expected disciplining effect probably clinched the vote in the Senate.[261] Despite several subsequent attempts to repeal indexing as a means of closing the budget deficit, the provision stuck,[262] likely because it came to be viewed, despite its fiscal impacts and the political motivations behind its passage, as a fairness-driven effort to protect taxpayers against unexpected increases in their liability.[263]

This splintered debate reveals the importance of reliance interests in indexing old laws. Failing to index means regulated persons, in this case taxpayers, will be at the mercy of variable laws. On the other hand, indexing

---

[252] Weaver, *supra* note 52, at 195–96.

[253] Weaver, *supra* note 52, at 201–03; Roberts, *supra* note 244, at 937; Economic Recovery Tax Act of 1981 (ERTA), § 104, Pub. L. No. 97-34, § 101, 95 Stat. 172, 179 (1981) (adding subsection (f) to 26 U.S.C. § 1).

[254] Economic Recovery Tax Act of 1981, § 104(c).

[255] S. Rep. 97-144, 118-19 (1981).

[256] *See* 97 CONG. REC. 19329, 19538 (1981).

[257] Weaver, *supra* note 52, at 200–01.

[258] *Id.*

[259] *See id.* at 200.

[260] *Id.* at 201; MONICA PRASAD, STARVING THE BEAST: RONALD REAGAN AND THE TAX CUT REVOLUTION 114 (2018) (explaining Republican theories that indexing would force Congress to take responsibility for otherwise hidden tax increases caused by inflation).

[261] Weaver, *supra* note 52, at 201–03.

[262] Weaver, *supra* note 52, at 205–07; Prasad, *supra* note 260, at 122.

[263] *See* Ronald Reagan, Remarks on Tax Reform to Concerned Citizens (May 29, 1985) ("By . . . indexing [credits] for inflation . . . we can make sure that the working families do not suffer under the burden of Federal taxation."). *Cf.* Shuldiner, *supra* note 18, at 547 n. 37 (describing the view that indexation protects vulnerable taxpayers).

harms those benefiting from the status quo. In the income tax case that was the public purse, which gained revenue from bracket creep. Moreover, indexing has important interaction effects. It can frustrate the desired function of related enactments, such as tax cuts that presume a certain amount of inflationary erosion.

The regular battles between these competing interests seem mostly to be decided by inertia. The multiple failed attempts to pass indexing show the difficulty of changing course, finally succeeding only after inflation hit record highs. The resiliency of indexing in the face of several repeal attempts further demonstrates the point. Important fiscal concerns were not enough to remove what had grown to be considered a mechanism of good governance.[264]

The momentum surrounding indexation continued into the next round of tax reform in 1986.[265] In the following years numerous other provisions were indexed, such as caps on contributions to Roth IRAs, certain deductions, and a variety of tax credits.[266] But this effort was incomplete. Despite pressure for further indexing, Congress failed to index a variety of other provisions, including the interest deduction and the basis for capital gains.[267]

Things stalled in the early 1990s with the failure to index the basis used in calculating capital gains for inflation. As previously explained, the appreciation (gain) on capital assets is calculated as the price upon a realization event (typically a sale) minus the basis (defined in statute as the "cost of such property").[268] Inflationary drift enters the equation through the basis because the cost at time of purchase is stated nominally, as is the price upon realization. Thus, the total gain can include increases that are due merely to changes in the value of the dollar rather than real appreciation of the asset. Put simply, the basis drifts with inflation, introducing error into the measurement of real income from the sale of the asset.

Supporters of indexation argue that it is unfair to tax these inflation-driven gains because it could impose tax even when the taxpayer suffers a real loss.[269] They also point out that the lack of indexation distorts the decision to save or consume because it untethers tax liability from real investment income.[270] Opponents point out that the preferential tax rates and delayed

---

264 *Cf.* Louis Kaplow, *Regional Cost-of-Living Adjustments in Tax-Transfer Schemes*, 1 (Nat'l Bureau of Econ. Rsch., Working Paper No. 5008, 1995) (providing view of prominent scholar in the field a decade later that "it is generally accepted that changes in the cost of living should be reflected in the tax system").

265 *See* Tax Reform Act of 1986, § 104, Pub. L. 99-514, 100 Stat. 2085 (Oct. 22, 1986). *See also* Jared Walczak, *Inflation Adjusting State Tax Codes: A Primer*, TAX FOUND. (Oct. 29, 2019) (describing momentum around state adoption of indexation during same period), https://taxfoundation.org/inflation-adjusting-state-tax-codes/.

266 *See* Kovach, *supra* note 250, at 606–07.

267 *See* Kovach, *supra* note 250, at 604 (collecting the numerous inconsistencies).

268 I.R.C. §§ 1001(a), 1011(a), 1012(a).

269 *See* Shuldiner, *supra* note 18, at 549 (neutrally describing this impact).

270 Hemel & Kamin, *supra* note 249, at 705.

realization applied to capital gain income more than account for the impact of inflation.[271]

Despite strong lobbying by indexation supporters, indexation had not been included in earlier reform bills. ERTA had not indexed capital gains in 1981 because of double counting concerns: it already dropped the top tax rate on capital gains by 20 percentage points, and at the time 60% of such gains could be excluded from income.[272] Then in 1986 Congress temporarily reversed course and removed the preferential treatment for capital income entirely out of a concern that it unfairly advantaged wealthy taxpayers.[273] Preferential treatment was soon reinstated in 1990,[274] but it proved politically impossible for the Bush administration to push through indexation in addition to the reinstatement.[275] This stall was the result of shifting political winds. Memories of the high inflation of the 1970s and 1980s were fading, and fiscal concerns had risen in importance among Republicans.[276] Many representatives, especially Democrats, were also concerned about being seen as giving rich taxpayers a windfall.[277]

Consequently, supporters of indexation looked to an unprecedented strategy—executive indexing. If the definition of cost was deemed ambiguous it could be interpreted by the Treasury to mean cost in real dollars (i.e. current dollars at the time of sale).[278] This argument was textually plausible but otherwise a long shot. In 1939, the Seventh Circuit, relying largely on nominalist principles and the 1792 Coinage Act, rejected a similar argument in *Bates v. United States*.[279] A taxpayer claimed that "cost" meant the amount he paid for securities based on the 1931 dollar, which was worth more than the dollar in 1934 because of FDR's gold redemption suspension.[280] The Seventh Circuit disagreed, holding that the gold value of the dollar was irrelevant because courts and statutes only dealt in units of "lawful money" which had been the same standard (the dollar) throughout the period in question.[281] In 1976 the Tax Court rejected a similar claim—that inflation should be deductible from nominal income—holding "our tax structure is not set up to take

271 *See id.* at 706; Bruce Bartlett, *Indexing Capital Gains by Fiat*, 135 TAX NOTES 883, 883 (2012).

272 Joint Economic Comm., The Economic Effects of Capital Gains Taxation 1 (1997).

273 Suyoung Moon, *Revisiting the 100-Year-Old Debate on the Preferential Treatment of Capital Gains*, 41 ABA TAX TIMES (2021).

274 *Id.*

275 Nathaniel C. Nash, *Here's a Twist: Gains Tax is Rising*, N.Y. TIMES (OCT. 31, 1990) (describing fiscally-driven "compromise" that frustrated presidential agenda for further revision of capital gains).

276 *Cf.* Justin Fox, *The Mostly Forgotten Tax Increases of 1982-1993*, BLOOMBERG (Dec. 15, 2017) (documenting fiscally-motivated correction to the tax cuts of the 1980s that first included indexation).

277 *See* David E. Rosenbaum, *Bush Drops Fight for Lower Tax on Capital Gains*, N.Y. TIMES (Sep. 30, 1990).

278 Hemel & Kamin, *supra* note 249, at 706–09.

279 Bates v. U.S., 108 F.2d 407, 408 (7th Cir. 1939).

280 *Id.*

281 *Id.*

into account the effects of inflation."[282] Nevertheless, proponents believed the recent development of *Chevron* deference would safeguard the interpretation from judicial reversal.[283]

This was the first time an argument for administrative indexing without congressional approval gained traction in the news cycle.[284] The momentum was such that President George H.W. Bush asked Attorney General William Barr to review the legality of the proposal.[285] After consideration, both the DOJ and the Treasury would independently determine unilateral indexation exceeded the Executive's lawful authority.[286] The main reason: the meaning of cost clearly meant the price paid at time of purchase, and thus *Chevron* was inapplicable.[287] Unless Congress had explicitly delegated authority to override the plain meaning of cost, the revisionist cost argument was inconsistent with separation of powers.[288] The DOJ's thorough opinion carried the day and efforts to renew similar arguments failed.[289]

Capital gains indexation was essentially dead in the water from that point on. Two repeated objections to subsequent efforts are worth mentioning. First, some commentators have observed that the rise of the major questions doctrine has made a strained interpretation of cost even less plausible.[290] Second, many observers oppose indexing capital gains in isolation because it might lead to tax arbitrage if other nominal provisions are not indexed simultaneously. A much discussed opportunity involves the business interest deduction, which allows for borrowers to deduct interest payments for non-personal loans.[291] These deductions are calculated, like capital gains, in nominal terms.[292] Interest has an inflationary component (lenders require compensation for the inflation-driven erosion in the value of their principal), so when one deducts interest payments the deduction is slightly larger than the real economic transfer because it accounts for both the true interest expense and an inflation-offset payment.[293] Under the current system taking out a loan to

282 Crossland v. Commissioner, 35 T.C.M. (CCH) 262, *3–*5 (1976).

283 *See* Memorandum from Charles J. Cooper, Michael A. Carvin & Vincent J. Colatriano to Dr. Lawrence A. Hunter, Exec. Vice Pres., Nat'l Chamber Found. 11–12 (Aug. 17, 1992), https://www.atr.org/sites/default/files/assets/shaw%20pittman%20potts.pdf.

284 Other administrative efforts by the U.S.S.C and the Civil Penalties Adjustment Act were explicitly authorized by Congress.

285 *See* Bruce Bartlett, *Indexing Capital Gains by Fiat*, 135 TAX NOTES 883, 884 (2012).

286 *Id.*

287 *See* Legal Authority of the Department of the Treasury to Issue Regulations Indexing Capital Gains for Inflation, 16 Op. O.L.C. 136, 146–51 (1992).

288 *See id.* at 138, 151–52.

289 See Hemel & Kamin, *supra* note 249, at 707–09.

290 *Id.* at 716–18; *but see* Charles J. Cooper & Vincent Colatriano, Law in an Age of Austerity: The Regulatory Authority of the Treasury Department to Index Capital Gains for Inflation: A Sequel, 35 HARV. J. L. PUB. POL. 487, 504–13 (2012) (arguing the argument is still viable).

291 I.R.C. § 163.

292 Hemel & Kamin, *supra* note 249, at 704.

293 John Bossons, *Indexing for Inflation and the Interest Deduction*, 30 WAYNE L. REV. 945, 954–58 (1984).

finance the purchase of the capital asset would give the borrower access to that bonus inflation-driven deduction but the borrower would also be taxed on inflationary gains at realization. Thus the impact of inflation nets out. But if only capital gains were indexed for inflation a person financing their investment would have an advantage over other purchasers because they would get the inflation-related deduction but not pay tax on inflation-related gains.[294]

The arbitrage critique is forceful but could go even further. It leaves out all the other interactions between capital gains and currently indexed provisions. These already create distortions and arbitrage potential. For example, it is well-documented that the failure to index capital income compared to other forms of income distorts individual savings decisions.[295] One estimate puts the costs of these distortions at roughly 1% of GDP per year.[296]

Stepping back, we can draw several lessons from the tax indexation saga. First, it is difficult to index complex schemes wholesale. It took multiple rounds of legislation to index the tax code even partially. Second, the legislature can be shortsighted and fail to account for the interactions in a statutory scheme. Thus even a politically responsive decision not to index capital gains introduced distortions that the political process never intended to produce. Finally, reliance and inertia matter. Once an indexation procedure is selected it can be difficult to adjust or replace legislatively.

When compared to Social Security, the tax experience suggests a qualification to the emerging thesis about how to respond to inflationary drift. While ad hoc legislative updates certainly cannot keep up with drift, indexation is not a panacea. Indexation may remedy the accuracy problems of ad hoc updating—in that it removes political influence over the adjustment amount—but the decision to index in the first place is still subject to political shortsightedness if the legislature fails to recognize the subtle costs of an incoherent approach to a statutory scheme. Delegation to an agency expert in administering the relevant statutes could help, but the tax experience demonstrates how hard it is to accomplish this under the status quo. Separation of powers bars unilateral agency action and the legislative process does not always delegate updating powers of sufficient scope.

---

294 Hemel & Kamin, *supra* note 249, at 704. Hemel and Kamin provide a helpful example: If real interest rates are zero but inflation is 10% the nominal interest paid on a $100 loan will be $10. If one uses the $100 loan to buy a stock, the value of which would also increase by 10% due to inflation, and sells it with a $10 nominal and $0 real gain. In the current system you could deduct the $10 interest expense but you would pay tax on the $10 nominal gain (at 20%). If only capital gains were indexed you could deduct the full $10 but pay no tax on the gain, making you $2 richer than an investor who did not finance the transaction. *Id.*

295 *See* Martin Feldstein, *Capital Income Taxes and the Benefit of Price Stability* 2 (Nat'l Bureau of Econ. Rsch, Working Paper No. 6200, 1997).

296 *Id.*

3. Civil Penalties round one: information gaps and flawed formulas

Roughly contemporaneously with early tax indexation efforts Congress began to grasp that drift could seriously erode financial sanctions.[297] This led Senator Frank Lautenberg to commence a decades-long effort to index monetary penalties to inflation. Congress approached this effort with more trepidation than its prior attempts and rightly so. Even after a decade of planning, the initial mechanism enacted in 1996 was a catastrophic failure, undermined by errors in the adjustment formula, misconceptions about the complexity of the U.S. code, and an inflexible attitude towards administrative agencies.

a. Early attempts to pass an inflation adjustment act

In 1986, Senator Lautenberg introduced two bills to address the erosion of monetary penalties due to inflation.[298] The Civil Penalties Inflation Adjustment Act would have required agencies to update civil penalties within their authority for changes to the cost of living, as measured by the annualized urban Consumer Price Index (CPI-U), between the year the penalty was last determined and 1986 rounded to the nearest $10.[299] Presumably to avoid unnecessary disruption to very old statutory schemes the first adjustments were capped at 1,000%.[300] Following this initial catch-up adjustment, agencies would be required to update their penalties annually using the same methodology.[301] The bill only contemplated upwards adjustments. Net deflation would not trigger updates.[302] The companion bill, the Federal Criminal Penalties Inflation Adjustment Act, would have required the U.S.S.C. to make similar adjustments of the fines in the federal Sentencing Guidelines but only every four years and rounding to the nearest $100.[303] Because the penalties had been set recently in 1984 there was no cap on the initial adjustment.[304] Both bills required the OMB to track the penalties and report annually on the amounts collected under them and any inflationary adjustments made.[305] Congress failed to act on these bills in 1986 and they were reintroduced with identical language in 1987.[306] They failed to make it out of the Judiciary Committee for uncertain reasons.

---

297 *See* David A. Lopez, *The Great Inflation: A Historical Overview and Lessons Learned*, PAGE ONE ECONOMICS (Oct. 2012).

298 S. 2558, 99th Cong. (1986); S. 2559, 99th Cong. (1986).

299 S. 2559, 99th Cong. § 4–5 (1986).

300 *Id.*

301 *Id.* § 4.

302 *Id.* § 4(c)(1) (Defining the adjustment only as the amount the current CPI "exceeds" the prior year).

303 S. 2558, 99th Cong. § 4 (1986).

304 *Cf.* 132 CONG. REC. S7595 (daily ed. June 16, 1986).

305 S. 2558 § 5; S. 2558 § 7.

306 133 CONG. REC. S5172–74 (daily ed. April 10, 1987).

What is clear is that the discussion around these bills showed Congress how little it knew about the scale of the inflationary drift problem. The legislative history reveals basic gaps in knowledge of even the number of fines and penalties.[307] Only as the legislative process developed did the dramatic scale of the issue come into focus, Congress even discovered the rather striking fact that some money penalties had not been updated since 1793.[308]

Sen. Lautenberg appeared to have two considerations in mind in sponsoring the bill—the recent high inflation of the early 1980s and the risk that inflationary erosion would undermine Congress' efforts be tough on crime, including corporate crime.[309] More broadly, he framed the issue to the Senate as a concern about Congressional intent being undermined.[310] He also mentioned revenue gains from increased fines in his floor speeches.[311]

The criminal penalty adjustment bill was never reintroduced, but a watered-down version of the civil penalty bill reemerged in 1989.[312] The bill was amended in response to DOJ criticism that automatic and broad-brush adjustments could lead to unfair impacts.[313] The revised bill merely directed the OMB to calculate what adjustments should be made to all civil money penalties to keep pace with inflation but contemplated individual legislation to actually implement the adjustments.[314] It also inserted a more elaborate six-tiered rounding scheme for the OMB to use in place of the nearest $10 language.[315] The bill was enacted into law on October 5th, 1990.[316]

Congress refrained from direct action until it received more information from the OMB, which reported that there were almost 1,000 unindexed penalties across the U.S. code.[317] Even with this clear evidence of the need for action, Congress still struggled with how to implement the adjustments,

307 *See, e.g.*, 136 CONG. REC. S1494 (daily ed. Feb, 22, 1990) (Sen. Lautenberg) ("Unfortunately, the OMB maintains no detailed central account that tracks penalty assessments and collections and matches them with the laws under which they are imposed. There is no accounting of which laws need updating the most. Apparently, hundreds of millions of dollars is seen as small change that is not worth watching more closely.").

308 *Id.* at 1494. *See also* H.R. REP. NO. 101–697, at 2 (1990) ("40 percent of [civil] penalties would need to be increased in amounts greater than $1000 (ranging up to greater than $1 million) in order to account for inflation").

309 *See* 132 CONG. REC. S7594–95 (daily ed. June 16, 1986) ("By its inaction, Congress each year pulls the punch of penalties for a variety of transgressions.").

310 *Id.*

311 *Id.*

312 H.R. REP. NO. 101–697, at 3 (1990).

313 *Id.*

314 136 CONG. REC. 1494 (daily ed. Feb. 12, 1990) (Sen. Lautenberg).

315 Pub. L. No. 101-410, 104 Stat. 890 (Oct. 5, 1990).

316 *Id*.

317 *See* U.S. GOV'T ACCOUNTABILITY OFF., GAO-03-409, CIVIL PENALTIES AGENCIES UNABLE TO FULLY ADJUST PENALTIES FOR INFLATION UNDER CURRENT LAW 5 (2003) (describing 1990 report findings).

especially if they should be automatic, and rejected a bill that would implement across-the-board, uncapped automatic adjustments in 1993.[318]

b. Congress gets its math wrong

Finally, in 1996 Congress authorized an actual updating procedure as part of the Omnibus Debt Collection Improvement Act of 1996. It applied to civil penalties across all but four statutes—the Internal Revenue Code, the 1930 Tariff Act, OSHA, and the Social Security Act were exempted.[319] The act resolved the debate over whether Congress should update by individual legislation or delegate to agencies by requiring agency heads to make automatic adjustments for inflation every year as well as a catch-up adjustment in 1997.[320] The same procedures applied to all agencies. Updates were to be based on the annual CPI-U calculations outlined in the 1990 Act with a critical exception: *all initial catch-up adjustments were to be capped at 10%.*[321] The legislative history does not explain why this cap was implemented or why the 1,000% cap initially proposed by Senator Lautenberg was reduced so dramatically. It seems likely that the result was a quickly considered political compromise between those who favored an automatic, agency-administered update system and those who worried about the unfairness of dramatic jumps in penalties.[322] It is also possible that in the chaotic final push of getting such a large omnibus bill passed some regulated entity pushed the change to avoid increasing punishments.

Whatever the reason, the change was massively counterproductive. By capping the initial adjustment at 10% the act ensured that values which had been eroded more than 10% in real terms could never catch up to current prices.[323] The impact was especially dramatic for large corporate fines. In 1996 the NTSB's fine relating to failure to provide crashworthiness information to consumers should have been updated from $800,000 (the initial fine set in the 1960s) to $2.45 million be equivalent in real terms.[324] Instead it was updated to $880,000.[325] A report by the GAO in 2003 found two other counterproductive methodological problems resulting in lower than intended adjustments. First, the act's decision to use the CPI-U for June of the year before the adjustment as the measure of current prices meant adjustments effectively lagged real price increases by one year.[326] And second, the

---

318 See James Ming Chen, *Inflation-Based Adjustments in Federal Civil Monetary Penalties*, YALE L. & POL. REV. 1, 18 (2016).

319 GAO-03-409, *supra* note 317, at 33–34.

320 Pub. L. No. 104-134, § 31001(s).

321 *Id.* at § 31001(s)(1)(C)(2).

322 *See* Chen, *supra* note 318, at 19.

323 *Id.* at 19–22.

324 Civil Penalties, 64 Fed. Reg. 37876-77 (July 14, 1999).

325 *Id.*

326 GAO-03-409, *supra* note 317, at 23.

rounding rules meant some fines at the bottom end of their rounding categories were not updated until massive inflationary drift had occurred.[327] For example, under the rounding rules a fine between $101 and $1,000 (inclusive) should have increases rounded to multiples of $100.[328] This means that a $110 dollar fine would not see any increase until cumulative inflation reached 45% (i.e. when the increase exceeded $50). At typical rates of inflation that would take decades. The GAO sampled six agencies and found that because of this rounding problem 90% of their penalties would not be adjusted for four years and 42% for at least ten years.[329] This was particularly problematic for small dollar value fines that were structured to be multiplied by a very large number of violations. For example, the NHTSA's $5.50 penalty for every 0.1 mpg exceeding CAFE fuel economy standards could not be adjusted for 28 years.[330]

As a result of these errors Congress' 1996 attempt to address impact drift did exactly the opposite of what was intended—locking in rather than resolving inflationary drift in many places. What institutional dynamics led to such a miserable performance? The obvious candidate is the use of the omnibus process to enact such a far-reaching and complex piece of legislation. It seems probable that with more consideration a staffer or lobbyist would have spotted the unintended effects of the 10% cap as well as the lag caused by using CPI data from the prior year. If one believes the 10% cap was not inadvertent but intended to prevent a large increase in the real values of penalties then the Congressional tendency to prioritize political rather than technocratic concerns gains explanatory power.

Besides the rushed passage and possible political compromise, a more fundamental limitation on Congress' ability to deal with inflationary drift is hinted at by the rounding errors. Even after a decade of legislative consideration, Congress simply did not understand how its procedure would impact the full variety of penalties covered. It had to paint with a fairly broad brush and had neither the time, expertise, nor the drafting creativity to accommodate hundreds of distinct penalty structures and amounts. Put plainly, perhaps updating over a thousand distinct provisions across radically different regulatory schemes via a single mandatory procedure was a bad idea from the start. Whatever the cause, the ultimate verdict on the 1996 act came from the agencies charged with implementing it: only 9 of 80 covered agencies implemented on time and six years later 20% had yet to implement at all.[331]

---

327 *Id.* at 26–27.
328 *Id.* at 27.
329 *Id.* at 28.
330 *Id.* at 29.
331 GAO-03-409, *supra* note 319, at 2–3.

4. Civil Penalties round two: flexible formulas but oversight challenges

a. Two decades later Congress corrects the procedure

Despite the major unintended consequences of the 1996 act Congress did not revise its updating structure until 2015, when it passed the Federal Civil Penalties Inflation Adjustment Improvements Act ("Improvements Act") as part of the Bipartisan Budget Act of 2015.[332] This retained the basic structure but attempted to address the inflexibility of the prior act in several ways. First, it required rounding to the nearest $1.00 and changed the CPI used to the October CPI, meaning the adjustment would always be based on CPI data from the same federal fiscal year.[333] Second, to correct for the inadvertent lock-in effect it required a new catch-up adjustment to be implemented by 2016.[334] But, perhaps recognizing the potential for outliers among the many penalties affected, Congress provided an escape hatch to the catch up adjustment. If the agency deemed the catch-up to have a negative economic impact or determined the social costs of updating outweighed the benefits it could implement a lower adjustment if the Director of the OMB concurred.[335] These exemptions do not apply to subsequent updates, though agencies are permitted to forgo a subsequent update if they updated the penalty for other reasons the same year and that increase was larger than the required update. It also capped the catch-up adjustment at 150% rather than 10%.[336] The 2015 improvements also subjected penalties under OSHA and the Social Security Act to inflation updates and required the OMB to provide updating guidance every December and the GAO to report annually on the adjustments made.[337] Finally, the improvements exempted adjustments other than the initial adjustment from § 553 of the APA, alleviating agencies from the burden of going through notice and comment.[338]

Implementation of the new provisions began fairly seamlessly. 46 of 52 agencies subjected to the Act published initial catch-up adjustments.[339]

---

332 Pub L. 114-74 (Nov. 2, 2015).

333 *Id.* at § 701(b)(2). Because the OMB publishes updating guidance in December the October numbers are generally the most recent available.

334 *Id.* at § 701(b)(1)(A).

335 *Id.* at § 701(c).

336 *Id.* at § 701(b)(2)(B).

337 *Id.* at § 701(b)(4).

338 *Id.* at § 701(b)(1).

339 U.S. GOV'T ACCOUNTABILITY OFF., GAO-17-634, CERTAIN FEDERAL AGENCIES NEED TO IMPROVE EFFORTS TO COMPLY WITH INFLATION ADJUSTMENT REQUIREMENTS 6 (2017). Three agencies did not adjust because they believed they were not subject to the act. *Id.*

Nearly universal compliance continued with all but a handful of agencies publishing updates annually through 2020.[340]

b. The Trump Administration attempts to evade the revised procedure

One notable drama marred this otherwise solid record—the NHTSA's delay of its catch-up adjustment to penalties associated with automobile fleet fuel economy standards. Ultimately this controversy would wind up in front of the Second Circuit and raise questions about conflict between the Improvement Act and executive power. In 2016, as with other agencies, the NHTSA published its initial catch-up to the penalty imposed on automakers for failing to comply with fleet fuel economy standards (CAFE standards). The penalty had been set for well over a decade at a $5.50 fee multiplied by the tenths of a mile per gallon an automaker's fleet exceeded a fuel economy target multiplied by the number of cars in the fleet.[341] While rarely imposed, the penalty could reach into the millions for a large manufacturer and is central to robust emissions credit trading.[342] The NHTSA proposed to update the penalty to $14, the maximum allowed under the 150% cap.[343] However, after lobbying by the auto industry, the agency determined that the adjustment should not be applied until 2019 to avoid unfair retroactivity but declined to reduce the adjustment because of economic harm.[344] Then, following the inauguration of President Trump, the agency received a communication from the White House instructing it to delay the final rulemaking.[345] It did so temporarily and then indefinitely suspended the update via final rule.[346] That indefinite suspension was challenged as beyond NHTSA's statutory authority and ultimately vacated by the Second Circuit, which held that the Inflation Adjustment Improvements Act did not grant agencies the discretion to delay adjustments indefinitely.[347]

---

340 U.S. GOV'T ACCOUNTABILITY OFF., GAO-19-567R, CIVIL MONETARY PENALTIES: REVIEW OF FEDERAL AGENCIES' COMPLIANCE WITH THE 2018 ANNUAL INFLATION ADJUSTMENT REQUIREMENTS 3 (2019); U.S. GOV'T ACCOUNTABILITY OFF., GAO-21-488R, CIVIL MONETARY PENALTIES: REVIEW OF FEDERAL AGENCIES' COMPLIANCE WITH THE 2020 ANNUAL INFLATION ADJUSTMENT REQUIREMENTS 3 (2021).

341 NHTSA Final Rule, 81 Fed Reg. 95489, 95489 (Dec. 28, 2016). Prior to the 1996 inflation adjustment the penalty had been set at $5.00 for decades.

342 *Id.*; *see also* ICCT, Briefing: Credit Trading in the US Corporate Average Fuel Economy (CAFE) Standard (2014), https://theicct.org/sites/default/files/publications/ICCTbriefing_CAFE-credits_20140307.pdf.

343 81 Fed. Reg. 95489.

344 *Id.* at 95490.

345 Delay of Effective Date of NHTSA Final Rule,82 Fed. Reg. 8694, 8694 (Jan. 30, 2017).

346 Indefinite Delay of NHTSA Final Rule, 82 Fed. Reg. 32139, 32139 (July 12, 2017).

347 NRDC v. NHTSA, 894 F.3d 95, 109 (2d Cir. 2018).

But the NHTSA was not finished. It engaged in a new rulemaking in hope of finding that either the CAFE penalties were not civil money penalties—in contravention of its consistent position since 1996—or that there was a negative economic impact which would justify lowering the catch-up adjustment.[348] It had the support of the OMB in this effort.[349] This was again challenged and the Second Circuit again vacated the NHTSA's rulemaking and ordered the $14 penalty to take effect immediately, holding that the CAFE provision was a civil penalty but not reaching the substance of the negative economic impact question.[350]

c. Lessons from the civil penalties saga

What does this convoluted tale tell us about inflation adjustment? In one sense it is a victory; the clear provisions of the 2015 Improvement Act allowed a court to enforce Congress' vision of economically current fines in the face of what has been described as a special-interest driven attempt to throw a bone to political allies.[351] The heavy fight put up by the auto industry against the increase also perhaps shows the magnitude of the unintended benefit drift had provided.

But the case also points to weaknesses in the current scheme. The Second Circuit's ruling that agencies lacked authority to forestall updates may be concerning if one believes future updates will cause real economic harm, especially in periods of higher inflation. And the interaction between the White House and the NHTSA makes one wonder about the usefulness of the requirement that the OMB concur in the use of a negative economic impact exemption. Did this really provide a check on arbitrary agency action? The concern may be theoretical now that all catch-up adjustments have been made, but it demonstrates the vulnerabilities of a flexible administrative updating system to interest group capture. The energy it took to prevent the political manipulation of the updating procedure was enormous, involving high-powered NGOs, plenty of press, and two appellate panels. And the efforts of the NHTSA to forestall the update were neither subtle nor particularly skillful,[352] but they still resulted in an almost three-year delay. It is easy to imagine a more competent effort utilizing a negative economic impact clause to avoid future updates with little accountability. When one compares the 2015 and 1996 efforts a clear tradeoff in updating emerges: flexibility versus

---

348 New York v. NHTSA, 974 F.3d 87, 90 (2d Cir. 2020).

349 *Id.*

350 *Id.* at 101.

351 Rebecca Beitsch, *Court for second time strikes down Trump admin rollback of automaker penalties*, THE HILL (Aug. 31, 2022) (quoting the Sierra club "The Trump Administration cannot give away polluting passes to automakers who lag behind on meeting standards required by law").

352 *See* Press Release, California Dep't of Justice, Attorney General Becerra Secures Victory in Lawsuit Challenging Trump Administration Decision to Cut Penalties for Automaker Violations of Fuel Efficiency Standards (Aug. 31, 2020) (characterizing the efforts as "wrong-headed").

accountability. The impact of drift is too variable across even a single type of provision for a standard updating procedure to be applied without some unintended or adverse effects. But more accommodating procedures can be repurposed for political ends.

Finally, the decision to cap the catch-up adjustment well below the actual deterioration, even in the 2015 bill, underscores that updating old, nominally drafted statutes implicates large reliance interests, as the resistance of the auto industry in the NHTSA cases demonstrates.[353] Either reliance concerns or effectuating the original design can be legitimately prioritized, but this choice requires weighing of values not just tweaking formulas. The challenge is when, as they did twice with civil penalty adjustments, legislators try to split the difference. The result is a wholly new law masquerading as technical updating that still diverges from the original legislative scheme but also imposes new costs on regulated entities. Using a broad-based process compounds the issue, making it more difficult to weigh reliance on individual provisions. Even the limited escape valve of negative economic impact in the 2015 law does not fully alleviate this problem because there may be non-economic forms of reliance that the updating procedures prohibit from being weighed.

5. Sentencing Guidelines: a successful administrative-legislative partnership of uncertain durability

The final case study is distinct in that the updating effort was driven by an independent agency. The United States Sentencing Commission is empowered to promulgate amendments to federal sentencing law subject to Congressional veto.[354] This unique structure illustrates the tradeoffs of a flexible, agency-driven updating process.

Congress' refusal to pass a criminal fine indexing bill was not the last word on the subject. In 2015 the Sentencing Commission, inspired by recent Congressional efforts, examined whether it should update the monetary tables in its guidelines.[355] These tables fell into two categories. The first set fine amounts for offenses of a particular level (fine tables). The second determined the offense level of the crime based on the amount of pecuniary loss involved (loss tables). Greater pecuniary loss results in a higher offense level which results in a harsher fine. These values had not been adjusted for

---

353 Other notable examples of interest groups fighting to maintain inflationary drift include the Angel Capital Association successfully lobbying against updating the revision of the eroded accredited investor standard (Monson, *supra* note 24, at 38) and construction unions fighting for the maintenance of the low Davis Bacon Act contract value threshold (Weaver, *supra* note 52, at 240).

354 *See* Mistretta v. United States, 488 U.S. 361, 374–77 (1989).

355 Notice of Proposed Amendments and Request for Public Comment on Sentencing Guidelines, 80 Fed. Reg. 2570 (Jan. 16, 2015).

inflation since the guidelines were first issued in 1987.[356] As a result, crimes resulting in smaller financial harm were triggering greater increases in offense level than before and thus longer sentences.[357] Notably, inflation's impact was not unidirectional. The tables containing fines had also been left alone since 1989.[358] These had become roughly 50% less punitive in real terms.[359]

Ultimately the Commission addressed these dual drifts by making a one time, CPI-based update. Several thorny questions complicated the decision. Foremost was the propriety of updating. The DOJ opposed adjustments out of fear that updating the loss tables for inflation would "lead to an unwarranted reduction in sentences."[360] Why these adjustments to reflect economic reality were unwarranted, the DOJ did not say. Possibly the opposition was another instance of reliance: the DOJ thought more punitive guidelines had become useful. Regardless, the argument failed. At the public hearing on the revisions the DOJ's representative actually retreated from the position.[361] And the Commission dismissed the concern about an unwarranted reduction in sentences, describing it as puzzling.[362] To the Commission, any debate about whether the now higher penalties were beneficial was besides the point: "good governance" required calibrating the penalties to match their effect at enactment.[363]

The DOJ was also concerned about administrative authority. It argued that the death in committee of Senator Lautenberg's bill to adjust criminal penalties indicated Congressional hesitation about indexing fines.[364] The Commission also rejected these concerns. In the final rule, it cited the Improvements Act as implicit evidence of Congressional support for updating.[365] In the public hearing commissioners also speculated that the choice of an agency to set sentencing guidelines implicitly authorized temporal

---

356 80 Fed. Reg. 2570, 2579 ("While some of the monetary values . . . have been revised since they were originally established in 1987 (e.g., the loss table in § 2B1.1 was substantially amended in 2001), they have never been revised specifically to account for inflation. Other monetary values . . . have never been revised.").

357 Notice of Submission to Congress of Sentencing Guidelines, 80 Fed. Reg. 25782, 25789 (May 5, 2015).

358 *Id.* at 25789–90.

359 *See id.*

360 DOJ Comment Letter on Proposed Amendments at 12 (March 9, 2015), https://www.ussc.gov/sites/default/files/pdf/amendment-process/public-hearings-and-meetings/20150312/DOJ.pdf.

361 *See* U.S.S.C., Public Hearing on Proposed Amendment to Sentencing Guidelines 139 (March 12, 2015) (response of Mr. Wagner) (responding "I'll do my best" to an invitation to address the inflation adjustments which Wagner had failed to discuss in his prior comments).

362 March 12, 2015, Hearing, *supra* note 360, at 141, 146, 149.

363 March 12, 2015, Hearing, *supra* note 360, at 149.

364 DOJ March 9, 2015, Comment Letter, *supra* note 360, at 12.

365 80 Fed. Reg. 25782, 25789.

updating.[366] One commissioner justified the effort as meeting the commission's obligation to reduce unwarranted disparities.[367] However, in the text of the final rule the Commission only cited its general amendment powers as a source of authority.[368] These powers are more than sufficient to justify the update, but the lack of citation to any explicit guiding factor raises questions about whether the Commission will feel compelled to engage in similar updates in future.

It was also observed that updating automatically and regularly could cause instability and strategic behavior.[369] They worried that trials occurring near an update during an inflationary period would be dragged out by defense attorneys to benefit from upwardly revised loss tables.[370] In response, the Commission ultimately removed language it had been using about considering updates every four years from the final amendment.[371] The Commission has not made another update since 2015.

The Commission also faced tricky technical questions, such as which measure of inflation to use, how to deal with ex post facto concerns, and the baseline year to use in adjustments. The Commission chose to use CPI despite a recommendation to use the GDP Deflator.[372] The reasoning behind this choice is unknown. To ensure the amendments were not unduly retroactive the Commission made a revealing hybrid decision. It applied the old, unadjusted fines to any offenses committed before the effective date of the final rule, but it subjected those old offenses to the newly updated offense level thresholds.[373] In short, it applied the adjustment that benefited defendants retroactively and not the one that increased punishment. One might wonder whether the need to make this choice is consistent with the Commission's perspective that it was engaged in a purely technocratic exercise. Finally, the Commission decided to assume that prior ad hoc adjustments had implicitly

---

366 *See* March 12, 2015, Hearing, *supra* note 361, at 151 ("But it is interesting that the last time [Congress] looked at it was right before the Commission was founded. I mean, you know, they thought about this in 1987 and then, you know, here we are . . . It seems temporally that the thought is there is an agency that can account for it.").

367 *See* March 12, 2015, Hearing, *supra* note 361, at 144. Commissioner Pryor referenced 28 U.S.C. § 3553(a)(6) in support of this claim, but that is guidance on the factors a court should consider in imposing a sentence, not a source of authority for the commission.

368 80 Fed. Reg. 25782, 25782 (citing 28 U.S.C. § 994(a),(o),(p)). One should also note that the position of the U.S.S.C. vis a vis Congress is different from other administrative agencies. Any amendments to the Sentencing Guidelines proposed by the Commission are reviewed and can be changed by Congress. *See* 28 U.S.C. § 994(p). Though in the vast majority of cases Congress makes no changes.

369 March 12, 2015, Hearing, *supra* note 361, at 143.

370 *Id.* ("If I were a defense attorney . . . I would do everything I could to stall my sentencing, until that next [inflation adjustment], to see if my guy was going to get a break.").

371 *Cf. id.* at 143 (discussing the initial proposal).

372 *See* 80 Fed. Reg. 25782, 25789; DOJ March 9, 2015, Comment Letter, *supra* note 360, at 14.

373 80 Fed. Reg. 25782, 25789-90.

accounted for inflation—though they had not done so in any formal way—and thus chose the last update as the year to base its new adjustments on.[374]

Per the unique procedures of the Commission the final guidelines were issued on May 5, 2015, but did not become law until the period for Congressional veto passed on November 5, 2015.[375] The impact was significant, even accounting for the discretionary nature of the Guidelines. The Commission estimated that the changes to loss thresholds would free up almost a thousand prison beds within five years.[376] The increases to the fines table likely resulted in greater revenues and potentially increased deterrence, though data is unavailable.

In retrospect, the updates can be considered a qualified success. They were well-informed, and coherent. Most critically they thoughtfully approached technically complex issues, such as retroactivity and the interaction between loss tables and fines. This nuanced approach demonstrates the potential of delegation.

However, the debates involved in updating the Guidelines reveal several challenges in implementing such a scheme. First, the difference between good governance and policymaking is artificial. The Commission favored the reliance interests of defendants over other stakeholders in rejecting the DOJ's opposition to offense level increases and in the selective retroactive application of the adjustments. Yet the commissioners' repeated insistence that they were merely engaged in good governance makes one wonder if they fully grasped the import of their choices. Any shift from the status quo will require picking winners and losers. Ensuring such decisions are democratically accountable is a key challenge. Second, the technical details of updating are nontrivial. How should one reflect prior updates that may or may not have implicitly accounted for inflation? What adjustment methodology should one use? These questions lack obvious answers and required careful consideration by the Commission. Finally, the lack of a plan for subsequent adjustments demonstrates a drawback in such an agency-driven procedure. The 2015 update was the last explicit inflation adjustment to the Guidelines. Over the intervening years the CPI has risen 25%.[377] A more accountable effort needs an external impetus for future efforts. Absent the inclination of a handful of appointed, busy commissioners fines and thresholds could easily drift in the future.

---

374 *See id.* at 25789-90.

375 *Id.* at 25782.

376 *Id.* at 25790.

377 *Historical CPI-U Data*, *supra* note 8.

IV. CONCLUSION: THE NEED FOR COLLABORATIVE LAWMAKING IN A DYNAMIC SOCIETY

Inflationary drift presents a battle between respect for written law and applying that law justly in new contexts. It is a duel between fighters at full strength. There is no clearer or more precise legal command than the dollar value enshrined in statute. Simultaneously, one is hard pressed to find a more reliable or objective measure of contextual change than the shifting value of money. The clarity of these opposing factors leaves no room to hide from the problem they present: what to do when context changes but the law does not?

The Constitution's authors understood this problem well. As the opening quote from Governour Morris in response to a proposal to index judicial salaries to the price of grain suggests, the framers recognized that social change could undermine written law and that no formula could fully anticipate those changes. Yet they also recognized that a specific value is practically required. Judges don't work for indeterminate pay and the law functions through detail. The framers had a solution: delegate to the legislature. Congress' regular sessions and relatively flexible lawmaking procedure would mitigate the obsolescence challenges of a constitutional provision only changeable by amendment. Meanwhile, the boundaries set out in the compensation clause and the democratic accountability of Congress would prevent overstepping. For a rather long time this sort of delegation worked. While prices were relatively stable and statutory law fairly sparse, Congress was able to keep statutes current.

But the framers did not anticipate one thing—that the value of money would change in faster and prolonged ways. It would have been rather odd if they had thought of this. A crucial goal of their convention was to create monetary stability, and no one designs a government for failure. Nevertheless, this blind spot shaped their choice of delegate. Had they known the future they might have picked a nimbler institution or one less likely to ignore subtle technicalities in favor of political interests. Or they might have given the other branches a more explicit role to play in updating the law. They did not. And over time a series of doctrines developed—nominalism, plain meaning, and nondelegation—that incorporated the blind spots of that original decision by reinforcing strict formal separation of lawmaking power.

Fast forward to today, on the other side of a sweeping upward price curve, and we are left to contend with their limited foresight. What we know is this: The legislative process envisioned by the framers did not keep up with the heightened pace of change. Many statutes drifted in harmful ways. Criminal sentences grew harsher. Government benefits eroded. And macroeconomic volatility was imported into the very boundaries of the law itself, boundaries that defined who could trade in what markets, which employers had to comply with which rules, and even who could get into court.

Doctrines rooted in separation of powers are largely to blame for this state of affairs. Nominalism and the plain meaning rule have hindered

agencies and courts from taking unilateral action. Likewise, nondelegation principles, if not the doctrine itself, have restricted Congressional responses. Congress has generally controlled the updating procedure itself or provided inflexible guidelines to agencies. It has never felt empowered to give executive agencies authority to adjust penalties or payouts without specific instruction or veto. To be sure, exceptions exist, but formal lines retain much force and have clearly chilled more flexible attempts to address inflationary drift.

This article has attempted to illustrate how much the existing system has broken down in the face of unprecedented change. It also aims to point the way to a new approach. One option is to take a page from the framers' book and delegate further. Clearly, ad hoc updating has proven unworkable and, as the taxation and civil penalties cases demonstrate, having Congress develop an indexing procedure for each vulnerable law presents its own problems. Congress lacks the time and expertise for individualized efforts and broad-brush approaches can lead to significant error. Compared to that, having agencies determine when and how to update seems attractive. Indeed, it seems almost unavoidable. But the case studies reviewed above should give one pause before embracing such a wholesale delegation. Determining how to update a monetary value is not pure mathematics. At the very least, updating involves picking winners and losers between those relying on the nominal status quo and those who have been harmed by erosion. How should sentencing indexation be applied retroactively? Which components of income should be indexed? Which price indices best reflect the policy goals of indexed provisions? Any answer to these questions will benefit some constituencies and disadvantage others. And even more fraught judgments occasionally arise. Should the fiscal impact of entitlement programs increase automatically based on criteria set 50 years ago? What size claims should be allowed into federal courts? Surely some of these questions are best suited for a deliberative and democratically responsive body. Moreover, as the White House-led attempt to subvert legislatively mandated CAFE penalty updates shows, even decisions that might seem proper to delegate to agencies present opportunities for corrupt or arbitrary decision making. This discretion can be concerning even when exercised in good faith. Criminal defendants in this country are likely not comforted by the fact that sentencing loss tables will only stay current if the sentencing commissioners decide to include inflation updates on their agenda.

In short, no one branch is capable of resolving inflationary drift of statutes without serious capability or accountability concerns. A collaborative solution is needed. The successes and failures exhibited in the above case studies point towards a more workable sharing of power. Congress, while completely unsuited for determining updating amounts or applying updates each year, is fit to make the macro policy choices about whether updates should be sacrificed to fiscal or other concerns. Agencies, while not well suited and perhaps unable to make decisions on when to update, are the ideal parties to decide how to update. They understand best the interactions within

the statutes they administer and are well-positioned to make the technical but impactful choices in arriving at an updating amount. But there is always the possibility of pretext in agency updating. As the CAFE updating scandal proved, the judiciary is absolutely up to the task of sniffing out when agencies are attempting to evade a Congressionally assigned mandate. Courts can also help police the lines between what levels of decisions are for agencies and which are for Congress.

If this approach sounds familiar that's because it is. A number of scholars and jurists have observed that a similar updating process plays out with non-monetary statutory provisions via interpretation.[378] But this approach is rarely acknowledged for what it is because it contradicts separation of powers formalism. The great value of studying inflationary drift is that it removes this noise. We get to observe the implications of that formalism without the safety valve of interpretation. The results are stark: insisting on such strict separation is totally unworkable in a society of sustained change. Congress alone cannot ensure a body of law that matches citizen expectations. The three branches must integrate their dispersed powers if we are to have law that fits the times we live in. Those who say such an approach is not within the constitutional design should take another look at Figure 1. The Constitution did not design for that sort of sweeping change either.

---

[378] *See, e.g.*, Jody Freeman & David B. Spence, *Old Statutes, New Problems*, 163 U. PENN. L. REV. 1, 18–19 (2014); Donald C. Langevoort, *Statutory Obsolescence and the Judicial Process: The Revisionist Role of the Courts in Federal Banking Regulation*, 85 MICH. L. REV. 672, 675 (1987); Cass R. Sunstein, *Interpreting Statutes in the Regulatory State*, 103 HARV. L. REV. 405, 493 (1989).

NOTE: USING EMPIRICAL ANALYSIS TO ASSESS THE ACCURACY OF EXPERT TESTIMONY PREDICTIONS FROM THE AT&T-TIME WARNER VERTICAL MERGER

Lillian Clark

INTRODUCTION

The primary goal of antitrust law is to protect competition, and the understanding of how to do so successfully has changed drastically over time since the initial passing of the Sherman Act and Clayton Act. As economics and industrial organization expand their understandings of how to measure efficiencies and costs, courts and federal agencies must expand with them. The Federal Trade Commission and U.S. Department of Justice routinely regulate mergers, both horizontal and vertical, under the Clayton Act. Periodically, it is necessary for the agencies to reevaluate their approach to this regulation.

Most recently, the agencies published their Vertical Merger Guidelines in 2020.[1] However, with new leadership at the FTC and Justice Department, the FTC withdrew support for these guidelines in 2021.[2] The Vertical Merger Guidelines are now being rewritten after only a short two years.[3] The politicization of antitrust has enhanced the inconsistent analysis of vertical mergers, where an increasing desire to regulate them within the agencies has been met with hesitation from the courts.[4]

The AT&T-Time Warner merger was the first major vertical merger challenge brought by either agency since 1979.[5] This case represented a landmark shift in agency regulation by turning the nation's focus back on vertical

---

[1] Press Release, Fed. Trade Comm'n, FTC and DOJ Issue Antitrust Guidelines for Evaluating Vertical Mergers (June 30, 2020), https://www.ftc.gov/news-events/news/press-releases/2020/06/ftc-doj-issue-antitrust-guidelines-evaluating-vertical-mergers.

[2] Press Release, Fed. Trade Comm'n, Federal Trade Commission Withdraws Vertical Merger Guidelines and Commentary (Sept. 15, 2021), https://www.ftc.gov/news-events/news/press-releases/2021/09/federal-trade-commission-withdraws-vertical-merger-guidelines-commentary.

[3] Press Release, Fed. Trade Comm'n, Federal Trade Commission and Justice Department Seek to Strengthen Enforcement Against Illegal Mergers (Jan. 18, 2022), https://www.ftc.gov/news-events/news/press-releases/2022/01/federal-trade-commission-justice-department-seek-strengthen-enforcement-against-illegal-mergers.

[4] *See* Larry Bumgardner, *AT&T and Time Warner's Vertical Merger: The Court Battle and Political Undercurrent*, 25 J.L. BUS. & ETH. 31, 34 (2019); *see also* Allison Neff, *A Reassessment of Vertical Mergers within the Context of Antitrust Laws: The Time Warner and AT&T Merger*, 44 DEL. J. CORP. L. 121, 125 (2020).

[5] Babette Boliek, *Antitrust and High Tech: A Tale of Two Mergers*, 71 EMORY L.J. 933, 940 (2022).

mergers after four decades.[6] The Justice Department worried that this merger would allow the new merged firm to raise its rivals' costs while also inhibiting future innovation.[7] However, the Justice Department lost at trial and again lost on appeal, as the courts found the agency expert had not adequately accounted for either the dynamic nature of the involved industries or the lack of actual evidence of similar behavior in practice.[8]

As the Vertical Merger Guidelines are currently being rewritten and there is minimal recent court precedent regarding vertical mergers, courts and the agencies do not have a clear direction regarding vertical merger regulation. Being four years out from the AT&T-Time Warner merger, there is now relevant data to assess the accuracy of the Justice Department's theories and whether both the United States District Court for the District of Columbia and the United States Court of Appeals for the District of Columbia Circuit were correct in assessing that the procompetitive benefits of the merger outweighed the anticompetitive effects. Understanding how the theories played out empirically can help the agencies in their expert analyses of future vertical mergers. Looking at the case and expert testimony, this paper aims to identify whether the merged firm of AT&T-Time Warner was successfully able to leverage Time Warner's packages to better DirecTV in the television provider market.

First, this paper explains the background of vertical merger analysis, the AT&T-Time Warner case, and the economic theories underlying the agency's decision-making in that case. The background section begins with a history of vertical merger regulation in the United States and how it has changed over time. It starts with the original intent behind antitrust laws and then discusses how and why regulation has expanded from the initial purpose to broader regulation today. There is a brief timeline and comparison of the different merger guidelines that the DOJ and FTC have passed over time.

Then, the paper deconstructs the AT&T-Time Warner case. Primarily, the paper focuses on the decision's seemingly political nature and why this decision by the Justice Department marks a shift in antitrust policy. This section also discusses the timeline of the case and the rationale behind the court opinions in refusing to block the merger. The background section then specifically dissects the theories and empirics found in the economic expert testimony from the Justice Department. It is followed by a brief summary of AT&T-Time Warner's decision to re-divide and specialize, although this will not impact the empirical research.

The next section of the paper will be the empirical analysis. The primary goal is to determine whether the Justice Department's expert was correct in

---

[6] Neff, *supra* note 4, at 126.

[7] *See* Complaint, ¶¶ 9, 31, United States v. AT&T, Inc., 310 F. Supp. 3d 161 (D.D.C. 2018) (No. 17 Civ. 2511).

[8] United States v. AT&T, Inc., 310 F. Supp. 3d 161, 194, 217–18 (D.D.C. 2018); United States v. AT&T, Inc., 916 F.3d 1029, 1032–34 (D.C. Cir. 2019).

anticipating that the AT&T-Time Warner merger would negatively impact DirecTV's competitors through raising rivals' costs, allowing DirecTV to unfairly gain subscribers and market share over its competitors. This section explains that the models will use the merger as the independent variable and assess its impact on monthly subscriber count, the dependent variable. The models also account for a variety of economic trend datapoints and events relevant to the cases. The section will explain the set of datapoints, their significance, their relevance, and their source.

Then, this section will present the statistical analysis and the underlying model. There are three regressions to compare, each similarly finding that the merger actually had a statistically significant and negative effect on DirecTV's monthly subscriber count; however, because subscriber count is the only available dependent variable, as opposed to price, the models can only demonstrate a reduced form equilibrium association rather than a demand model with causal interpretation. The paper then presents a sanity-check, comparing the regression analyses to how DirecTV's market share actually changed over time in the television provider market. The market share analysis and graph demonstrate that DirecTV not only lost monthly subscribers but also that this loss came with DirecTV's rivals being able to secure a larger market share.

Following the regression will be an explanation of how to both interpret and apply these results to future vertical merger analysis, as well as how to compare it to the original economic expert testimony. The results of the anecdotal empirical study of the AT&T-Time Warner merger show that the Justice Department may have overlooked the relevance of industry constraints from all layers of the newly merged vertical firm. Another key takeaway is that the more dynamic and innovative an industry is, the less likely anticompetitive conduct will be successful. The intention is that, while only anecdotal, the results of this merger can be used as actual evidence in future vertical merger cases when determining whether the Justice Department or FTC should challenge.

Finally, the paper will conclude with re-emphasizing the importance of proper vertical merger analysis. The Vertical Merger Guidelines are in dispute between those at the agencies, and courts lack direction when analyzing the complex weighing of efficiencies and costs in mergers that do not involve direct competitors. This paper seeks to bring light to some of the many complications in vertical merger analysis and hopefully restart the bipartisan effort to ensure that competition is protected efficiently.

BACKGROUND

I. HISTORY OF VERTICAL MERGER REGULATION

Antitrust laws in the United States began with the Sherman Act in 1890, followed by the Clayton Act in 1914.[9] The DOJ's Antitrust Division and FTC were tasked with regulating competition.[10] Until 1949, the laws at hand only permitted regulation of horizontal mergers.[11] The Celler-Kefauver Amendment of 1950 broadened the Clayton Act to allow agency regulation of vertical mergers, as well as conglomerate mergers or acquisition of assets, all of which were unable to be regulated prior to this amendment.[12] *United States v. Bethlehem Steel Corporation*[13] in 1958 and *Brown Shoe Company v. United States*[14] in 1962 confirmed this reach. These cases where characterized by an overall fear of acquisition between any two companies with some semblance of significant market share and primarily focused on the "size of the market [expected to be] foreclosed" as the result of the vertical merger.[15]

Although *Bethlehem Steel* and *Brown Shoe* are still considered good law, the actual analysis of vertical merger conduct has changed drastically since these early cases. The FTC and DOJ brought only two other significant vertical merger cases in the 1970s.[16] Since then, no major vertical mergers were challenged by either agency until the AT&T-Time Warner merger.[17]

In 1984, the Merger Guidelines included for the first time a section about regulating non-horizontal mergers.[18] It made clear that since the 1970s, the biggest difference between regulating horizontal and vertical mergers were that vertical merger efficiencies were given more weight than horizontal mergers.[19] The guidelines set the standard of only challenging a vertical merger when the projected HHI was over 1800.[20] It also suggested that the primary concerns should be harm to horizontal competitors within the relevant

---

[9] 15 U.S.C. §§ 12–27.

[10] *Id.*

[11] Charles J. Steele, *A Decade of the Celler-Kefauver Anti-Merger Act*, 14 VAND. L. REV. 1049, 1050–51 (1961).

[12] 64 Stat. 1125 (1950); 15 U.S.C. § 18.

[13] 168 F. Supp. 576 (S.D.N.Y. 1958).

[14] 370 U.S. 294 (1962).

[15] *Id.* at 328.

[16] *See generally* Ford Motor Co. v. U.S., 405 U.S. 562 (1972); Fruehauf Corp. v. F.T.C., 603 F.2d 345 (1979).

[17] Boliek, *supra* note 5, at 940.

[18] *See* U.S. Dep't of Justice, Non-Horizontal Merger Guidelines, § 4 (1984), https://www.justice.gov/sites/default/files/atr/legacy/2007/07/11/11249.pdf.

[19] *Id.* §4.24.

[20] *Id.* § 4.213.

markets, as well as harm to potential and actual entry.[21] After this, the agencies made no significant movement in their analyses or challenging of vertical mergers for almost forty years.[22]

In 2017, the Justice Department made the first move to challenge a vertical merger once again.[23] After almost four long decades without any change in vertical merger analysis by the agencies or the courts, much speculation was required to keep up with the current economic studies regarding vertical mergers and both their costs and efficiencies. In response to this new desire to again regulate vertical mergers, the DOJ and the FTC released the 2020 Vertical Merger Guidelines.[24] This was the first time the agencies released a formalized guide specific to vertical mergers.[25] It is distinct from the 1984 Non-Horizontal Merger Guidelines because it applies economic policy unique to vertical mergers and addresses how the main concerns of vertical mergers differ from those of other mergers.[26]

In particular, the 2020 Vertical Merger Guidelines laid out the primary costs of vertical mergers as market foreclosure for competitors and the raising of rivals' costs.[27] The second of these was the main theory of harm in the AT&T-Time Warner challenge.[28] These guidelines also demonstrate the efficiencies that need to be accounted for by the courts, such as the elimination of double marginalization, which was one of the main defenses used in the AT&T-Time Warner case.[29] However, in 2021, the FTC withdrew its support for these guidelines when the leadership at the agencies changed.[30]

The 2020 Vertical Merger Guidelines, as well as the AT&T-Time Warner challenge and subsequent withdrawal of the guidelines, rocketed vertical merger regulation into the antitrust limelight, leaving the agencies and courts uncertain of how to analyze vertical mergers. The agencies have little guidance on how and when to regulate vertical mergers under current precedent, but ideally an understanding of how the AT&T-Time Warner merger played out can assist antitrust lawyers by adding another real-world example to their repertoire.

---

21 *Id.*§ 4.1.

22 Boliek, *supra* note 5, at 940.

23 *See* Complaint, *supra* note 7; *AT&T*, 310 F. Supp. 3d at 161.

24 *See* U.S. Dep't of Justice & Fed. Trade Comm'n, Vertical Merger Guidelines (2020), https://www.justice.gov/atr/page/file/1290686/download [hereinafter Vertical Merger Guidelines (2020)].

25 *Id.*

26 *See id.* at 2; U.S. Dep't of Justice, *supra* note 18.

27 Vertical Merger Guidelines (2020), *supra* note 24, at 4.

28 Boliek, *supra* note 5, at 940–41; Complaint, *supra* note 7.

29 Vertical Merger Guidelines (2020), *supra* note 24; *AT&T*, 310 F. Supp. 3d at 197–98.

30 Press Release, *supra* note 2.

II. AT&T-TIME WARNER PROCEDURAL HISTORY

On October 22, 2016, AT&T announced its proposed acquisition of Time Warner.[31] The companies' press release described that the new firm would have "extensive customer relationships, world's largest pay TV subscriber base and leading scale in TV, mobile, and broadband distribution." [32] Unsurprisingly, the case garnered global attention, including that of antitrust regulators. However, the Antitrust Division at the Justice Department's Assistant Attorney General Makan Delrahim decided it was not, "a major antitrust problem."[33] However, the tables quickly turned as the Justice Department mobilized to block the merger, noting that the behavioral remedies it had prescribed to prior vertical mergers were difficult to enforce.[34]

The Justice Department filed its complaint on November 20, 2017, officially opening the case for investigation.[35] The United States District Court for the District of Columbia ruled on June 12, 2018, that the "Government's request to enjoin the proposed merger [was] denied," expressing that the Defendants' economic experts were more persuasive than those of the Justice Department.[36] The DOJ appealed the case, which was then affirmed by the United States Court of Appeals for the District of Columbia Circuit on February 26, 2019.[37]

One of the main reasons why the D.C. Circuit sided with the district court was that the Defendants' expert provided an "analysis of real-world data for prior vertical mergers in the industry that showed 'no statistically significant effect on content prices.'"[38] The lack of economic evidence in general of negative effects from vertical mergers dissuaded the court from taking further action in blocking the merger, especially noting that the government's "expert opinion and modeling predicting such increases failed to take into account Turner Broadcasting System's post-litigation irrevocable offers of no-blackout arbitration agreements, which a government expert acknowledged would require a new model."[39] The constantly changing market consisting of television, mobile phones, and other technology was credited as too dynamic for the Justice Department's concerns to be warranted.[40]

These two and a half years involved the first agency regulation of a vertical merger in decades, and although the economics applied by the agency

---

31 Bumgardner, *supra* note 4, at 33.

32 *Id.*

33 *Id.* at 36.

34 *Id.* at 37.

35 *See generally* Complaint, *supra* note 7.

36 *AT&T*, 310 F. Supp. 3d at 254.

37 *AT&T*, 916 F.3d at 1047.

38 *Id.* at 1031–32.

39 *Id.* at 1031.

40 *Id.* at 1031–32.

expert had incorporated modern economics and a balance between efficiencies and costs, the courts rejected the expert testimony of the agencies, so AT&T and Time Warner officially merged.[41] In a vertical merger challenge specifically, the burden of proof is much higher because there is no presumption of anticompetitive behavior.[42] The Justice Department was unable to meet the bar in this case without a real industry examples of a long-term blackout and evidence of efficient ways to successfully raise rivals' costs to the merged firm's benefit.[43]

Interestingly, after the merger was consummated, Warner Media quickly began "offering consumers bundled entertainment packages . . . driving new customers to sign up for its phone plans" while also increasing the price "for DirectTV Now Service, pushing consumers to downgrade their service for bundled packages."[44] Authors like Neff surmise that this implies that the Justice Department was correct in their analysis, although this evidence was not sufficient to reverse the case in the D.C Circuit.[45]

III. SUMMARY OF THE AGENCY EXPERT REPORT

Dr. Carl Shapiro served as the Justice Department's economic expert throughout this case.[46] Highly qualified, Dr. Shapiro is a professor at both the Graduate School at the Haas School of Business and the Department of Economics at the University of California at Berkeley.[47] He has extensive experience in antitrust economics and has served as an expert for the agencies on multiple occasions.[48] For the purposes of this paper, his testimony is the basis for the analysis and determining whether the Justice Department was able to accurately predict what would occur if the AT&T-Time Warner merger was completed.

Although the Justice Department proposed numerous theories of harm, the primary focus of the investigation and the only issue appealed was whether the AT&T-Time Warner merger would result in raising rivals' costs, "resulting in a blackout."[49] All antitrust cases must start with defining a relevant market, and the proposed market by the Justice Department and Dr. Shapiro was all "distributors of professionally produced, full-length video

---

41 *Id.*

42 James A. Keyte, *The AT&T/Time Warner Decision: More Than Meets the Eye*, 33-Sum ANTITRUST 20, 20–21 (2019).

43 Bumgardner, *supra* note 4, at 33.

44 Neff, *supra* note 4, at 128–29.

45 *Id.*; *AT&T*, 916 F.3d at 1047.

46 *See generally* Shapiro Expert Report, United States v. AT&T, Inc., 310 F. Supp. 3d 161 (D.D.C. 2018) (No. 17 Civ. 2511).

47 *Id.* at 1, Appendix A.

48 *Id.*

49 Complaint, *supra* note 7, ¶ 5; *AT&T*, 916 F.3d at 1035–36, 1047.

programming subscription services to residential customers in the United States."[50] In this paper, this market will be referred to as the television provider market.

The Justice Department estimated price increases by measuring:

> 1. *The Turner Subscriber Loss Rate*: The rate at which the rival MVPD [Multichannel Video Programming Distributor] would lose subscribers over time if it could not offer the Turner Content;
> 2. *The DTV* [DirecTV] *Diversion Ratio*: The proportion of the subscribers leaving an MVPD, if that MVPD could not offer Turner Content, that would shift to DTV;
> 3. *DTV's Contribution Margin*: The difference between DTV's PSPM subscription fee and the incremental cost to DTV of serving one more subscriber.[51]

These variables, which the Justice Department named the Turner Bargaining Model, after being introduced, were input into a model to estimate the harm to consumers.[52] The Turner Bargaining Model is based off of the Nash Bargaining Model, which was derived from the work of economist John Nash.[53] Even with a conversative estimate, Dr. Shapiro predicted consumer welfare would lose hundreds of millions of dollars each year.[54]

The primary theory predicted that AT&T and Time Warner's newly merged firm would have the ability to leverage Time Warner's premium content against DirecTV's competitors.[55] This would allow the firm to withhold Time Warner's content from the rivals unless they pay a much higher price for it, leaving competitors to choose between sacrificing content or paying higher prices, likely passing on those costs to consumers.[56] The Justice Department considers this leverage to be substantial, especially given that Time Warner owns HBO, which includes *Game of Thrones*, as well as the rights to NCAA March Madness and other major networks.[57]

---

50 Complaint, *supra* note 7, ¶ 27.

51 Expert Report, at 49–50, United States v. AT&T, Inc., 310 F. Supp. 3d 161 (D.D.C. 2018) (No. 17 Civ. 2511).

52 Trial Brief, at 34, United States v. AT&T, Inc., 310 F. Supp. 3d 161 (D.D.C. 2018) (No. 17 Civ. 2511).

53 *Id.* at 33.

54 *Id.* at 35–36.

55 Boliek, *supra* note 5, at 936.

56 *Id.* at 940–41.

57 Complaint, *supra* note 7, ¶ 4.

If these predictions were correct, DirecTV would be expected to provide its service without the upcharge of using Time Warner's content, as well as being able to ensure consumers access to that content.[58] If the theory is correct that DirecTV therefore would be able to unfairly outcompete its rivals after the merger, the data would likely show that DirecTV gained monthly subscribers and market share in the television provider network after the merger. Dr. Shapiro estimated rivals could lose anywhere from nine to fourteen percent of their monthly customers over time.[59] However, it is important to note that "[e]ven talented economists with access to extensive industry reporting data acknowledge that conclusions in this area are complicated and open to interpretation."[60]

The district court ruled that the industry was too dynamic for this theory to be a valid concern.[61] Judge Leon states clearly in his opinion that the empirical evidence did not properly account for the industry structure and previous natural data.[62] Within the television industry, there is no history of long-term blackouts of Turner networks, and the empirical evidence does not, in Judge Leon's opinion, adequately account for the binding long-term contracts that Time Warner was currently engaged in.[63] Therefore, the district court, affirmed by the D.C. Circuit, suggested that the Justice Department's concerns were unlikely to come true, especially with the ever-changing industries that AT&T and Time Warner are both reliant on and responsive to.[64]

IV. AT&T AND TIME WARNER'S DECISION TO SPLIT UP

As of May 2021, the merged firm of AT&T-Time Warner announced its decision to divest WarnerMedia and focus primarily on the communication business and acquisition of 5G networks for mobile phones, which was AT&T's primary industry before the merger.[65] This came as a result of the T-Mobile-Sprint horizontal merger, which the agencies did not challenge in 2020.[66]

Although this will likely impact any analysis of how the merged firm affected prices as a vertically integrated firm, its effect should be limited in the empirics here if the data is capped after May 2021, when the divestiture was announced.[67] This also reflects a response to AT&T's primary market,

58 Boliek, *supra* note 5, at 936.
59 *Id.* at 941.
60 *Id.* at 942.
61 *AT&T*, 310 F. Supp. 3d at 194, 254.
62 *Id.* at 253–54.
63 *Id.*
64 *Id.*; *AT&T, Inc.*, 916 F.3d at 1047.
65 Boliek, *supra* note 5, at 949.
66 *Id.* at 946–47.
67 *Id.* at 938.

not DirecTV's, which is the television provider market. While AT&T may have moved resources away from DirecTV in response to this merger earlier on, the effect of the merger on the television provider market should still be isolated enough to analyze.

For added security, the paper will not look past the first quarter of 2021, meaning the dataset will only reflect a time period through March 2021 and not any further. However, it is important to note that while this influence is one that was not predicted by the Justice Department, it does correspond with the ongoing industry dynamics alluded to by the defense experts and Judge Leon.[68]

ANALYSIS

I. DATA COLLECTION AND STATISTICS

For the analysis, the primary question is whether AT&T-Time Warner's merger affected DirecTV's ability to gain subscribers and attain a higher market share, as the Justice Department predicted. The DOJ's main theory is that the merger will allow AT&T-Time Warner to raise costs for rival television providers by increasing the cost for those networks to do business with Time Warner.[69] If this prediction is correct, the Justice Department surmises that it would be likely for DirecTV to see an increase in subscribers after the merger, as well as an increase in market share compared to the other cable networks.[70]

For the upcoming analysis, the dependent variable is the monthly subscriber count (Subscribers), and the independent variable is a dummy variable that is equal to "one" only for DirecTV after the merger has occurred (Merger); it is "zero" for all other companies, and "zero" for DirecTV before the merger. The intention is to see whether the merger had a positive impact on the monthly subscriber count for DirecTV because of the merger, affirming the predictions of the Justice Department, while accounting for other factors likely to impact monthly subscriber count to confirm that the change in subscriber count is the result of the merger itself, outside of growth or decline naturally expected from external factors.

To determine whether this growth for DirecTV happened as predicted, this paper will assess the subscriber count and market share over time for all the major cable networks while considering economic trends, relevant events, and the merger's occurrence. This data comes from a variety of sources and spans the time between the fourth quarter of 2016 and the third

---

68 *AT&T*, 310 F. Supp. 3d at 254.

69 Complaint, *supra* note 7, ¶¶ 3–5.

70 *Id*., ¶ 5; Boliek, *supra* note 5, at 941.

quarter of 2021. From Statista, quarterly data was pulled for each of the major cable networks, providing the Year, Quarter, Company, and Subscriber Count in millions (Subscribers).[71] These data were combined to create a mass dataset that included the companies DirecTV, Comcast, Charter, Dish, Verizon, Altice, and Mediacom, which account for over sixty-four million subscribers in the most recent quarter, the third quarter of 2021.[72] DirecTV, Comcast, and Charter, better known as Spectrum, are considered the largest three television providers, accounting for over 75% of the television provider market as of the third quarter of 2021, according to the Statista data.[73]

To address variables outside of the merger that may have affected subscriber count, the dataset includes Real Gross Domestic Product (GDP) from the Federal Reserve Bank of St. Louis database.[74] The data is quarterly to match the company data.[75] The Real GDP is measured in billions and is chained to the 2012 dollar and adjusted seasonally.[76] This ensures that the data is consistent over the entire dataset and inflation is accounted for separately. The Real GDP is important to demonstrate the overall flow of wealth in the U.S. economy, as this would affect aggregate demand for all goods, including cable networks.[77] Including GDP should account for any general economic ebbs and flows that would impact monthly subscriber count, isolating the effect of the merger on subscriber count.

Next, because of the Justice Department's fear that Time Warner's ownership of HBO would be a large factor in raising rivals' costs, there were two binary variables created to account for the effect of this ownership. Firstly, HBO producing and releasing new episodes of *Game of Thrones* (GoT) was likely to increase the demand for HBO, and therefore Time Warner products, during this time.[78] *Game of Thrones* stopped airing in the second quarter of 2019, so the binary variable is listed as a "one" before and in that quarter and a "zero" after.[79] HBO Max was launched in the second quarter of 2020, which allowed consumers to purchase HBO separately from their cable packages, potentially lowering demand for specific cable networks.[80] A "one" marks

---

71 Julia Stoll, *Pay TV in the United States - Statistics & Facts*, STATISTA (Mar. 31, 2022), https://www.statista.com/topics/1309/pay-tv/.

72 *Id.*

73 *Id.*

74 U.S. Bureau of Econ. Analysis, *Real Gross Domestic Product*, THE FED. RSRV. BANK OF ST. LOUIS (July 1, 2022), https://fred.stlouisfed.org/series/GDPC1.

75 *Id.*

76 *Id.*

77 *Id.*

78 Complaint, *supra* note 7, ¶ 4.

79 IMDB, *GAME OF THRONES* EPISODE LIST (2019), https://www.imdb.com/title/tt0944947/episodes?year=2019.

80 Todd Spangler, *HBO Max Sets Official Launch Date*, VARIETY (Apr. 21, 2020), https://variety.com/2020/digital/news/hbo-max-launch-date-price-streaming-1234585776/; Complaint ¶ 4, United States v. AT&T, Inc., 310 F. Supp. 3d 161 (D.D.C. 2018) (No. 17 Civ. 2511).

the second quarter of 2020 as well as any time after, and a "zero" before that period signifies that HBO Max did not yet exist. *House of the Dragon* began airing in the third quarter of 2022, after our dataset ends, so this does not need to be considered.[81]

Finally, a last set of binary variables was created to account for the impact of the merger's announcement, challenge, success, and end, as well as the potential impact of Covid-19 on overall aggregate demand. The merger was announced in the fourth quarter of 2016 (Merger Announced), the Justice Department challenged the merger in the fourth quarter of 2017 (DOJ Challenged), the DOJ was prevented from blocking the merger in the first quarter of 2019 (DOJ Lost), and finally the merger was then undone in the second quarter of 2021 (Merger Ended).[82] Then, the Covid-19 variable (Covid) considers the first quarter of 2020 as the starting impact of Covid-19 on the U.S. economy.[83] Starting with when each event took place and from then on, each corresponding variable is marked with a "one"; before each event took place, there is a "zero" for that variable.

This set of variables makes up the panel dataset. A panel dataset is a multi-dimensional dataset that covers a variety of variables over time.[84] This is most appropriate in this setting, as the status of the economy, inflation, and population, as well as the timeline of the merger, would affect subscriber count over time and in correlation with one another.[85] A fixed effects model is then applied to identify the trend between the dependent variable and independent variable while accounting for the other constants that likely affected the dependent variable, monthly subscriber count, for all the firms regardless of merger status.[86]

II. REGRESSIONS ANALYZING THE CORRELATION BETWEEN THE AT&T-TIME WARNER MERGER AND SUBSCRIBER COUNT

Using R, the data was input into the fixed effects panel regression model. The first regression uses the monthly subscriber count as the dependent variable and the merger binary variable as the independent variable. The model also includes the Real GDP, which is used to account for how economic trends may have impacted monthly subscriber count. This first regression is shown in Figure 1.1.

---

[81] IMDB, *HOUSE OF THE DRAGON* EPISODE LIST (2022), https://www.imdb.com/title/tt11198330/episodes?year=2022.

[82] Boliek, *supra* note 5, at 938, 952.

[83] CONGRESSIONAL RESEARCH SERVICE, COVID-19 AND THE U.S. ECONOMY 1 (2021).

[84] Mahbubul Alam, *Panel Data Regression: A Powerful Time Series Modeling Technique*, TOWARD DATA SCIENCE (Feb. 26, 2020), https://towardsdatascience.com/panel-data-regression-a-powerful-time-series-modeling-technique-7509ce043fa8.

[85] *Id.*

[86] *Id.*

Given this model, the AT&T-Time Warner merger (Merger) was estimated to have a coefficient of -3.34. This implies that the merger actually decreased the subscribers for DirecTV by about 3.34 million subscribers monthly on average. The effect had a p-value of less than 0.001, meaning the effect is statistically significant and the null hypothesis that the merger had no effect on subscribers can be rejected. To reject the null hypothesis that a variable had a statistically significant effect, the p-value must be less than 0.05, and so it is noteworthy that the p-value in this regression is even closer to zero than necessary, demonstrating even more statistical significance.[87] The merger actually had a negative effect on subscribers, according to this model. It is important to note, however, that these data points show a reduced form equilibrium association rather than a full demand model, as many of the variables needed to do the full assessment are unavailable to the public.

Figure 1.1

Variable	Estimate	P Value	Significance
Merger	-3.34	0.00	***
GDP	0.00	0.92	
F-Statistic	112.40		
Adjusted R-Squared	0.90		

The adjusted R-squared indicates how much of the variation in the dependent variable can be associated with the input variables and is measured between zero and one.[88] An adjusted R-squared of 0.90, as produced from this model, can be read as that the variation in monthly subscribers can be ninety-percent accounted for as being from the given input variables.[89] The higher the adjusted R-squared, the more likely it is that the model can act as a helpful predictive tool.[90] The large F-statistic of 112.40 also demonstrates a high likelihood that these results were statistically significant rather than the product of random chance.[91] Given the low p-value, high adjusted R-squared, and high F-statistic, it is statistically likely with the limited data available that the merger had a significant negative impact on monthly subscribers for DirecTV, as shown by the three asterisks.

This regression was repeated twice more to account for other potential variables. The second regression, shown in Figure 1.2, does not consider the

---

87 Zach, How to Interpret the F-Value and P-Value in ANOVA, STATOLOGY (Aug. 16, 2021), https://www.statology.org/anova-f-value-p-value/.

88 Zach, *How to Interpret Adjusted R-Squared (With Examples)*, STATOLOGY (Mar. 24, 2022), https://www.statology.org/adjusted-r-squared-interpretation/.

89 *Id.*

90 *Id.*

91 Zach, *How to Interpret the F-Value and P-Value in ANOVA*, STATOLOGY (Aug. 16, 2021), https://www.statology.org/anova-f-value-p-value/.

economic trend variable, but instead focuses on the influence of *Game of Thrones*, HBO Max, Covid-19, the merger announcement, merger ending, the challenge to the merger by the Justice Department, and the decision by the D.C. Circuit Court of Appeals to prevent the Justice Department from blocking the merger.

Similarly, the merger was the only statistically significant variable, showing an effect of about -3.45 million subscribers each month. *Game of Thrones*, the merger announcement, and the Justice Department's loss at trial were all positive influences on the subscriber count, though minimally, while the introduction of HBO Max, Covid-19, the end of the merger, and the challenge by the Justice Department all negatively, but insignificantly, impacted the subscriber count. The p-value for the merger was less than 0.001 in this regression, as well, showing that the merger's influence was again statistically significant. In conjunction with the high F-statistic and high adjusted R-squared, this regression demonstrates the same result as the first.

Figure 1.2

Variable	Estimate	P Value	Significance
Merger	-3.45	0.00	***
GoT	0.41	0.68	
HBO Max	-0.50	0.66	
Covid	-0.32	0.88	
Merger Announced	0.94	0.42	
Merger Ended	-0.49	0.67	
DOJ Challenged	-0.11	0.93	
DOJ Lost	0.24	0.90	
F-Statistic	91.72		
Adjusted R-Squared	0.90		

Finally, a third regression, shown in in Figure 1.3, demonstrates the effects of all the variables used thus far together. Similarly, the merger showed a statistically significant effect on monthly subscriber count for DirecTV of about -3.45 million subscribers each month. The Real GDP again had no real effect on subscriber count, while *Game of Thrones* and the merger announcement and a positive, but not statistically significant, impact on monthly subscriber count. HBO Max, Covid-19, the end of the merger, the Justice Department's challenge, and the Department's loss at trial all had an insignificant but negative effect on monthly subscriber count. Again, the p-value, F-statistic, and adjusted R-squared indicate the persuasiveness of this regression, demonstrating that it is likely that the merger between AT&T and Time Warner had a significantly negative effect on the monthly subscriber count for DirecTV.

Figure 1.3

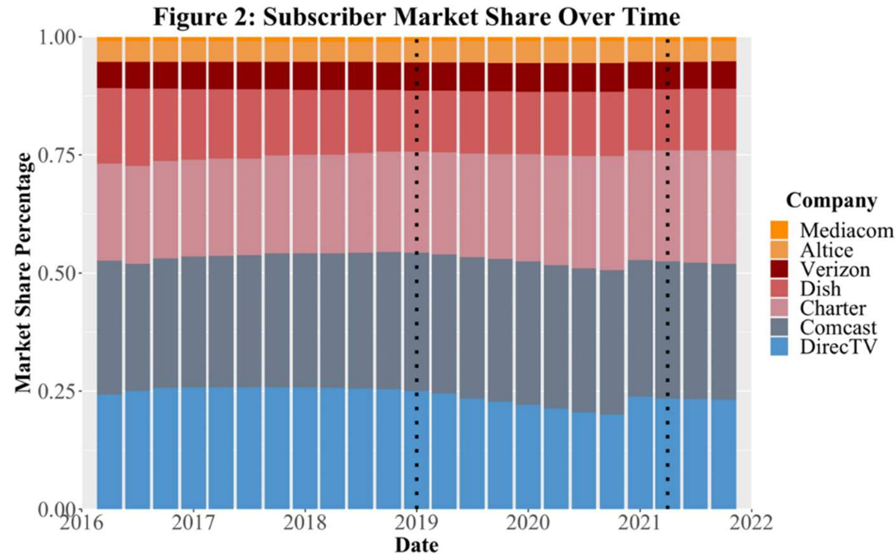
Variable	Estimate	P Value	Significance
Merger	-3.45	0.00	***
GDP	0.00	0.82	
GoT	0.44	0.66	
HBO Max	-0.69	0.63	
Covid	0.03	0.99	
Merger Announced	0.90	0.45	
Merger Ended	-0.43	0.72	
DOJ Challenged	-0.12	0.92	
DOJ Lost	0.74	0.80	
F-Statistic	87.51		
Adjusted R-Squared	0.90		

These results come with a few caveats. Of course, the data was limited to what was discoverable. The actual data used by the experts was obtained through a formal discovery process and is private information, so it is unavailable for public use or research. Because of this, the analysis is not a direct comparison to what was predicted. Rather, the analysis compares the current situation of DirecTV to the overall theory of the Justice Department expert. This limitation also means that the regression simply represents a reduced form equilibrium association rather than a full demand model with causal implications.

Also, as with any economic study, it is impossible to fully know what would have occurred had the merger not been initiated. It is possible DirecTV's subscriber count and market share were always going to decline after 2019. While the study attempts to account for all outside variables, it can never be positive that this effect did not come from a source other than the merger. For example, while Covid-19, HBO Max, and the economic trend variable (GDP) strive to account for movement from television providers to streaming services, there is no definitive variable that can account for these preference changes in the data that were available.

To perform a sanity check and assess how the merger affected DirecTV on a larger scale, one may look at market shares. The market shares over time demonstrate how DirecTV has performed since the merger relative to other television providers. Figure 2 provides a graphic that visualizes the market shares, showing the downward trend in DirecTV's subscriber count. Because the DirecTV market shares are declining after the merger, not just subscriber count alone, one can assume the decline in subscriber count is not solely from a decline in the demand for television providers in general. The subscribers are being captured by the other television providers who are gaining market

share, signifying that this negative impact is truly from the merger and the consequences of the conduct by DirecTV.



In 2018, the data show that DirecTV had a market share of approximately 26% and was considered the second largest television provider, second to Comcast. By 2019, DirecTV's market share dropped to 25%, while Comcast, Charter, Dish, and Verizon all saw an increase in market share. In 2020, DirecTV's share dropped again to 22%, only rising in 2021 after the merger had ended back to about 24%, still under where it was before the merger was announced.

The regressions and market share analysis demonstrate that the merger was not successful for DirecTV in terms of retaining or increasing total monthly subscribers. While it could have been successful for AT&T-Time Warner in other ways, DirecTV saw a decline in monthly subscribers and market share during the time, departing from the predictions of the expert for the Justice Department. While the DOJ anticipated this merger would specifically be used to increase market control for DirecTV, raising its subscriber count and market share, as well as its profitability, in actuality the merger resulted in a decrease in market power and control for DirecTV. This decline may be related to the merger, and DirecTV's decline was met with its rivals' success in gaining market share and subscriber count.

III. COMPARING THE REGRESSION RESULTS TO THE JUSTICE DEPARTMENT'S EXPERT TESTIMONY

As previously noted, Dr. Carl Shapiro's expert testimony laid out several estimations, all of which suggested that the Justice Department believed

the merger between AT&T and Time Warner would harm consumers specifically in the television provider market.[92] The fear that the new merged firm would withhold Time Warner products from DirecTV's competitors drove the case, and the Justice Department assumed that this added leverage to DirecTV's already strong position in the television provider market would allow the company to gain subscribers relative to its rivals while increasing rivals' costs assuming more market share.[93]

Judge Leon in his opinion pointed out that this fear was unwarranted given the industry's dynamic nature to endure constant change and technological development.[94] He, along with the expert team for AT&T-Time Warner, noted that this industry's nature would prevent a stall in innovation as predicted by the Justice Department.[95] The case was decided for AT&T-Time Warner, allowing the vertical merger to consummate without fear of anticompetitive effect.[96]

The data analyzed in this paper show that it is more likely the defense and Judge Leon correctly assessed that the changing technological environment in the telecommunications industry as a whole, as well as the consistency of long-term contracts, were sufficient to prevent any serious threat to the rivals of DirecTV. While this does not mean that the Justice Department's estimations and analyses were invalid or should not have been trusted, it suggests that industry is one of the key factors in antitrust cases. A vertical merger is already significantly less likely to be anticompetitive, as compared to a horizontal merger, and this study concludes that in a dynamic industry, a vertical merger is even less likely to have an anticompetitive influence, regardless of the firm's intention.

Interestingly, the Justice Department most likely did accurately predict AT&T-Time Warner's intentions.[97] Shortly after the firms merged, the newly established firm began bundling practices and changed products in a way that did not benefit most consumers.[98] The practices were intended to benefit the new firm through pushing customers towards new packages with higher pricing.[99] However, although the intentions may align with what the agency expert predicted, the actual data do not show that this conduct in practice actually benefitted DirecTV.

In antitrust and economic theory, a firm has no incentive to maintain anticompetitive practices if these practices do not end up being profitable or beneficial to the company. In the present case, the loss in market share and monthly subscribers indicates that this strategy implemented by DirecTV

---

92 Complaint, *supra* note 7, ¶¶ 3-7.
93 *Id.*
94 *AT&T*, 310 F. Supp. 3d at 171–73.
95 *Id.* at 173–74.
96 *Id.* at 254.
97 Complaint, *supra* note 7, ¶¶ 3-7; Neff, *supra* note 4, at 128–29.
98 Neff, *supra* note 4, at 128–29.
99 *Id.*

after the merger was not effective and did not give DirecTV any serious leverage against its rivals. Any harm to consumer welfare caused by the changing prices and product offerings from AT&T-Time Warner will therefore likely be undone as DirecTV adapts to recoup its losses.

Because AT&T and Time Warner are heavily involved in so many other industries outside of the television provider market, the vertical integration here almost acts as a further limit on anticompetitive conduct. AT&T and Time Warner must as one firm react to changes in a multitude of other dynamic industries, so DirecTV, while bundling, had to address trends and new innovations in markets outside of the television provider market. For example, the T-Mobile-Sprint horizontal merger in 2020, following the AT&T-Time Warner vertical merger, encouraged AT&T and refocus some of its resources back on the mobile phone industry and take resources away from DirecTV.[100] While these changes may not have been easily foreseeable by the Justice Department, the real evidence from this natural experiment demonstrate that it is necessary to acknowledge an industry's momentous nature when determining if anticompetitive conduct will be successful.

Overall, the data demonstrates that the Justice Department may have overreached in trying to block this vertical merger. The past forty years without much vertical merger intervention demonstrate that economics and antitrust regulation have found most vertical mergers to be generally free of serious anticompetitive effect.[101] Early on after the merger, the Justice Department seemed unphased by the merger announcement.[102] Whether the decision to attempt blocking the merger came from political influence or a changing economic perspective, the data suggest that the merger was not one worth blocking. The markets in which AT&T and Time Warner are involved worked to self-regulate and prevent DirecTV from imposing any actions that would have had anticompetitive effects. Therefore, it was likely an oversight that the Justice Department excluded the industries' changing natures from the bulk of the analyses, leading the agency expert to predict that DirecTV would benefit from the merger and harm consumer welfare.

IV. APPLYING THESE RESULTS TO FUTURE VERTICAL MERGER ANALYSIS

Going forward, the future of vertical merger regulation is not clear. As the FTC and the DOJ work on re-writing the Vertical Merger Guidelines, it is necessary that they consider the outcome of the AT&T-Time Warner challenge, as well as the observed results of the actual merger. Considering this case is especially important given how few vertical merger challenges there have been in the last forty years.

---

100 Boliek, *supra* note 5, at 951–52.

101 *Id.* at 940.

102 Bumgardner, *supra* note 4, at 36.

A few important lessons can be learned from the retrospective of this case. However, it is important to note that this is one case in a very specific industry, so this evidence is anecdotal. Every merger must be studied in a fact-specific manner that considers the industry, that industry's rigidity or fluidity, the market shares of rivals, and the influence of innovation and brand recognition in that industry. Regardless, the AT&T-Time Warner vertical merger was one of the only vertical merger challenges in recent history, so it is one of the only examples to draw from in future vertical merger trials.

The first suggestion this natural experiment presents is that a vertical merger faces market constraints on intentionally anticompetitive conduct from multiple product markets. In this case, it appears that while the Justice Department correctly assumed that AT&T-Time Warner would bundle products and push new products on customers, this bundling did not prove to aid DirecTV in gaining market power or monthly subscribers. Whether this is because of increasing innovation in the streaming market competing with the television provider markets or from increased pressure on AT&T to compete with the newly merged T-Mobile-Sprint, there were downward pressures on price from a variety of sources.[103] These separate product markets in which the vertically merged firm operated demonstrate pressures on multiple different fronts that the firm must coordinate. Moving resources between different sectors of the firm add to the balance required to ensure competition can still function. It appears that these pressures prevented DirecTV from successfully engaging in anticompetitive conduct, and even resulted in DirecTV losing market power and monthly subscribers.

The second suggestion is that the agencies should consider the level of innovation and change in each sub-industry of the future vertically merged firm. Following the first suggestion, it was critical in this case that the introduction of 5G in the mobile phone market prevented AT&T from distributing more resources to DirecTV.[104] While the Justice Department was concerned with DirecTV's potential leveraging of Time Warner's content, the agency's expert report did not sufficiently flag the budding innovations in AT&T's primary market.[105]

Although arguably the television provider market is as equally innovative as the mobile phone market when one includes streaming, the Justice Department excluded streaming from the market definition**,** making the defined market less dynamic than if streaming were included. However, this oversight should not have prevented the Justice Department from seeing that potential anticompetitive conduct could have been constrained by other dynamic industries, such as the mobile phone market and the expansion into 5G as a relevant factor. AT&T being so connected to the mobile phone market, even while it has branches in other large industries, seems to have effectively

103 *See generally* Boliek, *supra* note 5; Neff, *supra* note 4.

104 Boliek, *supra* note 5, at 951–52.

105 *See generally* Expert Report, United States v. AT&T, Inc., 310 F. Supp. 3d 161 (D.D.C. 2018) (No. 17 Civ. 2511).

worked to constrain its anticompetitive behavior, and therefore the agencies should look more closely at the innovative drive in each related industry to those relevant in a vertical merger case.

The third suggestion that the data reflect is that the more dynamic an industry is, the more likely it is that anticompetitive conduct will be unsuccessful. The different telecommunications sectors and industries are each booming in their own way. Had either the television provider or mobile phone market been less innovative, DirecTV under AT&T may have been more successful at creating bundling packages and increasing price for its rivals and its own consumers. While it is impossible to know how this natural experiment would have played out differently in other circumstances, it is important that the agencies recognize the level of innovation and productivity in these industries contributed greatly to the competitive constraints on anticompetitive conduct.

In other industries where not a lot of change or progress is occurring, such as shoe sales in *Brown Shoe*, new ideas or products are much less common and therefore anticompetitive behavior is harder to constrain.[106] However, in big technology, like the AT&T-Time Warner merger, new technology is constantly replacing old technology and tech giants become relics of the past, a process known as creative destruction.[107] This cycle has been seen with Facebook replacing Myspace, Netflix replacing Blockbuster, and overall the movement for people to replace television providers as a whole with streaming. These facts indicate that markets and industries that involve rapid changes and constantly evolving trends must be presumed as more capable of fighting anticompetitive conduct naturally.

Finally, the fourth suggestion is to not underestimate the effect of long-term contracts on an industry. Most likely, this was one of the reasons DirecTV was unable to affect rivals' costs significantly. Most competitors were already in longstanding contracts with Time Warner, and the change in ownership did not affect Time Warner's current contracts.[108] Judge Leon himself explained that this lack of mobility from contracts would likely keep rivals' costs from shifting too much.[109] The fact that the data shows DirecTV's rivals were actually more successful in acquiring market share and subscribers after the merger points to the fact that this estimation by the agency was not particularly accurate, and in the future, long-term contracts should be weighed more heavily by the agencies as factors affecting a firm's ability to behave anticompetitively.

---

106 *See Brown Shoe*, 370 U.S. at 334.

107 Michael D. Goldman & Eileen M. Filliben, *Corporate Governance: Current Trends And Likely Developments For The Twenty-First Century*, 25 DEL. J. CORP. L. 683, 684–85 (2000).

108 Keyte, *supra* note 42, at 20–21.

109 *AT&T*, 310 F. Supp. 3d at 200–01.

CONCLUSION

There is much to be learned by studying the past. Being the first major vertical merger challenge from the agencies in over forty years, the AT&T-Time Warner merger is the primary contemporary example of a vertical merger litigation for modern regulators and antitrust defense teams. The case brought vertical mergers into the center of political antitrust debate.

With new antitrust ideologies and a shift in regulatory policy, as well as the re-writing of the Vertical Merger Guidelines, vertical merger regulation is currently unsettled. To restabilize regulatory policy regarding vertical mergers, the agencies must begin with looking at actual evidence and natural experiments involving vertical mergers from the past. Having an example of an agency challenge, supported by expert testimony anticipating anticompetitive effects that was followed by results that do not support that claim demonstrate that there is still much to learn about vertical mergers and their effect on competition within their particular industries.

While all economic and statistical models are extremely difficult to use for predictive measures, regulators can always learn and adapt to new observations. By incorporating the level of innovation involved in an industry, as well as how the industries of a vertical merger are integrated and the long-term contracts that the merging firms are engaged in, the agencies are one step closer to having a fuller understanding of vertical merger effects. Hopefully, the AT&T-Time Warner merger can provide that example and become the blueprint for future vertical merger analysis, giving a better understanding of what indicators suggest that a vertical merger is unlikely to have anticompetitive effects.

AGENTS AS UNDERWRITERS: INCORPORATING AGENCY LAW TO MAKE SECTION 11 OF THE SECURITIES ACT BETTER APPLICABLE TO NON-TRADITIONAL IPOS

Ethan McCarthy

INTRODUCTION

In the run up to the Great Depression, the Dow Jones Industrial Average ("DJIA") fell over twelve percent on October 28, 1929.[1] The next day, known as Black Tuesday, it lost another almost twelve percent.[2] In total, the DJIA would drop eighty-nine percent during the Stock Market Crash of 1929.[3] The federal government responded to the crash and Great Depression by implementing increased regulation, including enacting the Securities Act of 1933 (Securities Act).[4]

The Securities Act, among other things, regulates the initial offering of securities to the public.[5] The legislation's purpose was to provide greater protection for consumers.[6] As part of this regulatory framework, those issuing securities must file a registration statement providing the information necessary for potential investors to properly evaluate the security.[7] The registration statement is filed with the Securities and Exchange Commission (SEC).[8] Part of this registration is a prospectus that is provided to investors.[9] Section 11 of the Securities Act holds strictly liable those that sign the registration or contribute to the registration statement as a director, expert, or underwriter if the statement "contained an untrue statement of a material fact or omitted to

---

1 Gary Richardson, Alejandro Komai, Michael Gou & Daniel Park, *Stock Market Crash of 1929*, FED. RES. HISTORY (Nov. 22. 2013), https://www.federalreservehistory.org/essays/stock-market-crash-of-1929#:~:text=On%20Black%20Monday%2C%20October%2028%2C%201929%2C%20the%20Dow%20declined,almost%20half%20of%20its%20value.

2 *Id.*

3 André Douglas Pond Cummings, *"Ain't No Glory in Pain": How the 1994 Republican Revolution and the Private Securities Litigation Reform Act Contributed to the Collapse of the United States Capital Markets*, 83 NEB. L. REV. 979, 990 (2005).

4 Elisabeth Keller & Gregory A. Gehlmann, *Introductory Comment: A Historical Introduction to the Securities Act of 1933 and the Securities Exchange Act of 1934*, 49 OHIO ST. L.J. 329, 329 (1988).

5 *Id.* at 330.

6 *Id.*

7 *Id.*

8 Tayler Tanner, *Spotify's Direct Listing and Foreign Private Issuers: Protecting Investors When Foreign Private Issuers List on a U.S. Exchange But Not on Their Home Exchange*, 2019 B.Y.U. L. REV. 573, 579 (2019).

9 *Id.*

state a material fact required to be stated therein or necessary to make the statements therein not misleading."[10]

The Securities Act established the traditional initial public offering (IPO) process through which a company may go public.[11] During a standard IPO, a company will employ lawyers and investment bankers to develop the comprehensive financial reports that are required as part of the registration statement.[12] Investment bankers also market the shares, underwrite the offering, and usually take ownership of the shares for sale, which is called a firm-commitment underwriting.[13] The Securities Act's purpose in developing such a strenuous and time-consuming process is to best ensure that investors are provided all information necessary to safely invest in a newly offered security.[14]

Recently, however, companies have more frequently experimented with alternative means of going public,[15] and Section 11's applicability in these cases may not be clear. In 2018, Spotify went public on the New York Stock Exchange (NYSE) through a direct listing.[16] Additionally, in 2020, Special Purpose Acquisition Companies (SPACs) "accounted for more than 50% of new publicly listed U.S. companies."[17] The increased use of direct listings and SPACs have created two key issues with applying Section 11 liability. First, potential plaintiffs have greater difficulty establishing Section 11 standing when bringing claims arising from nontraditional IPO methods.[18] Second, it is not clear who can be liable as an "underwriter" under Section 11.[19] To address these two issues, Congress should amend Section 11 to include agency law definitions and concepts and to provide an easier means for plaintiffs to establish standing for their Section 11 claims. Incorporating agency law into the statute would enable plaintiffs to more easily establish standing while better defining those that may be liable for misstatements or omissions in a registration statement. Thus, this amended Section 11 would achieve both goals of protecting consumers while clearly identifying who may be held liable and when under the statute.

---

[10] 15 U.S.C.A. § 77k; *see also* Elisabeth Keller & Gregory A. Gehlmann, *supra* note 4, at 345–46.

[11] Tayler Tanner, *supra* note 8, at 579.

[12] *Id.* at 579–80.

[13] *Id.* at 580.

[14] *Id.* at 595.

[15] Deloitte, *A CFO's Guide to Traditional and Nontraditional IPOs*, WALL ST. J.: CFO J. (Aug. 17, 2022, 3:00 PM), https://deloitte.wsj.com/articles/a-cfos-guide-to-traditional-and-nontraditional-ipos-01660679589.

[16] Brent J. Horton, *Spotify's Direct Listing: Is It a Recipe for Gatekeeper Failure?*, 72 SMU L. REV. 177, 179 (2019).

[17] Max H. Bazerman & Paresh Patel, *SPACs: What You Need to Know*, July–Aug. 2021 HARV. BUS. REV. ¶ 1, ¶ 1 (2021), https://hbr.org/2021/07/spacs-what-you-need-to-know.

[18] Anat Alon-Beck, Robert Rapp & John Livingstone, *Investment Bankers as Underwriters—Barbarians or Gatekeepers? A Response to Brent Horton on Direct Listings*, 73 SMU L. REV. F. 251, 255 (2020).

[19] Benjamin J. Nickerson, *The Underlying Underwriter: An Analysis of the Spotify Direct Listing*, 86 U. CHI. L. REV. 985, 989 (2019).

Agency law enables plaintiffs to still bring claims against those who delegate authority to another but still retain some control over that other person's actions. In the context of initial offerings of securities, agency law principles can be used to classify certain parties as agents based solely on their relationship to the company issuing the securities. Including an agency law component in Section 11 would allow the statute to identify parties liable for registration statement errors or omissions without strictly needing to rely on the term "underwriter." Section 11 would thus be more flexible and more easily applicable to the changing landscape of firm public offerings.

Part One of this comment will explain how Section 11 regulates traditional IPOs, discuss increasingly popular IPO alternatives and the problems they pose regarding Section 11, and provide a brief overview of agency law. Part Two will provide a proposed amended Section 11 and analyze how an agency law framework and relaxed standing requirements can address the issues current Section 11 has with regulating non-traditional IPOs.

I. BACKGROUND

In analyzing the challenges that nontraditional IPOs present regarding Section 11, this section first details how Section 11 functions in the traditional IPO context. It then explains the direct listing and SPAC processes. Finally, this section provides an overview of the agency law provisions relevant to the proposed solution.

A. *Traditional IPOs*

The traditional IPO process is time-consuming and costly.[20] The Securities Act requires that a firm issuing securities make a "full and fair disclosure of information" regarding the issuance so that investors are protected.[21] This is accomplished through the registration requirement, which mandates the issuing company file a registration statement that includes all material information regarding the company and the securities being offered.[22] Section 11 of the Securities Act enables purchasers of the securities to bring claims against specified participants in the issuance, such as directors, partners, and underwriters, if the registration statement contains material misstatements or omissions.[23]

---

[20] Tayler Tanner, *supra* note 8, at 579.

[21] Omnicare, Inc. v. Laborers Dist. Council Const. Indus. Pension Fund, 575 U.S. 175, 178 (2015) (quoting Pinter v. Dahl, 486 U.S. 622, 646 (1988)).

[22] *Id.*

[23] *Id.* at 179; 15 U.S.C.A. § 77k(a).

Through an IPO, a company raises equity financing by making its stock publicly available for the first time by selling it through an exchange.[24] To do so, the Securities Act requires that the firm develop a prospectus and file a registration statement with the SEC.[25] The purpose of the registration statement is to provide all necessary information about the company to potential investors.[26] Thus, the registration statement must include a description of the company, the company's audited financial statements from previous years, and an analysis of potential risks that could affect the company.[27]

To develop this registration statement, firms employ investment banks and lawyers.[28] The investment bank underwrites the IPO and generally commits to purchasing the entirety of the stock offering, which is known as a firm-commitment underwriting.[29] As the underwriter, once the registration statement is filed with the SEC and has been reviewed by the agency, the investment bank then markets the company's shares to large institutional investors, developing the initial offering price in the process.[30] The shares can be sold once the registration statement becomes effective, meaning that the shares are now registered in accordance with the Securities Act.[31] After issuance, the company must continue to file specified disclosures with the SEC, such as the annual 10-K and quarterly 10-Q financial reports.[32]

In an IPO, the investment bank makes money from (1) the difference in the price for which it purchases the stock from the issuing firm and the price at which it sells the stock to institutional investors, and (2) the industry standard seven percent commission on the total offering amount.[33] As the underwriter, the bank may also buy back shares as a price stabilization mechanism, as this reduces the share supply and thus helps prop-up the price.[34] Additionally, the investment bank imposes a lockup period on the issuing company's employees and previous investors.[35] These insiders may not sell their shares in the issuing company for a specified period after the IPO, typically six months.[36] This is another supply control measure.[37] Importantly, these shares

---

24 Benjamin J. Nickerson, *supra* note 19, at 990.

25 *Id.* at 991; 15 U.S.C.A. § 77e.

26 Benjamin J. Nickerson, *supra* note 19, at 991.

27 Tayler Tanner, *supra* note 8, at 579–80.

28 *Id.* at 580.

29 *Id.*

30 Benjamin J. Nickerson, *supra* note 19, at 991–92, 992 n.25 (quoting Securities and Exchange Commission Release No 33-8565, 70 Fed. Reg. 19672, 19674-75 (Apr. 13, 2005) explaining this process known as "book building.").

31 Tayler Tanner, *supra* note 8, at 581.

32 *Id.* at 581–82.

33 *Id.* at 580.

34 Benjamin J. Nickerson, *supra* note 19, at 993.

35 *Id.* at 993.

36 *Id.* at 993–94.

37 *Id.*

owned by the issuer's management and early investors are "unregistered" because they are not part of the offering to which the registrations statement applies.[38]

The investment bank is liable as an underwriter under Section 11.[39] Purchasers of the securities may therefore hold the bank strictly liable if the registration statement "contained an untrue statement of a material fact or omitted to state a material fact," making the statement misleading.[40] During the book building process, the investment bank must also comply with Section 5 of the Securities Act and SEC rules that regulate what actions it can take while the registration statement is not yet effective.[41]

B. *The Tracing Requirement*

Section 11 provides a cause of action for purchasers of a security when the registration statement under which "such security" is registered contains material misstatements or omissions.[42] In *Barnes v. Osofsky*, the Second Circuit interpreted the meaning of this phrase "such security" in § 77k(a).[43] In *Barnes*, Aileen, a women's sportswear company, publicly offered shares under a registration statement in 1963.[44] After the price fell upon the company missing its sales projections, purchasers of the company's stock comprising this sale brought claims under Section 11 arguing that the company failed to disclose known risks in the registration statement.[45] These claimants, however, could not establish that all their purchased shares were those issued under the registration statement for this offering and not shares previously issued under a different registration.[46]

The Second Circuit faced the question of whether to read the phrase "such security" to provide a cause of action for anyone purchasing a security of the same type as that issued under the registration statement, or for only those that purchase the specific shares issued under the registration statement.[47] The court, in its opinion authored by Judge Friendly, adopted the later, narrower reading.[48] The court stated that this is the more natural reading of Section 11's language, and noted that the statute imposes "stringent

38 *See* Pirani v. Slack Techs., 13 F.4th 940, 943 (9th Cir. 2021).

39 *Id.* at 990; 15 U.S.C.A. § 77k(a)(5).

40 15 U.S.C.A. § 77k(a); *see also* Benjamin J. Nickerson, *supra* note 19, at 990.

41 Tayler Tanner, *supra* note 8, at 581 (discussing these "gun-jumping rules" in place during a "quiet period" before there is an effective registration statement).

42 § 77k(a).

43 Barnes v. Osofsky, 373 F.2d 269, 271 (2d Cir. 1967).

44 *Id.* at 270.

45 *Id.*

46 *Id.* at 271.

47 *Id.*

48 *Id.* at 272.

penalties" on those found liable under it.[49] The other circuits agreed with Judge Friendly's interpretation of the language.[50]

The Ninth Circuit recently examined Section 11's tracing requirement in the context of a direct listing.[51] *Pirani* arises from the technology firm Slack's direct listing on the NYSE, where the plaintiff, Pirani, could not prove whether he purchased registered or unregistered shares.[52] The court ultimately had to decide whether the plaintiff had standing to bring the suit given the "such security" phrase in § 77k(a) and concluded that Pirani did have standing.[53]

As previously mentioned, direct listings, unlike IPOs, do not include a lockup period.[54] Therefore, when Slack's direct listing was executed on June 20, 2019, the process made 118 million registered shares and 165 million unregistered shares available through the public exchange.[55] The court portrayed this as a case of first impression.[56] After doing so, the court held that the reading of "such security" adopted by *Barnes* was not applicable in the direct listing context.[57] The Ninth Circuit based this conclusion on the fact that both the registered and unregistered shares could be sold in the manner they were because of the registration statement, thus eliminating any traceability concerns.[58]

Judge Miller disagreed.[59] He did not believe that *Pirani* was a case of first impression and argued in dissent that the court should have followed the precedent of *Barnes*, meaning Pirani's claim would fail because he could not demonstrate that he purchased shares registered under the challenged statement.[60]

---

49 *Barnes*, 373 F.2d at 272.

50 *See e.g.*, Krim v. pcOrder.com, Inc., 402 F.3d 489, 499 (5th Cir. 2005); Lee v. Ernst & Young, LLP, 294 F.3d 969, 976 (8th Cir. 2002); APA Excelsior III L.P. v. Premiere Techs, Inc., 476 F.3d 1261, 1271 (11th Cir. 2001); *see also* Pirani v. Slack Techs., 13 F.4th 940, 952 (9th Cir. 2021) (listing cases from other circuits that apply the same reading to the phrase "such security" in Section 11).

51 *Pirani*, 13 F.4th at 946.

52 *Id.* at 943.

53 *Id.*

54 *Id.* at 944.

55 *Id.*

56 *Id.* at 946.

57 *Pirani*, 13 F.4th at 946–47.

58 *Id.* at 947.

59 *See id.* at 950 (Miller, J., dissenting).

60 *Id.* at 951–52.

C. *Direct Listings*

"Unicorns," private startups with a valuation exceeding one billion dollars,[61] have recently been delaying conducting IPOs because they now have more options to raise capital.[62] For these firms, the value of being able to sell stock to the public is the liquidity it provides, not the capital that it can obtain .[63] In fact, direct listings initially could not be used to raise capital.[64] The direct listing process provides the liquidity unicorns desire without diluting the current shareholders' control and with costs lower than those associated with a traditional IPO.

Spotify, then a unicorn firm, was the first major company to go public through a direct listing and did so in 2018.[65] In 2020, the Securities and Exchange Commission (SEC) permitted the NYSE to begin allowing companies to raise capital by selling shares through direct listings, which could lead to increased use of direct listings in the future.[66]

In a direct listing, the company (directly) lists its stock on an exchange without going through the book building process that would result in specific institutional investors being the initial purchasers and establishing a determined price.[67] Once the shares are listed on the exchange, the market ultimately determines the price.[68] This also means that there can be none of the price stabilization measures underwriters may use in the IPO context.[69] The key benefit of a direct listing is the lower cost associated with it than with a traditional IPO.[70] In addition to there being no book building, the "roadshow" is often shorter, and there is no lockup period.[71]

When a company conducts a direct listing, it still must follow the same regulations for disclosing information as with an IPO.[72] Thus, firms conducting a direct listing still must file a registration statement with the SEC.[73]

---

[61] James Chen, *Unicorn: What It Means in Investing, With Examples*, INVESTOPEDIA (May 31, 2022), https://www.investopedia.com/terms/u/unicorn.asp; Benjamin J. Nickerson, *supra* note 19, at 987 n.1.

[62] Anat Alon-Beck, Robert Rapp & John Livingstone, *supra* note 18, at 259 (noting that companies backed by venture capital are now on average remaining public for in excess of eleven years as opposed to previously going public within four years).

[63] *Id.* at 260.

[64] *Id.* at 261.

[65] Brent J. Horton, *supra* note 16, at 179.

[66] James J. Park, *Investor Protection in an Age of Entrepreneurship*, 12 HARV. BUS. L. REV. 107, 138 (2022).

[67] *Id.*

[68] *Id.*

[69] Benjamin J. Nickerson, *supra* note 19, at 994.

[70] *Id.*

[71] *Id.*

[72] Tayler Tanner, *supra* note 8, at 583.

[73] Benjamin J. Nickerson, *supra* note 19, at 994.

Direct listings, however, do not involve a typical "underwriter," and the issuing company instead works with "financial advisors." [74] It is unclear whether these financial advisors would qualify as "underwriters" under current Section 11.[75]

D. *SPACs*

A Special Purpose Acquisition Company (SPAC) is a company that goes public through its own IPO for the purpose of later acquiring a private company that becomes public itself through this acquisition.[76] When the SPAC merges with the targeted private company, the target company is now essentially public as an asset of the SPAC.[77] The entity that initially forms the SPAC and takes part in its IPO is called the "sponsor" and is generally a limited liability company.[78] The sponsor selects the SPAC's officers and directors, who are generally the people that own the sponsor.[79] During the IPO, the SPAC sells an offering consisting of a share, a warrant, and possibly a right to acquire a fraction of a share.[80] This offering is called a "unit," and the standard is to price them at $10.[81] The funds received from the SPAC's IPO are placed in a trust.[82] A SPAC's articles of incorporation typically provide that the SPAC has two years to complete a merger, or the SPAC must liquidate the funds in the trust and distribute the funds to the shareholders.[83] When the SPAC proposes a merger with a target private company, a shareholder can redeem her share for the initial ten-dollar price plus interest, instead of retaining the shares and retaining the equity interest in what is essentially the previously private company post-merger.[84]

The use of SPACs ballooned in 2020 and 2021.[85] This apparent bubble, though, burst in the Spring of 2021..[86] There is still debate over how popular SPACs will be in the future.[87] Generally, the number of SPACs each year will depend on market situations; the same is true for traditional IPOs. Regardless,

---

74 James J. Park, *supra* note 66, at 111, 138.

75 *Compare* Benjamin J. Nickerson, *supra* note 19, at 986 *with* James J. Park, *supra* note 66, at 140.

76 Michael Klausner, Michael Ohlrogge & Emily Ruan, *A Sober Look at SPACs*, 39 YALE J. ON REG. 228, 235 (2022).

77 James J. Park, *supra* note 66, at 111, 133.

78 Michael Klausner, Michael Ohlrogge & Emily Ruan, *supra* note 76, at 236.

79 *Id.*

80 *Id.*

81 *Id.* at 247.

82 *Id.* at 237.

83 *Id.*

84 Michael Klausner, Michael Ohlrogge & Emily Ruan, *supra* note 76, at 237.

85 Max H. Bazerman & Paresh Patel, *supra* note 17, at ¶ 1.

86 Michael Klausner, Michael Ohlrogge & Emily Ruan, *supra* note 76, at 231.

87 *See* Max H. Bazerman & Paresh Patel, *supra* note 17, at ¶ 4.

the current use of SPACs has helped identify weaknesses in Section 11, which will be helpful in developing a framework that can be applied to an array of IPO alternatives beyond just SPACs.

There are three main challenges to applying current Section 11 in the context of SPACs. The first is that some SPACs issue unregistered shares to their target shareholders as opposed to issuing registered shares, sometimes preventing shareholders from being able to bring claims under Section 11.[88] Second, establishing standing under Section 11 after SPAC shares have been sold to secondary parties is challenging given the tracing requirement.[89] Lastly, the SPAC process does not use underwriters, so there may be fewer parties from which plaintiffs can try to recover damages.[90]

E. *Agency Law*

Agency law applies when one person employs the assistance of another to achieve something for the first person's benefit.[91] The person who employs another person is the principal, while the person working on behalf of the principal is the agent.[92] The purpose of agency law is to properly categorize parties to a transaction or agreement completed using an agent or agents, so responsibilities and potential liabilities derived from these arrangements are correctly assigned.[93] Generally, agency law works to hold the principal responsible for her agent's actions done in furtherance of the principal's instructions and purpose.[94]

An agency relationship is a fiduciary relationship established when (1) one person, the principal, "manifests assent" to have another person, the agent, act on behalf of the principal, (2) the principal has some control over the agent, and (3) the agent "manifests assent" to act on principal's behalf.[95] A principal is liable for the agent's actions when the agent acts with express actual authority, implied actual authority, or apparent authority.[96] An agent has actual authority to act to bind the principal when the agent (1) reasonably believes that the principal wishes the agent to act in that way, and (2) that belief is based on a manifestation by the principal to the agent.[97] The agent has express actual authority when the principal instructs the agent to perform

88 Michael Klausner, Michael Ohlrogge & Emily Ruan, *supra* note 76, at 285–86.

89 *Id.* at 286.

90 *Id.*

91 *See* RESTATEMENT (THIRD) OF AGENCY § 1.01 (2006).

92 *See id.*

93 *See* Paula J. Dalley, *A Theory of Agency Law*, 72 U. PITT. L. REV. 495, 497 (2011).

94 *Id.*

95 RESTATEMENT (THIRD) OF AGENCY § 1.01 (2006).

96 Salyers v. Metro. Life Ins. Co., 871 F.3d 934, 940 (9th Cir. 2017) (citing RESTATEMENT (THIRD) OF AGENCY § 2 intro. note (2006)).

97 RESTATEMENT (THIRD) OF AGENCY § 2.01 (2006).

the specific act.[98] Additionally, the agent has the implied actual authority to take the steps that reasonably further the actions the principal expressly instructed the agent to complete or obtain.[99] An agent has apparent authority to bind a principal through the agent's actions when (1) a third party "reasonably believes the actor has authority to act on behalf of the principal," and (2) that belief derives from a manifestation by the principal.[100]

Agency law "restore[s] the status quo" of obligations and liability when someone employs an agent when doing business with another.[101] This is done so that a principal cannot avoid liability to which she would otherwise be subject solely because he chose to use an agent.[102] This comment argues that amending Section 11 to include these definitions would more clearly identify the parties liable under the provision and enable plaintiffs to more easily establish standing by tracing harm through specified agency relationships. In this context, the focus is holding people that may be defined as agents liable for the errors and omissions in a registration statement. Thus, the question that will be answered is how certain participants that aid a corporation in going public, like the financial advisors in direct listings, can be defined as agents of that corporation for that purpose.

What makes agency law so useful for addressing this problem is its uniformity and simplicity. The Supreme Court cites the Restatement of Agency in decisions when needing to analyze agency law.[103] Furthermore, the definitions of principal and agent can be applied in any situation in which a company enlists the aid of others in going public. Courts could therefore apply Section 11 more easily regardless of the specific methods companies may use to go public because they are not restricted by how the statute may define a term like "underwriter" but can apply the developed agency common law. In turn, Congress would not need to devise a liability-tracing framework from scratch and could refer to an established area of law. The fiduciary relationship between a principal and agent and the different types of authority under which an agent may act on behalf of the principal are relatively straightforward and clearly defined in the Restatement. Individuals and firms could thus more easily determine what was required of them in their roles in a going public process and what liability they may face. In general, connecting agency law to Section 11 creates a more workable framework that courts may apply and promotes fairness by providing greater clarity to parties to securities transactions regulated by Section 11.

---

98 Salyers, 871 F.3d at 940 (citing NLRB v. Dist. Council Iron Workers State Cal. & Vicinity, 124 F.3d 1094, 1098 (9th Cir. 1997)).

99 *Id.*

100 RESTATEMENT (THIRD) OF AGENCY § 2.03 (2006).

101 Paula J. Dalley, *supra* note 93, at 497.

102 *Id.* at 497–98. Dalley refers to this as the "cost-benefit internalization theory."

103 *Salyers*, 871 F.3d at 939–40.

F. *Strict Liability and the Need for Clarity*

Section 11 imposes strict liability for misstatements or omissions in a registration statement.[104] Likewise, Judge Friendly noted that the "stringent penalties" imposed on those who violate Section 11.[105] Strict liability is relevant in the criminal law context, and Congress should consider criminal law's principles regarding strict liability when amending Section 1. In criminal law, statutes that criminalize a certain behavior per se without any mens rea requirement, meaning laws imposing strict liability, are generally disfavored.[106] Additionally, the Supreme Court generally refrains from applying strict liability when criminal statutes contain ambiguous provisions.[107]

When considering whether someone may be liable under Section 11, the question arising from the statute's ambiguity differs from that typically asked in the criminal law context. Criminal laws, generally, prohibit specified conduct and apply to everyone. Contrastingly, the uncertainty regarding Section 11 is not what conduct is prohibited but who is liable under the statute. Those that contribute to a company going public should not be required to guess as to whether their level of involvement resultingly qualifies them as underwriters responsible for filing a complete and accurate registration statement.

G. *Rule 144*

SEC Rule 144 provides that, in specified circumstances, an owner of a company's shares of stock may sell those shares to the public without the company first needing to register these shares.[108] When an underwriter sells the securities it owns, the sale is treated like a public offering itself because of the extensive role the underwriter plays in introducing the securities to the public markets.[109] Furthermore, Section 4(1) of the Securities Act creates an exemption for shareholders that are not underwriters that permits these shareholders to sell their shares without the company first needing to register the shares.[110]

Section 2(a)(11) of the Securities Act defines the term "underwriter", and the definition is notably broad.[111] The broad definition of the term

---

104 Elisabeth Keller & Gregory A. Gehlmann, *supra* note 4, at 345–46.

105 *Barnes*, 373 F.2d at 272.

106 Note, *The New Rule of Lenity*, 119 HARV. L. REV. 2420, 2432 (2006).

107 Leonid (Lenny) Traps, *"Knowingly" Ignorant: Mens Rea Distribution in Federal Criminal law After Flores-Figueroa*, 112 COLUM. L. REV. 628, 628 (2012).

108 *Pirani*, 13 F.4th at 944; 17 C.F.R. § 230.144.

109 Marina Petrova, *Capital Formation for Internet Companies: Why Facebook Stayed Private for So Long and What that Means for Investors*, 12 J. BUS. & SEC. L. 305, 317 (2012).

110 *Id.*; 15 U.S.C.A. § 77(a)(7), (d).

111 17 C.F.R. § 230.144; 15 U.S.C.A. § 77b (defining an underwriter as "any person who has purchased from an issuer with a view to, or offers or sells for an issuer in connection with, the distribution of

"underwriter" would seemingly include any investor who purchased the stock with the intention of then reselling, or "distributing," the shares for a profit, imposing additional costs on purchasing these shares and therefore reducing the demand for them.[112] In response to this market issue, the SEC promulgated Rule 144.[113] Rule 144 provides an analytical framework to establish that a particular investor is not an underwriter by determining that the investor is "not engaged in distribution."[114]

The first question under Rule 144 is whether the company that issued the securities is required to file reports in accordance with the Exchange Act.[115] In Rule 144, The SEC included fewer requirements for selling unrecorded shares of a company that files public reports because investors have clear access to the financial information regarding the company and the securities themselves.[116]

The second question is whether the securities' owner is an "affiliate" of the issuer; Rule 144 includes more requirements that an affiliate must satisfy before she can sell her shares than are imposed on non-affiliates.[117] A person is an affiliate of an issuer if the issuer has some control over her or she has some control over the issuer.[118] This term therefore includes members of the issuer's board of directors, the issuer's executives, and the issuer's stockholders that have voting rights.[119] Affiliates may only sell their shares under Rule 144 if the company publicly discloses the financial information required by the Exchange Act.[120] Additionally, they must file a Form 144 with the SEC when selling the securities, and Rule 144 limits the number of securities they can sell and how widely they can advertise their sale.[121] Lastly, affiliates cannot sell their securities until six months after acquiring them if the company reports financial disclosures or until after one year if the company does not formally report.[122] Non-affiliates cannot sell their securities until after they have owned them for one year but face none of the other limitations imposed on affiliates.[123]

As was seen in *Pirani*, these provisions become relevant in the direct listing context because some shares being sold upon the filing of a

---

any security, or participates or has a direct or indirect participation in any such undertaking, or participates or has a participation in the direct or indirect underwriting of any such undertaking").

112 Petrova, *supra* note 109, at 318.

113 *Id.*

114 17 C.F.R. § 230.144, Preliminary Note; Petrova, *supra* note 109, at 318.

115 Petrova, *supra* note 109, at 318.

116 *Id.*

117 *Id.* at 318–319.

118 *Id.* at fn. 57; 17 C.F.R. § 230.144(a)(1).

119 Petrova, *supra* note, 109 at 318.

120 *Id.* at 319.

121 *Id.*

122 *Id.*

123 *Id.*

registration statement will be registered shares while others may be unregistered shares.[124] The majority in *Pirani* argued that both Slack's registered and unregistered shares could only be sold publicly as a result of the registration statement filed in accordance with the direct listing.[125] Based on this assertion, the majority ruled that both those that purchased registered shares and those that purchased unregistered shares (and therefore those plaintiffs that could not determine which category of shares they purchased) could bring claims under Section 11 relating to that registration statement.[126]

Under Rule 144 and the provisions referenced in its preliminary note, however, unregistered shares may sometimes lawfully be sold publicly without the need for a registration statement.[127] Thus, concluding that a registration statement filed in the course of a direct listing alone enables all the company's shares to be sold publicly may not always be correct. In a situation where a plaintiff knows that she purchased unregistered shares, it also necessitates determining if the seller of those unregistered shares satisfies the requirements of Rule 144 that would permit the seller to sell those shares with the company filing a registration statement. If the seller does not meet these requirements, she would only be permitted to sell those shares under the company's registration statement, so the purchaser would have standing to bring claims related to errors in the registration statement pursuant to Section 11. Congress, however, can and should amend Section 11 to make this distinction in unregistered shares clearer and to provide standing for purchasers of unregistered shares that may have been sold without the filing of a registration statement but are nonetheless sold in such proximity with the filing of a registration statement that the purchaser would have reasonably relied on the registration statement's assertions when evaluating the purchase of the securities.

II. ANALYSIS

Section 11's current language does not clearly define who can be held liable as an underwriter in nontraditional IPO settings, and plaintiffs face troublesome difficulty in establishing standing under this provision in these settings. Congress should amend Section 11 to include an agency law framework to clarify liability and enable plaintiffs to establish standing more easily.

Because Section 11 necessarily imposes strict liability for consumer protection, its language must be clear as to who is liable under it. This agency law framework solution provides that. Establishing agency relationships

124 *See Pirani*, 13 F.4th at 944.
125 *Id.* at 947.
126 *Id.* at 943.
127 *Id.* at 944; 17 C.F.R. § 230.144.

within the alternative processes of going public also helps plaintiffs show standing, which is a problem under current Section 11 language. Section 11 only offers a remedy to a plaintiff who relies on a misleading registration statement filed to sell a security and purchases "such security."[128] In *Pirani*, the Ninth Circuit held that the plaintiff had standing even though he could not prove that he had purchased registered shares, and not unregistered shares, under Slack's direct listing.[129] This may be the desired result, but the current language of Section 11 does not support it. *Pirani* thus demonstrates that Section 11 should be amended to include the proposed agency framework to effectively protect investors. Lastly, this solution is a framework that can be used proactively if other nontraditional IPO methods arise.

A. *Proposed Additional Section 11 Framework*

Current Section 11 enumerates those that purchasers of securities registered under a registration statement may sue based on errors or omissions in the registration statement.[130] Congress should not alter or remove any of the current text within Section 11. The current statute applies liability in the traditional IPO context, and that understanding should not be changed. Therefore, the language of 15 U.S.C.A. § 77k(a)(1)-(5), including holding liable any "underwriter," should remain the same. Instead, a § 77k(a)(6) should be added. The relevant definitions of an agency relationship and the different types of authority could also be added to the amended Securities Act, or the statute could simply provide that courts should interpret agency law terms and concepts using the Restatement of Agency.

The proposed amended Section 11 would thus read in relevant part:

> **(a) Persons possessing cause of action; persons liable**
> In case any part of the registration statement, when such part became effective, contained an untrue statement of a

---

[128] 15 U.S.C. § 77(k); *Pirani*, 13 F.4th at 946.

[129] *See Pirani*, 13 F.4th at 948–49.

[130] 15 U.S.C. § 77(k). The portion of this section that this comment proposes to amend reads:
(a) Persons possessing cause of action; persons liable
In case any part of the registration statement, when such part became effective, contained an untrue statement of a material fact or omitted to state a material fact required to be stated therein or necessary to make the statements therein not misleading, any person acquiring such security (unless it is proved that at the time of such acquisition he knew of such untruth or omission) may, either at law or in equity, in any court of competent jurisdiction, sue—
(1) every person who signed the registration statement;
(2) every person who was a director of (or person performing similar functions) or partner in the issuer at the time of the filing of the part of the registration statement with respect to which his liability is asserted;
…
(5) every underwriter with respect to such security.

> material fact or omitted to state a material fact required to be stated therein or necessary to make the statements therein not misleading, any person acquiring such security or, if in such temporal proximity to the registration statement becoming effective that justifies the person's reliance on the registration statement, the same type of security issued by the same issuer (unless it is proved that at the time of such acquisition he knew of such untruth or omission) may, either at law or in equity, in any court of competent jurisdiction, sue—
>
> (1) every person who signed the registration statement;
>
> (2) every person who was a director of (or person performing similarfunctions) or partner in the issuer at the time of the filing of the part of the registration statement with respect to which his liability is asserted;
>
> …
>
> (5) every underwriter with respect to such security;
>
> (6) every person who otherwise acts as an agent on behalf of the issuer having prepared or certified all or any portion of the registration statementwith respect to that portion or all the registration statement the agent prepared or certified.

This proposed version of Section 11 makes two substantial changes. The first is the additional language following the phrase "such security." The federal courts interpret that phrase as requiring plaintiffs to have purchased securities registered under that specific registration statement.[131] The additional language is meant to expand the number of plaintiffs who can successfully plead suits under Section 11 for material misstatements or omissions from registration statements. The second substantial change is the addition of sub-subsection (6). This is the agency law framework that provides that those that play certain roles in helping companies go public may still be held liable even if they do not fit the strict definition of "underwriter" in 15 U.S.C. § 77(k)(a)(5).

B. *The Effects on Standing*

The addition of the language following "such security" in 15 U.S.C. § 77(k)(a) resolves Section 11's standing issues by broadening what investors may effectively bring claims. Because potential plaintiffs currently must be able to prove that the securities they purchased were those specifically registered for public distribution under the registration statement, The Ninth Circuit reasoning is incorrect, or at least other courts including the Supreme

---

[131] *Barnes*, 373 F.2d at 271–72.

Court could conclude so. The best course of action is therefore for Congress to amend Section 11 to better protect potential plaintiffs.

Another relevant consideration is that owners of unregistered shares are sometimes permitted to sell these shares publicly when the shareholders meet certain conditions under SEC Rule 144.[132] The agency law framework also functions in this case. The premise of Rule 144 is to provide a method by which owners of unregistered shares may sell their shares to public buyers without needing the issuing company to first file a registration statement covering the shares.[133] If the firm conducts a direct listing, the person looking to sell the unregistered shares of that firm no longer needs to take those additional steps outlined in Rule 144. Thus, the direct listing's registration statement enables the public sale of these shares. This is essentially the controversial holding in *Pirani*.[134]

The Ninth Circuit, however, oversimplified this in *Pirani*. The Ninth Circuit held that the plaintiff had standing regardless of whether the shares he purchased were registered or unregistered because, in the context of a direct listing, the filed registration statement is what enabled both categories of shares to be sold publicly.[135] This may not always be true because certain shareholders may have met the requirements of Rule 144, which would have enabled these shareholders to sell the shares without needing Slack to first file the registration statement.[136] Plaintiffs would therefore not have standing even under *Pirani* because the broad holding in *Pirani* is that plaintiffs may sue whenever they purchase shares, even unregistered ones, that they could only purchase as a result of the registration statement becoming effective.[137] *Pirani* thus fails to provide standing for those that purchase unregistered shares that could have been sold under Rule 144, even if the potential plaintiffs purchased the shares the same day the registrations statement became active.

The Ninth Circuit in *Pirani* also limited its holding to the direct listing context, emphasizing current Section 11's inability to provide a framework that can be applied to other methods of going public. The Ninth Circuit stated that the question it was answering was "what does 'such security' mean under Section 11 in the context of a direct listing."[138] The holding thus suggests a court must treat the first instance of filing a registration statement for each new IPO alternative that may arise as a matter of first impression and that the court must determine how it is to apply the meaning of "such security" to each of the new methods individually. This would surely prevent firms from trying to conduct new means of going public beyond the traditional IPO

132 *See* 17 C.F.R. § 230.144A.
133 *See generally*, 17 C.F.R. § 230.144.
134 *Pirani*, 13 F.4th at 947.
135 *Id.* at 943.
136 17 C.F.R. § 230.144.
137 *Pirani*, 13 F.4th at 943.
138 *Id.* at 946.

because they and those individuals or entities that would attempt to help would face too great a risk by surrendering themselves to the unknown will of a court.

Applying this new language to the facts in *Pirani*, the plaintiff can establish standing under Section 11 without the court having to resort to the questionable application of the statute found in the Ninth Circuit's opinion. According to precedent, the meaning of "such security" in Section 11 is that a plaintiff may only bring a claim if she purchased shares specifically registered under the registration statement she is challenging. This definition has no room to include unregistered securities of any type. Moreover, the court could not have known whether the shares Pirani purchased, assuming they were unregistered, could not otherwise be sold under Rule 144. If they were unregistered shares that could be sold without Slack filing a registration statement, then *Pirani*'s holding is incorrect because the registration statement did not initially enable the shareholder to sell these unregistered shares.

To solve this issue, the proposed additional language to Section 11 considers the time a buyer purchases shares in relation to when the registration statement at issue in a matter becomes effective. This takes the focus away from the potentially overly formalistic tracing requirement and looks to what the investor reasonably expected when purchasing the securities. Imagine a situation where two people, (A) and (B), learn that a corporation, who has already publicly issued stock previously, is going to publicly issue new shares of its stock. The corporation will sell new shares under a registration statement filed for this offering, but a past investor, (X), who had previously purchased unregistered shares pursuant to Rule 144 decides he also wishes to sell his unregistered shares. On the day that the registration statement becomes effective, (A) and (B) both review the financial information made in the disclosure related to the issuing corporation and the shares themselves. Both decide to purchase shares the next day, but (A) purchases shares registered under the registration statement while (B) purchases unregistered shares from (X). Neither (A) nor (B) know what type of shares they purchased. Unfortunately for them, there was a material mistake in the registration statement both viewed, and the shares greatly decreased in value.

Under current Section 11, (A) would have standing to bring a claim because (A) purchased registered shares. (B) would not have standing to bring a claim even though (B) took the same steps in reviewing the registration statement as (A) but unknowingly purchased unregistered shares. This is the form of unfair outcome that can result from overly formalistic requirements. The proposed Section 11, however, would allow both (A) and (B) to establish standing because they both purchased shares in the corporation in reliance on the registration statement the corporation filed only a day before they bought their shares.

The proposed new Section 11 addresses the traceability problem by considering when an investor purchases securities in relation to a registration statement that covers them or that type of security. Purchasing securities

registered under the specific registration statement obviously still provides standing. The crucial portion of the new language is that even purchasing unregistered shares, regardless of whether these unregistered shares could otherwise be sold under Rule 144, would provide standing for a plaintiff to challenge the registration statement so long as the purchase was made within a reasonable time of the statement becoming effective. This would largely alleviate tracing issues because plaintiffs would only need to demonstrate that they purchased the same type of security as was covered by the registration statement and that the purchase was made within a time suggesting that the investor would have reasonably relied on the registration statement when evaluating the security.

This result may seem to negatively impact owners of unregistered shares, but most owners of unregistered shares are generally the issuing firm's management and early investors.[139] They therefore would likely have control over the registration statement over which they would potentially be held liable. The filing of a registration statement may also increase the value of their unregistered shares because more investors may be willing to purchase the shares with greater access to information and greater trading volume. The proposed additional language to follow "such security" ultimately better promotes consumer protection by allowing purchasers of certain securities to seek remedies based on reasonable reliance on a registration statement filed shortly before the time of purchase. Additionally, fewer plaintiffs will have their complaints dismissed simply because the plaintiffs could not accurately trace their shares to a specific registration statement.

Technology may develop that will enable the easy tracing of previous owners of specified securities.[140] This could alleviate the traceability issues hindering Section 11's usefulness.[141] That shares will one day be perfectly traceable is however in no way ensured or imminent. Amending Section 11 thus offers a solution that can be implemented now. Additionally, the new proposed Section 11 language would allow even plaintiffs who could trace their shares and determine that they were unregistered to still bring claims related to a registration statement in reasonable circumstances. This proposal better protects public market participants now and in the future.

C. *The Effects on Who Is Held Liable*

Adding an agency law framework to Section 11 in a new sub-subsection (e) addresses the concern about alternatives to the traditional IPO enabling those that contribute to a registration statement avoiding liability for any

---

139 *Id.* at 943.

140 *See generally* George S. Geis, *Traceable Shares and Corporate Law*, 113 NW. U. L. REV. 227 (2018).

141 *Id.* at 238–39.

errors or omissions the statement may have. IPO alternatives, like the direct listing and its use of a financial advisor, may not have an investment bank acting as an underwriter for the offering as there would be if the company were going public through a traditional IPO.[142] Current Section 11 imposes liability on underwriters but may not do so for financial advisors or others who do not strictly fit into the underwriter classification.[143] The concern is that investment banks or other participants in a non-traditional IPO will intentionally classify themselves as something other than underwriters to avoid potential liability under Section 11 and thus not have the appropriate incentive to conduct sufficient due diligence related to the offering.[144] This could then harm those who purchase securities because the filings pertaining to the offering would be lower quality and more likely to contain errors.[145]

Expanding Section 11 to include a provision providing that an issuer's agents who assist the issuer in registering or distributing securities ensures that harmed purchasers of the securities can recover from these agents for wrongdoing in their work on behalf of the issuer. Purchasers could rely on standard agency law to first establish that one of the parties, like a financial advisor, was in an agency relationship with the issuer. This requires establishing the three elements of an agency relationship, namely that (1) the principal manifests assent to have the agent act on her behalf, (2) the principal has some control over the agent, and (3) the agent manifests assent to act on principal's behalf.[146] The relationship between the issuing firm and the financial advisor in a direct listing satisfies these elements for an agency relationship. In a direct listing, financial advisors must value the company going public and certify this valuation to the NYSE.[147] The financial advisor also helps determine an initial selling price for the shares, helps file necessary forms, and creates investor presentations.[148] The issuing firm and the investment bank acting as financial advisor enter a contract with one another for this work, so both parties agree to enter the principal-agent relationship with one another. The issuing firm necessarily controls the financial advisor regarding drafting and submitting filings, creating marketing presentations, and ultimately establishing a price for the shares. Thus, all elements of an agency relationship are met. Financial advisors would therefore be liable under the amended Section 11 for mistakes or omissions in the portions of filings to which they contribute.

The greatest benefit of this added agency framework is that it can be readily applied to other IPO alternatives. The same analysis of agency

---

142 Brent J. Horton, *supra* note 16, at 195.

143 *See id.* at 206–07.

144 *See id.*

145 *See id.*

146 RESTATEMENT (THIRD) OF AGENCY § 1.01 (2006).

147 Brent J. Horton, *supra* note 16, at 195.

148 *Id.*

relationships done above for financial advisors in direct listings can be done for other parties in other contexts. It removes the concern that investment banks may skirt Section 11 liability simply by avoiding classifying themselves as underwriters because they would be acting as the issuer's agent under standard agency law. SPACs provide another example of this because they likewise do not use traditional underwriters.[149]

On the reverse side, a hypothetical can demonstrate that the amended Section 11 would not be overinclusive and improperly hold certain parties liable. In the traditional IPO and the direct listing, both the investment bank and the financial advisor play instrumental roles in developing the registration statement and ultimately enabling the shares to be sold to the public market. In contrast, imagine the issuing company hires a consulting firm to help identify potential risks the issuing company may face in a certain segment and draft an overview of this risk in its disclosure documents relating to a public offering of new securities. The Securities Act requires companies issuing securities to publicly disclose these risks.[150] Now imagine the registration statement contains a factual error relating to the firm's financial performance during the previous two years. An investor, in reliance on this error, purchases the company's securities and then wishes to recover damages when the security loses value and upon learning of this error. The desirable result would include the investor being able to recover from the issuing company, which would be easy enough under both the current and proposed Section 11. The desired result would also include that the consulting firm would only be liable for errors or omissions in the part of the registration statement to which it contributed, so it therefore should not be liable in this hypothetical.

Current Section 11 would desirably not make the consulting firm liable for errors in historical financial information. As long as the consulting firm did not take ownership of the shares registered under the statement, it would not be deemed an underwriter under Section 11.[151] Subsection 15 U.S.C.A. § 77k(a)(4) would make the consulting firm liable for the portion of the registration statement it completed and certified.[152] The concern would be that the introduction of agency law would then improperly make the consulting firm liable for this error to which it did not contribute. The proposed agency law framework, however, would maintain this desired result because the consulting firm would only be deemed an agent of the issuing company with respect to the actions the issuing company gave the consulting firm authority to complete.

The consulting firm would only have the actual authority to complete the actions the issuing company explicitly instructed it to complete or those

149 Michael Klausner, Michael Ohlrogge & Emily Ruan, *supra* note 76, at 286.
150 Benjamin J. Nickerson, *supra* note 19, at 988.
151 15 U.S.C.A. § 77k(a)(5).
152 15 U.S.C.A. § 77k(a)(4).

actions taken in furtherance of completing what the issuing company instructed.[153] In the hypothetical, that would only include the portion relating to the potential future risks because that is the only issue the company hired the consulting firm to address. The issuing company would have explicitly asked the consulting firm to analyze and summarize this specific risk and would not have asked the consulting firm to audit past company financial statements for accuracy. Moreover, there would be no implied actual authority for the consulting firm to audit and then reproduce past company financial reports into the registration statement because this would not contribute to the completion of the portion relating to future risk.

Apparent authority arises from manifestations made from the principal to the third party to the agency relationship that the third party reasonably believes indicate that the principal has given the agent the authority to act in a certain way.[154] In the present hypothetical, this would mean that the issuing company indicated to the investor who purchased the shares that the consulting firm had the authority to review and certify historical financial information in the registration statement and that the investor reasonably believed these indications. There would be no manifestations like this from the issuing company to the investor that could establish apparent authority. Certain professions may be hired in a limited capacity, and Section 11 acknowledges this regarding the lengthy process of completing a registration statement.[155] The issuing company would have its agreement with the consulting firm limit the scope of the consulting firm's work to the specified risk analysis. The issuing company could announce this limited agreement, but the company would not be making any indications that the consulting firm has authority to complete assignments on the issuing company's behalf that do not relate to analyzing certain future risks. The result would be no different than how accountants who contribute to only a portion of the registration statement are liable for errors or omissions in only that portion under current 15 U.S.C.A. § 77k(a)(4).

Any solution to the threat to consumer protection that an IPO alternative may pose must go beyond requiring the issuer to make additional disclosures. Most importantly, this practice would not establish a statutory framework that can be consistently applied to multiple methods of going public. It requires specific tailoring for each method. Thus, when new methods first appear, consumers may be left unprotected during these initial periods. Additionally, there is no clear way of determining what disclosures should be required or at what point the costs associated with increased disclosures imposed on companies exceed the benefit to consumers. A solution that only addresses the alternative methods one at a time is untenable, especially if further innovation in the area is expected.

---

153 RESTATEMENT (THIRD) OF AGENCY § 2.01 (2006).

154 RESTATEMENT (THIRD) OF AGENCY § 2.03 (2006).

155 *See* 15 U.S.C.A. § 77k(a)(4).

As was done in *Pirani*, an alternative solution could be to completely distinguish alternatives like direct listings from the traditional IPO.[156] The problem with this alternative is that Section 11 imposes strict liability on those found guilty of violating it. The statute should therefore clearly indicate what conduct it prohibits and when someone can be held liable under it to best promote fairness for those assisting firms in going public. If courts were to treat each variation to how firms may go public as matter of first impression, the risk to those who would play a similar role to that of the financial advisor in a direct listing would present a cost barrier too high to warrant that participation and would thus stifle innovation in the going public process. Adding the proposed agency framework would establish the clarity that is necessary for the imposition of strict liability to be appropriate while also achieving Section 11's goal of maintaining consumer protection.

A wholesale rejection of an IPO alternative like the direct listing is not necessary and is potentially harmful. Unicorns have not universally opted to use a direct listing as opposed to the traditional IPO. This is because the direct listing process, like all corporate endeavors, has its own risks and benefits. A company like Spotify may be in a better position to realize the benefits while another firm is not. That does not mean a firm like Spotify should be prohibited from making use of an innovative means of going public, so long as appropriate consumer protections remain in place.

The proposed agency law framework addition to Section 11 is best suited to addressing the need for a flexible procedure that can be used to ensure that investors are protected and that issuers and those that work with them understand what obligations they have and what consequences they face for not meeting those obligations. Much of the concern with applying Section 11 to non-traditional IPO procedures is the emphasis on the word "underwriter."[157] Expanding the range of parties plaintiffs may sue for flawed registration statements to include those that act as the issuer's agents provides a framework that can be applied to these relationships, regardless of how they specifically function or what terms people choose to describe them. Unlike in *Pirani*, courts will not need to decide whether they should reinterpret Section 11 when faced with a new IPO alternative. Agency law likewise provides clarity in a statute that imposes strict liability and uses a well-established area of law. This solution properly balances consumer protection, clarity, and proactivity. That is the goal that can and should be achieved through amending Section 11.

---

156 *Pirani*, 13 F.4th at 946.

157 *See* Brent J. Horton, *supra* note 16, at 206–07; Michael Klausner, Michael Ohlrogge & Emily Ruan, *supra* note 76, at 286.

CONCLUSION

Current Section 11 of the Securities Act does not properly define liability or provide standing when companies go public through nontraditional IPOs. The increase in popularity in these non-traditional IPOs has given rise to situations that demonstrate this. Direct listings and SPACs have been particularly popular in this space and pose obstacles to Section 11 achieving its purpose by complicating tracing issues for plaintiffs and moving away from the standard practice of employing underwriters.

Innovation and greater choice in how companies go public can be economically beneficial and foster growth. The proper consumer protections, however, must remain in place. Lack of this protection would cause damage to the public in the short term and erode confidence in the long run. Because Section 11 is a crucial provision for protection of securities purchasers, it is important that it achieves its consumer protection purpose while clearly specifying when someone is liable under it. Congress should therefore amend Section 11 to alleviate the challenges plaintiffs may face with needing to trace their shares to a specific registration statement outside of the IPO, and to ensure that entities, like investment banks that help companies go public, cannot skirt liability under Section 11.

Incorporating an agency law framework into Section 11 and permitting standing to also be established by demonstrating the purchase of securities around the time a registration statement covering that type of security becomes effective would alleviate these problems and can be broadly applied to different means by which a company may go public. The agency law framework would more clearly identify who is liable for misstatements or omissions in a registration statement and allow potential plaintiffs to establish standing more easily by tracing liability through agency relationships. Adding additional language after the mention of "such security" in Section 11 to include a temporal proximity consideration would likewise allow plaintiffs to more easily establish standing. This proposed amended Section 11 is better than other solutions because it can be applied in multiple situations and helps establish the clarity that should be in place when a statute imposes strict liability. Thus, Congress amending Section 11 in this manner would help achieve the Security Act's goal of consumer protection in the financial markets.

A POLITICAL QUESTION SOLUTION TO THE ALARMING RISE OF NATIONWIDE INJUNCTIONS

Daniel Perez

INTRODUCTION

Legal watchers and scholars have paid more attention to nationwide injunctions in recent years because district courts are issuing them in remarkable numbers,[1] especially given the fact that courts denied them entirely as a remedy for much of our country's history.[2] Nationwide, or universal,[3] injunctions have surged dramatically in the last half-century. An injunction is a "court order commanding or preventing an action."[4] An injunction becomes nationwide when it enjoins the federal government from enforcing a policy against the particular plaintiffs and non-plaintiffs alike.[5] In short, a single district court[6] can enjoin the government's policies on a nationwide scope. This extraordinary form of relief is especially alarming because politically motivated litigants easily obtain nationwide injunctions to halt an administration's policy through forum and judge shopping, effectively circumventing the political process and turning the judiciary into a political weapon.

---

1 Sam Heavenrich, *An Appellate Solution to Nationwide Injunctions*, 96 IND. L.J. SUPP. 1, 4 (2020).

2 *Id.*

3 Nationwide injunctions are also called "universal" injunctions. The term used matters in the sense that each emphasizes a distinct, key element of what distinguishes a nationwide/universal injunction from a typical injunction. Calling it a nationwide injunction has the rhetorical advantage of communicating the "territorial breadth" of these injunctions. Trump v. Hawaii, 138 S. Ct. 2392, 2425 n.1 (Thomas, J., concurring). This is useful when the focus of your criticism of nationwide injunctions is its alarmingly sweeping scope given its extraterritorial application. Calling it a universal injunction has the academic advantage of communicating that the injunction's unique quality is that "they prohibit the Government from enforcing a policy with respect to anyone, including nonparties." *Id.* (Thomas, J., concurring). This is useful when the focus of the critique is the nationwide injunction's deviation from the traditional scope of courts' equity powers.

The term used does not substantively alter the nature of the injunction. I have opted to call them nationwide injunctions because the driving critique of this Comment is that the appeal of nationwide injunctions is its ability to halt a President's policy instantaneously and on a national level, an achievement which by other means would entail significantly higher costs for interested parties.

4 *Injunction*, BLACK'S LAW DICTIONARY (11th ed. 2019).

5 Samuel L. Bray, *Multiple Chancellors: Reforming the National Injunction*, 131 HARV. L. REV. 417, 425 (2017).

6 Not always a district court, but mostly. Getzel Berger, *Nationwide Injunctions Against the Federal Government: A Structural Approach*, 92 N.Y.U. L. REV. 1068, 1106 (2017) ("As the primary fact finders, district courts are usually the ones issuing injunctions, which are then reviewed by the courts of appeal. Appellate courts, however, can and do issue broad injunctions of their own.").

There are numerous problematic consequences of issuing nationwide injunctions to halt executive actions. First, it erodes trust in the nonpartisanship and judgment of the judiciary, as the issues at the heart of these suits are often highly politicized. Second, it allows litigants to act strategically: they not only forum shop for a jurisdiction inclined to enjoin the policy, but they shop for judges inclined to order such drastic remedy.[7] It is no coincidence that litigants turned to red states when seeking nationwide injunctions to the Obama Administration's policies and turned to blue states when seeking them against the Trump Administration's policies.[8]

This paper argues that to protect the judiciary's apolitical integrity and address the alarming rise of this extraordinary form of relief, district courts should apply the political question doctrine to determine whether or not to grant a nationwide injunction. Applying the political question doctrine in those circumstances would alter the incentives making nationwide injunctions a lucrative and less costly move to halt an administration's policies than going through the political process. The political question doctrine is "primarily a function of the separation of powers."[9] It precludes the judiciary from entertaining certain suits if the matters implicated are better suited for resolution by the political branches. The justiciability and prudential concerns that motivate courts to apply the political question doctrine exist in the prolific granting of politically charged nationwide injunctions.

Scholars have written extensively on both nationwide injunctions and the political question doctrine. However, very few have even mentioned the two concepts together. There is a wide range of literature both for and against nationwide injunctions.[10] And there is also an extensive debate over the political question doctrine and whether it is even relevant or necessary today.[11]

---

7 Alex Botoman, Note, *Divisional Judge-Shopping*, 49 COLUM. HUM. RTS. L. REV. 297, 298–308 (2018).

8 Bray, *supra* note 5, at 459–60.

9 Baker v. Carr, 369 U.S. 186, 210 (1962).

10 *See, e.g.,* Ezra Ishmael Young, *The Chancellors Are Alright: Nationwide Injunctions and an Abstention Doctrine to Salve What Ails Us*, 69 CLEV. ST. L. REV. 859 (2021); Heavenrich, *supra* note 1, at 4; Botoman, *supra* note 7, at 298–308; Howard M. Wasserman, *"Nationwide" Injunctions Are Really "Universal" Injunctions and They Are Never Appropriate*, 22 LEWIS & CLARK L. REV. 335 (2018); Amanda Frost, *In Defense of Nationwide Injunctions*, 93 N.Y.U. L. REV. 1065 (2018); Bray, *supra* note 5, at 459–60.

11 *See, e.g.,* Elizabeth Earle Beske, *Political Question Disconnects*, 67 AM. U.L. REV. F. 35 (2018); Gwynne Skinner, *Misunderstood, Misconstrued, and Now Clearly Dead: The "Political Question Doctrine" As A Justiciability Doctrine*, 29 J.L. & POL. 427, 459 (2014) ("The fact that the Court rejected the 'political question doctrine' in these cases raises significant questions about whether the doctrine continues to exist at all."); Zachary Baron Shemtob, *The Political Question Doctrines:* Zivotofsky v. Clinton *and Getting Beyond the Textual-Prudential Paradigm*, 104 GEO. L.J. 1001, 1026 (2016) ("[P]olitical questions themselves are increasingly ignored. . . . Thus, . . . perhaps a future Court will abolish political questions altogether."); Rachel E. Barkow, *More Supreme Than Court? The Fall of the Political Question Doctrine and the Rise of Judicial Supremacy*, 102 COLUM. L. REV. 237, 300 (2002) ("The Supreme Court's failure even to consider the political question doctrine reflects a broader trend in which the Court

Therefore, this Comment contributes to the discussion on nationwide injunctions and what can be done about them and demonstrates that there is still life and relevancy in the political question doctrine.

Part I argues that the increasing prevalence of nationwide injunctions is due to the changing attitudes of courts towards granting them and the resulting shift in litigants' incentives to seek them as an easier way of achieving policy victories against an administration's policies. Part II discusses how the potential for nationwide injunctions to be used as a political weapon raises concerns about the propriety of the judiciary's role in resolving certain controversies, which are rooted in the principle of separation of powers and the need for judicial impartiality. Part II will also explore how the political question doctrine can be applied in considering requests for nationwide injunctions. Part III examines the current state of the political question doctrine and how the application of the political question doctrine in considering requests for nationwide injunctions could address the potential harm to the judiciary's impartiality and alter the current approach of litigants seeking such injunctions compared to other proposed approaches and their potential advantages and disadvantages.

I. THE TROUBLESOME RISE OF NATIONWIDE INJUNCTIONS

In a world where courts steadfastly refuse to issue nationwide injunctions, litigants could not view nationwide injunctions as a viable means to achieve their policy goals in one fell swoop. It follows that litigants' behaviors and incentive structure changes when courts start showing their willingness to issue nationwide injunctions. This Part explains the emergence of nationwide injunctions through the judiciary's evolving view on the propriety of courts to afford relief to nonparties. There is no single factor that explains rise of nationwide injunctions,[12] but this Part argues that courts' willingness to issue nationwide injunctions changed litigants' incentive structure, as they discovered that they could achieve political wins against an administration's policies through the cheaper and more expedient method of asking for nationwide injunctions.

---

overestimates its own powers and prowess vis-à-vis the political branches."); Louis Henkin, *Is There A "Political Question" Doctrine?*, 85 YALE L.J. 597, 598 (1976) ("One needs no special doctrine to describe the ordinary respect of the courts for the political domain.").

[12] *See* Frost, *supra* note 10, at 1090 ("Although nationwide injunctions have been issued more frequently over the last fifty years, that may be the natural response to the expansion of federal law and the recent increase in major policy changes made through unilateral executive action."); Bray, *supra* note 5, at 547 (attributing the rise of nationwide injunctions to "shifts in how judges thought about legal challenges and invalid laws, accompanied by changes in agency practice and new reasons for judicial confidence."); Suzette M. Malveaux, *Response, Class Actions, Civil Rights, and the National Injunction*, 131 HARV. L. REV. F. 56, 60 (2017) ("As the availability of the class action device goes down, the need for the national injunction goes up.").

A. *The Dramatic Shift from Disavowal to Widespread Acceptance*

Plaintiffs naturally seek as wide relief as possible.[13] This is especially true with politically motivated plaintiffs seeking a nationwide injunction to achieve a broader policy goal. Take, for example, the Trump Administration's "Muslim travel ban."[14] Within days, two district court judges issued a nationwide injunction blocking the executive order.[15] Spencer E. Amdur and David Hausman, attorneys for the American Civil Liberties Union Immigration Rights' Project, defended nationwide injunctions on the basis that "the same equities that require protection for the plaintiff often support protection for those who are similarly situated."[16] Indeed, nationwide injunctions rest, in part, on the theory that nonparties similarly situated should be afforded relief.[17]

As early as 1897, the Supreme Court rejected the "similarly situated" proposition as "too conjectural to furnish a safe basis upon which a court of equity ought to grant an injunction."[18] In *Scott v. Donald*,[19] the Court noted that virtually every suit is of interest to nonparties, and that there is an interest in preventing a multiplicity of suits, but to "'allow them to be made parties to the suit would confound the established order of judicial proceedings, and lead to endless perplexity and confusion.'"[20] In spite of a strong judicial interest in preventing a multiplicity of suits,[21] the Court refused to issue a nationwide injunction.

---

13 Maureen Carroll, *Aggregation for Me, but Not for Thee: The Rise of Common Claims in Non-Class Litigation*, 36 CARDOZO L. REV. 2017, 2027 (2015) ("Subject to the caveat presented by the necessity doctrine, plaintiffs in a range of substantive areas have the ability to choose whether to bring a class action or a quasi-individual claim.").

14 On January 27, 2017, mere days after his inauguration, President Donald Trump issued an executive order barring entry into the U.S. by nationals from certain countries. Exec. Order No. 13,769, 82 Fed. Reg. 8977 (Jan. 27, 2017).

15 Washington v. Trump, No. C17-0141JLR, 2017 WL 462040 (W.D. Wash. Feb. 3, 2017); Darweesh v. Trump, No. 17 CIV. 480 (AMD), 2017 WL 388504 (E.D.N.Y. Jan. 28, 2017).

16 Spencer E. Amdur, David Hausman, *Nationwide Injunctions and Nationwide Harm*, 131 HARV. L. REV. F. 49, 51 (2017).

17 *See, e.g.,* Wasserman, *supra* note 10, at 1080; Szymon S. Barnas, *Can and Should Universal Injunctions Be Saved?*, 72 VAND. L. REV. 1675, 1695 (2019); Susan R. Klein, *Movements in the Discretionary Authority of Federal District Court Judges over the Last 50 Years*, 50 LOY. U. CHI. L.J. 933, 963 (2019).

18 Scott v. Donald, 165 U.S. 107, 115 (1897).

19 *Id.*

20 *Id.* at 116 (quoting Cutting v. Gilbert, 6 F. Cas. 1079 (C.C.S.D.N.Y. 1865)).

21 *See, e.g.*, Elizabeth Martin, *Getting A Second Bite at the Apple: The Res Judicata Exception for Seeking Foreclosure Deficiencies in Illinois*, 2016 U. ILL. L. REV. 2271, 2280 (2016); Robert G. Bone, *Mapping the Boundaries of the Dispute: Conceptions of Ideal Lawsuit Structure from the Field Code to the Federal Rules*, 89 COLUM. L. REV. 1, 29 (1989).

In *Frothingham v. Mellon*,[22] the Supreme Court displayed sensitivity to the judicial propriety of issuing nationwide injunctions in light of the political branches' roles. *Frothingham* is now commonly known as a "taxpayer standing" case,[23] but by peeling back the reasons why the Court rejected taxpayer standing, we see that the case is really a rejection of nationwide injunctions because the judiciary should not "invade the province" of the other branches.[24] *Frothingham* concerned, for our purposes,[25] an individual plaintiff challenging an Act of Congress on the basis that she is a taxpayer and the appropriations from the Act would increase her taxation burden and amount to a taking of property without due process of law.[26] Because a taxpayer's "interest in the moneys of the treasury . . . is shared with millions of others,"[27] to order an injunction would in effect be a nationwide injunction that applies to Frothingham and nonparties alike.[28] The Court also clarified what an injunction does: "the court enjoins, . . . not the execution of the statute, but the acts of the official."[29] To enjoin the execution of the statute, the effect of which is a national injunction that enjoins the federal government from applying the statute against nonparties, "would be, not to decide a judicial controversy, but to assume a position of authority over the governmental acts of another and coequal department, an authority which plainly we do not possess."[30]

The Court changed its tune in *Wirtz v. Baldor Electric Co.*[31] and reflected a departure from the concerns that led the *Scott* and *Frothingham* courts from issuing a nationwide injunction. The plaintiffs in *Wirtz* sought to enjoin the application of the Secretary of Labor's determination about the prevailing wage in in the electrical motors and generators industry.[32] The

---

22 Massachusetts v. Mellon (*Frothingham*), 262 U.S. 447 (1923).

23 *See, e.g.*, Kyle B. Gee, Daimlerchrysler Corp. v. Cuno-*Denying State Taxpayers Standing in Federal Court: Are Municipal Taxpayers Next*?, 38 U. TOL. L. REV. 1241 (2007).

24 *Frothingham*, 262 U.S. at 488.

25 The Supreme Court heard *Frothingham* as a consolidated case. The first was a suit by the Commonwealth of Massachusetts that came to the Supreme Court under its original jurisdiction. The second, on which I focus, is a suit brought by an individual plaintiff named Harriet Frothingham. Frothingham, 262 U.S. at 478. I focus on the discussion of Mrs. Frothingham's claim because that is where the injunction issue shines.

26 *Frothingham*, 262 U.S. at 486.

27 *Id.* at 487.

28 Due to the fungibility of tax revenues, there would be no way to single out Frothingham's taxes and issue an injunction prohibiting the federal government from using *her* money to fund the program created by the Act of Congress. Bray, *supra* note 5, at 431.

29 *Frothingham*, 262 U.S. at 488.

30 *Id.* at 489.

31 337 F.2d 518 (D.C. Cir. 1963).

32 *Id.* at 520.

D.C. Circuit held that the district court could issue[33] a nationwide injunction because plaintiffs with standing sue "to vindicate the public interest"[34] and it would be inequitable[35] to enjoin the Secretary's wage determination only as to the plaintiff. Thus, the court's interest in uniformity and equity alleviated it from the "artificial restriction[]"[36] of plaintiff-protective injunctions. This is a far cry from *Frothingham*, which rejected nationwide injunctions for the more abstract reasons of preserving the constitutional structure.

After *Wirtz*, the Supreme Court weighed in and gave an even stronger signal that plaintiffs could receive nationwide injunctions. In *Flast v. Cohen*,[37] plaintiffs challenged federal funds to religious schools in violation of the Establishment Clause and, thinking that a nationwide injunction is not a proper remedy, limited its prayer for injunctive relief to issue against the New York City Board of Education.[38] The Supreme Court held that plaintiffs did not have to so limit the injunction, reasoning that if the district court issued such narrow injunction, "that decision would cast sufficient doubt on similar programs elsewhere."[39] The Supreme Court found the finality and certainty of a nationwide injunction appealing.

Courts had never been blind to the public interest argument for issuing a nationwide injunction.[40] However, no court before *Wirtz* had accepted that argument as sufficient to cast aside "'the established order of judicial proceedings.'"[41] The consequence of this shift is a signal to plaintiffs that they can seek nationwide injunctions to achieve desired policy outcomes through the judiciary and circumvent the inertia of the political branches. Nationwide injunctions did not immediately take off with the full force and prominence immediately after *Flast*.[42] Litigants started asking for them—and courts began granting them—with noticeable force beginning in the George W. Bush administration.[43]

---

33 The court did not outright issue a nationwide injunction but rather held that, if upon remand, the district court found that the plaintiffs had standing, the district court "should enjoin the effectiveness of the Secretary's determination with respect to the entire industry." *Id.* at 535.

34 *Id.* at 534.

35 The court worried that to enjoin the Secretary's wage determination only as to the plaintiff would give the plaintiff an "unconscionable bargaining advantage over other firms in the industry" still subject to the Secretary's higher wage requirement. *Id.*

36 *Id.* at 535.

37 392 U.S. 83 (1968).

38 *Id.* at 89.

39 *Id.* at 90.

40 *Frothingham*, 262 U.S. at 487 ("The administration of any statute . . . is essentially a matter of public and not of individual concern."); Scott, 165 U.S. at 116 (1897) ("There is scarcely a suit at law or in equity which settles a principle . . . that does not involve a question in which other parties are interested.").

41 *Scott*, 165 U.S. at 116 (quoting Cutting v. Gilbert, 6 F. Cas. 1079 (C.C.S.D.N.Y. 1865)).

42 Bray, *supra* note 5, at 444.

43 *Id.*

B. *Adverse Incentives of Nationwide Injunctions*

Nationwide injunctions appear to now be entrenched in judicial practice. Some argue that nationwide injunctions are unconstitutional and outside the "judicial power."[44] These arguments ultimately imply that the entire concept of a nationwide injunction should cease entirely. This Comment does not go that far (although it is sympathetic to that recommendation). Believing that nationwide injunctions are here to stay, this Comment argues that the politicization of the judiciary that nationwide injunctions incentivizes should play a prominent role in how the benefits and harms of issuing nationwide injunctions are analyzed. This Section recaps some, but not all,[45] common harms attributed to nationwide injunctions—forum/judge shopping and asymmetrical issue preclusion—and contextualizes these harms in the politicization of the judiciary they produce.

1. Greater Certainty: Forum/Judge Shopping

The political pattern that has emerged with regard to nationwide injunctions are unmistakable. Litigants bring suit against Republican presidents in blue states like California, Washington, and New York, and against Democratic presidents in red states like Texas.[46] The problem goes beyond just shopping for a favorable court, but also singling out favorable judges. In *Texas v. United States*,[47] for example, the State of Texas and its cohort of twenty-five Republican states filed suit against the Obama administration's DACA and DAPA policies, discussed in more detail below, in the Brownsville division of the Southern District of Texas where there are only two

---

44 Trump v. Hawaii, 138 S. Ct. 2392, 2425 (2018) (Thomas, J., concurring) ("[Universal injunctions] appear to be inconsistent with longstanding limits on equitable relief and the power of Article III courts."); Wasserman, *supra* note 10, at 340 ("[U]niversal injunctions are inappropriate as a matter of equitable principle, judicial decisionmaking, and Article III of the Constitution.").

45 There are several other harms associated with nationwide injunctions. These include the risk of conflicting injunctions and doctrinal inconsistencies with class actions nonmutual offensive issue preclusion's inapplicability to the federal government. Bray, *supra* note 5, at 462. Nationwide injunctions also undermine the "value in percolation among the circuits" of novel issues to provide a rich body of law from which the Supreme Court can determine is worthy of attention. Matthew Erickson, *Who, What, and Where: A Case for A Multifactor Balancing Test As A Solution to Abuse of Nationwide Injunctions*, 113 NW. U. L. REV. 331, 341 (2018).

46 *See, e.g.*, Wasserman, *supra* note 10, at 363; Bray, *supra* note 5, at 459–60. *But see* Doug Rendleman, *Preserving the Nationwide National Government Injunction to Stop Illegal Executive Branch Activity*, 91 U. COLO. L. REV. 887, 937–39 (2020) (arguing that forum shopping is an inherent and unavoidable advantage of the plaintiff, and doctrines of personal jurisdiction and venue already exist to prevent forum shopping to the extent that it can be abused).

47 86 F. Supp. 3d 591 (S.D. Tex. 2015).

active federal district judges.[48] Judge Andrew Hanen, an openly conservative critic of President Obama's immigration policies took the suit and issued a nationwide injunction.[49]

The *Texas v. United States* case is a perfect example of the damage that nationwide injunctions have on the judiciary. Twenty-six politically motivated Republican states challenged a Democratic president's policy and judge shopped to obtain a nationwide injunction effectively halting that president's policy on a matter of high political significance—amid the backdrop of a presidential election no less. The judiciary looks politicized and partial when district courts grant nationwide injunctions on highly political matters stemming from obvious forum and judge shopping. "Inserting the judiciary into quintessentially political fights, even when there is a substantial legal issue to be decided on recognizably legal grounds, plainly risks the perception that judges base decisions on political preferences or at least are affected by those preferences."[50]

2. Asymmetrical Outcome Costs: Losing Is Not the End of the Matter

The rush to enjoin an administration's policy often comes from multiple politically motivated groups.[51] These plaintiffs enjoy a national forum from which to shop to bring their suit. Plaintiffs enjoy an additional advantage that the government does not: they can litigate and relitigate until they win. Different plaintiffs can bring multiple actions against the same policy in different district courts and circuits to see what it sticks. The government, meanwhile, has to defend itself on multiple fronts. The stakes for both parties are asymmetrical:[52] the plaintiff can lose with the opportunity to relitigate the issue elsewhere and ultimately enjoin the policy; the government can win and still face the risk that it will lose in a different court, or it can lose in one court and the nationwide effect of that strikes down the policy entirely.

---

48 Andrew Kent, *Nationwide Injunctions and the Lower Federal Courts*, LAWFARE (Feb. 3, 2017, 3:02 PM), https://www.lawfareblog.com/nationwide-injunctions-and-lower-federal-courts.

49 *Id.*

50 Ronald A. Cass, *Nationwide Injunctions' Governance Problems: Forum-Shopping, Politicizing Courts, and Eroding Constitutional Structure*, 27 GEORGE MASON L. REV. 30 (2020).

51 State attorneys general, for example, have "unique . . . incentives that enable them to frequently sue the federal government and successfully secure nationwide injunctions." Elysa M. Dishman, *Generals of the Resistance: Multistate Actions and Nationwide Injunctions*, 54 ARIZ. ST. L.J. 359, 362 (2022).

52 *See, e.g.,* Erickson, *supra* note 45, at 339; Berger, *supra* note 6, at 1090.

C. *Nationwide Injunctions Are Political Tools*

The Department of Justice estimates that district courts issued twelve nationwide injunctions during the Bush administration, nineteen during the Obama administration, and at least fifty-five during the Trump administration.[53] During the Obama administration, for example, district courts enjoined the U.S. military's "Don't Ask Don't Tell" policy,[54] the transgender school bathroom policy,[55] and the DACA and DAPA policies.[56] During the Trump administration, courts enjoined the Muslim travel ban[57] and the plan to defund sanctuary cities.[58] During the Biden administration, we have already seen nationwide injunctions in regard to the COVID mask mandate in public transportation[59] and the partial student debt forgiveness plan.[60]

This Section argues that nationwide injunctions encourage litigants to weaponize the judiciary as a work-around to challenge political policies that they do not like and cannot as easily defeat through the political process. A couple of case studies illustrate the point and bring some patterns to the surface. First, the judiciary has become the political battlefield over which fiercely political matters have played out. The difficulty and failures of achieving desired policies motivated political interests to ask courts to do the hard work, and the nationwide injunction has been the primary tool. Second, nationwide injunctions are bipartisan tools: Democrats and Republicans use them against the administration of the opposing party. One party supports or opposes nationwide injunctions according to who sits in the White House. Third, while interested Democratic and Republican coalitions use them, presidents hate them.

---

53 Deputy Attorney General Jeffrey A. Rosen Delivers Opening Remarks at Forum on Nationwide Injunctions and Federal Regulatory Programs, DEP'T OF JUST. (Feb. 12, 2020), https://www.justice.gov/opa/speech/deputy-attorney-general-jeffrey-rosen-delivers-opening-remarks-forum-nationwide.

54 Log Cabin Republicans v. United States, 716 F. Supp. 2d 884, 969 (C.D. Cal. 2010).

55 Texas v. United States, 201 F. Supp. 3d 810, 815 (N.D. Tex. 2016).

56 Texas v. United States, 86 F. Supp. 3d 591, 676 (S.D. Tex. 2015), *aff'd*, 809 F.3d 134 (5th Cir. 2015).

57 Washington v. Trump, No. C17-0141JLR, 2017 WL 462040 (W.D. Wash. Feb. 3, 2017); Darweesh v. Trump, No. 17 CIV. 480 (AMD), 2017 WL 388504 (E.D.N.Y. Jan. 28, 2017).

58 County of Santa Clara v. Trump, No. 17-cv-00574-WHO, 2017 WL 1459081, at *22 (N.D. Cal. Apr. 25, 2017).

59 Jacob Gershman & Alison Sider, *Judge Throws Out Federal Mask Mandate for Public Transportation*, WALL ST. J. (Apr. 18, 2022), https://www.wsj.com/articles/judge-throws-out-federal-mask-mandate-for-public-transportation-11650306480.

60 Danielle Douglas-Gabriel, *Appeals Court Temporarily Halts Biden's Student Debt Relief Program*, WASH. POST (Oct. 21, 2022), https://www.washingtonpost.com/education/2022/10/21/student-loan-forgiveness-temporary-stay/.

1. DACA and DAPA

Members of Congress have introduced some form of what is generally called the Development, Relief, and Education for Alien Minors Act (DREAM Act) since 2001.[61] However, no version of the DREAM Act has attained a majority of both houses of Congress.[62] President Obama then decided that "We Can't Wait"[63] and announced the Deferred Action for Childhood Arrivals (DACA) program, an executive policy to halt the deportation of illegal immigrants who met certain criteria.[64] He later expanded DACA and introduced the Deferred Action for Parents of Americans and Lawful Permanent Residents (DAPA) program.[65]

Just two weeks after announcing DAPA, twenty-six Republican states sought a nationwide injunction which a single district judge in Texas, a Republican appointee, granted.[66] The Fifth Circuit affirmed,[67] and, amid the backdrop of the 2016 presidential election, an equally divided Supreme Court affirmed.[68] President Obama called the decision "heartbreaking" and "frustrating,"[69] but having determined that his policies could simply not get through Congress, remained committed to having the political battle over his policies play out in the courts.[70] The decision also became a political rallying cry ahead of the election, and the Supreme Court was stuck right in the center of political cross-fire. Protests in front of the Supreme Court building sprung up immediately.[71] The tie vote at the Supreme Court added more salt to the

---

[61] Dylan Matthews, *Rep. Luis Gutiérrez Explains How Immigration Reform Gets Out of the House*, WASH. POST (Jul. 22, 2013), https://www.washingtonpost.com/news/wonk/wp/2013/07/22/rep-luis-gutierrez-explains-how-immigration-reform-gets-out-of-the-house/.

[62] *Texas*, 86 F. Supp. 3d at 608 ("To date, however, neither the President nor any member of Congress has proposed legislation capable of resolving these issues in a manner that could garner the necessary support to be passed into law.").

[63] Josh Blackman, *The Constitutionality of DAPA Part II: Faithfully Executing the Law*, 19 TEX. REV. L. & POL. 213, 268 (2015).

[64] *Texas*, 86 F. Supp. 3d at 608, *aff'd per curiam*, 136 S. Ct. 2271 (2016).

[65] *Id.* at 610.

[66] State of Texas; State of Alabama; State of Georgia; State of Idaho; State of Indiana; State of Kansas; State of Louisiana; State of Montana; State of Nebraska; State of South Carolina; State of South Dakota; State of Utah; State of West Virginia; State of Wisconsin; Governor Phil Bryant, State of Mississippi; Governor Paul R. Lepage, State of Maine; Governor Patrick L. McCrory, State of North Carolina; and Governor C.L. "Butch" Otter, State of Idaho, 2014 WL 6806231.

[67] Texas v. United States, 809 F.3d 134, 146 (5th Cir. 2015).

[68] United States v. Texas, 579 U.S. 547 (2016) (per curiam).

[69] *Remarks by the President on the Supreme Court Decision on U.S. Versus Texas*, THE WHITE HOUSE PRESIDENT BARACK OBAMA (June 23, 2016, 11:53 AM), https://obamawhitehouse.archives.gov/the-press-office/2016/06/23/remarks-president-supreme-court-decision-us-versus-texas.

[70] *Id.* ("[W]e're going to have to abide by that ruling until an election and a confirmation of a ninth justice of the Supreme Court so that they can break this tie.").

[71] *Dems: Immigration Decision Will 'Energize' Hispanic Voters*, THE HILL (June 23, 2016), https://thehill.com/latino/284659-dems-scotus-immigration-decision-will-energize-latino-voters/.

political wound that existed when the Republican-controlled Senate steadfastly refused to act on President Obama's nomination of Merrick Garland to the Supreme Court.[72]

Republicans' success was short-lived, however. President Trump fulfilled his campaign promise and rescinded DACA and DAPA.[73] The Trump Administration claimed that DACA and DAPA exceeded Executive power and that by rescinding the programs, the matter would be returned to Congress.[74] Predictably, however, Democratic groups challenged the rescission in court. And predictably, the battlefields were district courts in California,[75] New York,[76] and Washington, D.C.[77] The California and New York district courts issued nationwide injunctions,[78] and the Supreme Court ultimately ruled that the Trump administration's decision to end DACA violated the Administrative Procedure Act.[79] Again amidst a presidential election in 2020, President Trump called the decision "politically charged."[80]

2. President Trump's Travel Ban

The saga of President Trump's travel ban tells a similar story of nationwide injunctions as the primary tool in the tug-of-war between the administration and courts. The speed with which nationwide injunctions are sought and issued are remarkable and contribute to their appeal. Indeed, just three days after President Trump issued the first version of the travel ban,[81] the State of Washington filed suit.[82] Four days later, a district court judge issued a nationwide injunction.[83] The Ninth Circuit affirmed six days later.[84] Congressional action could certainly not move as fast; Congress would not have

---

72 Nina Totenberg, *170-Plus Days and Counting: GOP Unlikely to End Supreme Court Blockade Soon*, NPR (Sept. 6, 2016), https://www.npr.org/2016/09/06/492857860/173-days-and-counting-gop-unlikely-to-end-blockade-on-garland-nomination-soon.

73 Dep't of Homeland Sec. v. Regents of the Univ. of California, 140 S. Ct. 1891, 1903 (2020).

74 Regents of Univ. of California v. United States Dep't of Homeland Sec., 279 F. Supp. 3d 1011, 1026 (N.D. Cal. 2018).

75 *Id.*

76 Batalla Vidal v. Duke, 295 F. Supp. 3d 127 (E.D.N.Y. 2017).

77 Nat'l Ass'n for the Advancement of Colored People v. Trump, 298 F. Supp. 3d 209 (D.D.C. 2018).

78 Dep't of Homeland Sec., 140 S. Ct. at 1904.

79 *Id.* at 1916.

80 Nina Totenberg, *Supreme Court Rules For DREAMers, Against* Trump, NPR (June 18, 2020), https://www.npr.org/2020/06/18/829858289/supreme-court-upholds-daca-in-blow-to-trump-administration.

81 Exec. Order No. 13769, 82 Fed. Reg. 8977 (Jan. 27, 2017).

82 Complaint, State v. Trump, No. 2:17-cv-00141-JLR, 2017 WL 443297 (W.D. Wash. Jan. 30, 2017).

83 Washington v. Trump, No. C17-0141JLR, 2017 WL 462040 (W.D. Wash. Feb. 3, 2017).

84 Washington v. Trump, 847 F.3d 1151 (9th Cir. 2017).

moved at all, in fact, since Republicans controlled both houses. Under this set of political circumstances, Democrats' only viable recourse was the nationwide injunction.

Indeed, nationwide injunctions are effective recourses. The Trump administration was forced to defend the travel ban in various courts and at different stages. After the Ninth Circuit enjoined the travel ban, the Trump administration rescinded the executive order and issued a new one, clarifying the policy and hoping to make it less susceptible to legal challenge.[85] With the nationwide injunction at play, such a hope was a non-starter. Two district courts promptly issued another nationwide injunction, and the respective appeals courts affirmed.[86] The speed and ease with which litigants obtain nationwide injunctions demonstrate that nationwide injunctions give litigants the unbounded opportunity to challenge the statute in various courts simultaneously and "[s]hop 'til the statute drops."[87]

II. APPLICATION OF THE POLITICAL QUESTION DOCTRINE TO NATIONWIDE INJUNCTIONS

The concerns that motivated the Supreme Court to create the political question doctrine arose from textual and prudential concerns related to the propriety of the judiciary to settle certain controversies. This ultimately reflects a judicial interest in respecting the separation of powers and the judiciary's duty to be impartial. The political weaponization of the judiciary resulting from nationwide injunctions implicates these concerns. This Section will explain how courts can apply the political question doctrine when confronted with a request to issue a nationwide injunction.

The seminal case that established the political question doctrine is *Baker v. Carr*.[88] *Baker* concerned an equal protection challenge to the State of Tennessee's apportionment statute for seats in the state legislature.[89] The Court rejected arguments that the case was a nonjusticiable political question and clarified its precedent on the doctrine by clearly articulating six factors, the presence of any one of which is cause for dismissal as nonjusticiable. These six factors are:

> [(1)] a textually demonstrable constitutional commitment of the issue to a coordinate political department; or [(2)] a lack of judicially discoverable and manageable standards for resolving it; or [(3)] the impossibility of deciding without an initial policy determination of a kind clearly for non-judicial discretion; or [(4)] the impossibility of a court's

---

85 Exec. Order No. 13780, 82 Fed. Reg. 13209 (Mar. 6, 2017).

86 Int'l Refugee Assistance Project v. Trump, 857 F.3d 554 (4th Cir. 2017); Hawaii v. Trump, 859 F.3d 741 (9th Cir. 2017) (per curiam).

87 Bray, *supra* note 5, at 460.

88 369 U.S. 186 (1962).

89 *Id.* at 198–200.

> undertaking independent resolution without expressing lack of the respect due coordinate branches of government; or [(5)] an unusual need for unquestioning adherence to a political decision already made; or [(6)] the potentiality of embarrassment from multifarious pronouncements by various departments on one question.[90]

The Court qualified that "the mere fact that the suit seeks protection of a political right does not mean it presents a political question."[91] Rather, "a political question is primarily a function of the separation of powers"[92] requiring "discriminating inquiry into the precise facts and posture of the particular case."[93]

The *Baker* test merges textual and prudential concerns into one unifying doctrine, giving each equal weight.[94] Yet the Court since *Baker* has inconsistently applied the political question doctrine, as Zachary Baron Shemtob discovered after surveying all thirty-eight cases involving political questions since *Baker* and cataloging the different interpretations and manipulations the *Baker* factors have undergone.[95] The Court has particularly wavered in its commitment to the prudential *Baker* factors.[96] The following Section will explain why the prudential factors fell into disfavor and why that devolution is unfortunate due to the increasing need for prudentialism to maintain the judiciary's apolitical integrity—especially in light of the unique threats to the judiciary's integrity that nationwide injunctions pose.

A. *Revival of Prudentialism*

The prudential factors in *Baker* reflect an acknowledgment that there are matters over which the judiciary is institutionally incompetent to entertain without risking undue backlash from the other two branches or risking its integrity as a non-self-enforcing branch. The Guarantee Clause of the Constitution provided the foundation for this prudential concern. Thus, in *Luther v. Borden*,[97] where the Court was asked to interpret the Guarantee Clause, Chief Justice Taney stressed the practical consequences of the Court wading into those waters.[98] The Court reconfirmed these concerns in *Pacific States Telephone & Telegraph Co. v. Oregon*,[99] where the Court was asked to define

---

[90] *Id.* at 217.

[91] *Id.* at 209.

[92] *Id.* at 210.

[93] *Id.* at 217.

[94] Shemtob, *supra* note 11, at 1008.

[95] *Id.* at 1002.

[96] Barkow, *supra* note 11, at 267–68; Vieth v. Jubelirer, 541 U.S. 267, 278 (2004) ("The [*Baker*] test [is] probably listed in descending order of both importance and certainty.").

[97] 48 U.S. 1 (1849).

[98] *Id.* at 41 (raising concerns over what further complications await a court that injected itself in the matter).

[99] 223 U.S. 118 (1912).

when a state lacked a republican form of government. Allowing the judiciary to settle matter, the Court reasoned, would empower courts to tear down and fashion new state governments.[100]

Alexander Bickel famously articulated facets of prudential concerns distinct from *Baker*'s prudential factors by illuminating of the processes that courts consider:

> (a) the strangeness of the issue and its intractability to principled resolution; (b) the sheer momentousness of it, which tends to unbalance judicial judgment; (c) the anxiety, not so much that the judicial judgment will be ignored, as that perhaps it should but will not be; (d) finally ('in a mature democracy'), the inner vulnerability, the self-doubt of an institution which is electorally irresponsible and has no earth to draw strength from.[101]

Bickel's vision of political question and judicial integrity has been attributed to a post-New Deal concern over judicial overreaching after the court packing crisis and radical changes in the power of government.[102] As a "tool of avoidance" and judicial restraint, the political question doctrine is the judiciary's way to restore its legitimacy in the public mind.[103]

Yet the prudential concerns have fallen into disfavor for a number of reasons. First are those who criticize prudential factors as ungrounded in the Constitution or any textual support.[104] The prudential factors of the political question doctrine, they argue, are less grounded in precedent as are the textual factors.[105] Prudentialism has also died because of the full-scape societal and judicial acceptance of judicial supremacy and the propriety and expectation for courts to wade into the pressing social issues of the day.[106] Thus, the Supreme Court is more reluctant to exercise judicial restraint in matters that have significant social and political impact. In light of the decreasing

---

100 *Id.* at 142 ("And as a consequence of the existence of such judicial authority, a power in the judiciary must be implied . . . to build by judicial action upon the ruins of the previously established government a new one[.]").

101 Alexander M. Bickel, THE LEAST DANGEROUS BRANCH: THE SUPREME COURT AT THE BAR OF POLITICS, 184 (1962).

102 Mark Tushnet, *Law and Prudence in the Law of Justiciability: The Transformation and Disappearance of the Political Question Doctrine*, 80 N.C. L. REV. 1203, 1231 (2002) ("They were dealing with a Supreme Court that had only recently emerged from the crisis precipitated by the Court's obstruction of the New Deal and that was simultaneously attempting to redefine the scope of government power by developing civil rights and civil liberties restrictions on government power."); Henkin, *supra* note 11, at 625 ("The political question doctrine . . . was perhaps an expression of a wider, deeper mood by Justices appointed to restore judicial self-restraint and allow the elected governors to govern.").

103 Barkow, *supra* note 11, at 258.

104 Szurkowski, *The Return of Classical Political Question Doctrine in* Zivotofsky Ex Rel. Zivotofsky v. Clinton, 132 S. Ct. 1421 (2012), 37 HARV. J.L. & PUB. POL'Y 347, 353 (2014).

105 Shemtob, *supra* note 11, at 1008.

106 *Id.* at 1025 (2016) ("[M]any Americans have come to accept a strong judiciary ready to wade into the most momentous issues of the day and assert judicial supremacy over these."); Tushnet, *supra* note 102, at 1233; Barkow, *supra* note 11, at 300.

confidence the public places in the judiciary,[107] the concerns motivating prudential considerations should be revisited. The prudential factors are not completely dead: despite the Supreme Court's infrequent use of the prudential elements of the political question doctrine, lower courts have been more willing to adopt them.[108]

The remarkable rise of district courts issuing nationwide injunctions threatens to implicate the very concerns that militate in favor of the prudential judicial self-restraint of the political question doctrine. Litigants seeking nationwide injunctions on politically salient matters to stop the administration from acting do so because it is an easier alternative than working through the political process to achieve political goals. A district court that does these litigants' bidding and wades into politically controversial matters in such early stages risks the perception that they are just another partisan actor. Thus, news articles find it relevant to mention what president and what party appointed any judge mentioned.[109] Before her appointment to the Supreme Court, then-Judge Amy Coney Barrett stated that the greatest threat to the judiciary today was "people perceiving [it] as partisan."[110] Prudential factors could become more crucial as these warning signs amplify.

B. *How the Doctrine's Application Would Play Out*

How the political question doctrine would factor into a district court's practical decision of whether to issue a nationwide injunction is the next key question. To illustrate how to square the political question doctrine, this Section takes a few case studies of district courts issuing nationwide injunctions.

1. The DACA and DAPA Nationwide Injunction

As discussed in Part I of this Comment, just two weeks after President Obama announced the DAPA policy and DACA extension, twenty-six Republican states sought a nationwide injunction which a single district judge

---

107 Jeffrey M. Jones, *Confidence in U.S. Supreme Court Sinks to Histori Lows*, GALLUP (Jun. 23, 2022), https://news.gallup.com/poll/394103/confidence-supreme-court-sinks-historic-low.aspx.

108 Skinner, *supra* note 11, at 465; Szurkowski, *supra* note 104, at 356; Linda Champlin Alan, *Political Question Doctrine and Allocation of the Foreign Affairs Power*, 13 HOFSTRA L. REV. 215, 217 (1985).

109 Jess Bravin, *No Obama or Trump Judges Here, Appointees of Both Declare*, WALL ST. J. (Sept. 15, 2019), https://www.wsj.com/articles/judges-say-they-arent-extensions-of-presidents-who-appointed-them-11568566598 (recounting various federal judges criticizing news articles that emphasize which party appointed a judge as if they were party actors).

110 *Id.*

in Texas, a Republican appointee, granted and[111] the Fifth Circuit affirmed.[112] In *Texas v. United States*, the district court found as a preliminary matter that the plaintiffs had constitutional standing, prudential standing (generalized grievances and zone of interest), and standing under the Administrative Procedure Act.[113] But by granting a nationwide injunction, the district court did not limit the injunction as applied to the particular plaintiffs; by its nature, nationwide injunctions afford relief to people who are similarly situated but may or may not have standing to seek an injunction. This incongruency in embarking upon an extensive standing analysis to then afford relief to individuals who might not have standing demonstrates the paradox of nationwide injunctions. In the ruling granting the nationwide injunction, the district court did not discuss the propriety of the nationwide scope of the injunctive relief. Instead, the court relied on Fifth Circuit precedent to support the granting of a preliminary injunction.[114]

The political question doctrine being a doctrine of justiciability, its application would come before the court gets to the merits of the case. That is, as another barrier in the standing analysis in addition to the constitutional, prudential, and APA standing questions. *Texas v. United States* would be an example of a political question militating for limiting injunctive relief to the particular plaintiffs. The twenty-six Republican states who filed the challenge against DACA and DAPA were clearly politically motivated to halt President Obama's immigration policies because they disagreed with them. Plus, they strategically filed suit in a favorable district court in Texas hoping to obtain such drastic nationwide relief. This implicated many of the factors in the political question doctrine.

First, there is "the impossibility of deciding without an initial policy determination of a kind clearly for non-judicial discretion."[115] This policy determination dovetails with the standing analysis: while the district court ruled that the particular plaintiff States met injury in fact, causation, and redressability, the court could not make that same determination as to non-plaintiffs

---

111 State of Texas; State of Alabama; State of Georgia; State of Idaho; State of Indiana; State of Kansas; State of Louisiana; State of Montana; State of Nebraska; State of South Carolina; State of South Dakota; State of Utah; State of West Virginia; State of Wisconsin; Governor Phil Bryant, State of Mississippi; Governor Paul R. Lepage, State of Maine; Governor Patrick L. McCrory, State of North Carolina; and Governor C.L. "Butch" Otter, State of Idaho,, 2014 WL 6806231.

112 Texas v. United States, 809 F.3d 134, 146 (5th Cir. 2015).

113 Texas v. United States, 86 F. Supp. 3d 591, 614 (S.D. Tex. 2015), *aff'd per curiam*, 136 S. Ct. 2271 (2016).

114 *Id.* at 646. The elements a plaintiff needs to establish are: "(1) a substantial likelihood of success on the merits; (2) a substantial threat that the [States] will suffer irreparable injury if the injunction is denied; (3) that the threatened injury outweighs any damage that the injunction might cause [Defendants]; and (4) that the injunction will not disserve the public interest." *Id*. These factors are generally the same across jurisdictions. *See* Arthur D. Wolf, *Preliminary Injunction Standards in Massachusetts State and Federal Courts*, 35 W. NEW ENG. L. REV. 1 (2013).

115 *Baker*, 369 U.S. at 217.

to whom the nationwide injunction applies. The plaintiff States' injury stemmed primarily from an allegation that the DACA and DAPA policies would "create a new class of individuals eligible to apply for driver's licenses, the processing of which will impose substantial costs on [the States'] budget."[116] No determination was made as to States not parties to the suit, and yet the nationwide injunction sweeps them in as well.

Second, this case prevents the court from "undertaking [an] independent resolution without expressing lack of the respect due coordinate branches of government."[117] The DACA-DAPA saga which the court detailed demonstrated how politically charged the matter was. President Obama took unilateral executive action after failed attempts at legislative reform in Congress.[118] Republicans, of course, opposed the policies and rather than passing legislation to counter it, the strategy was to seek a nationwide injunction—a faster and politically cheaper route. A court wading into the issue on those circumstances ends up making determinations that the political branches could not themselves work out. Thus, it does not give due respect to the political branches, but rather inserts itself into one side or the other.

2. The Travel Ban Nationwide Injunction

This case implicated the sixth *Baker* factor: "the potentiality of embarrassment from multifarious pronouncements by various departments on one question."[119] The ability to challenge a policy in various courts also produces the potential that two district courts will issue conflicting nationwide injunctions.[120] The district court in Hawaii that enjoined the travel ban, for example, enjoined sections two and six[121] whereas the district court in Maryland just two weeks earlier enjoined only section two.[122] The Hawaii court took cognizance of this fact when it initially refused to rule on plaintiffs' motion for a nationwide injunction because a district court in Washington had issued a nationwide injunction that very day.[123] The hope is that judicial comity and judicial restraint would make judges take cognizance of other lawsuits and decide on whether to tailor their injunction accordingly. But that may not

116 *Texas*, 86 F. Supp. 3d at 616.

117 *Baker*, 369 U.S. at 217.

118 Josh Blackman, *The Constitutionality of DAPA Part II: Faithfully Executing the Law*, 19 TEX. REV. L. & POL. 213, 268 (2015).

119 *Baker*, 369 U.S. at 217.

120 Wasserman, *supra* note 10, at 383; Bray, *supra* note 5, at 462.

121 Hawai'i v. Trump, 245 F. Supp. 3d 1227, 1239 (D. Haw. 2017).

122 Int'l Refugee Assistance Project v. Trump, 241 F. Supp. 3d 539, 566 (D. Md. 2017).

123 *Hawai'i*, 241 F. Supp. 3d at 1123.

always be the case,[124] especially as litigants shop for courts and specific judges they believe are predisposed to ruling their way.

III. EVALUATING THE APPLICATION OF THE POLITICAL QUESTION DOCTRINE

The Supreme Court has never weighed in on the limits and contours of nationwide injunctions,[125] leaving it to each circuit to define the propriety of issuing them.[126] This status quo is untenable because it does not adequately mitigate the harms of nationwide injunctions to the judiciary's apolitical integrity. Instead, the current regime perpetuates the "[s]hop 'til the statute drops"[127] mentality of litigants quickly filing to obtain nationwide injunctions in various jurisdictions until they are successful. This Section explains how the political question doctrine would apply and alter a district court's analysis when faced with a nationwide injunction, how this approach compares to others that scholars have proposed, and the pros and cons of applying the political question doctrine instead.

A. *Applying the Doctrine*

If a court finds that a lawsuit requesting a nationwide injunction presents a nonjusticiable political question, what should the court do? The political question doctrine is a doctrine of justiciability: if a case presents a political question, a court is barred from hearing it altogether. The harshness of this doctrine is admitted, and indeed, it contributes to why the Supreme Court has so rarely invoked it.[128] But in light of the threats posed to the judiciary's apolitical integrity from the prolific granting of nationwide injunctions for fresh

---

124 Bray, *supra* note 5, at 464 ("Conflicting injunctions can be avoided with judicial restraint and good luck, but neither one is sure to last forever.").

125 Trump v. Hawaii, 138 S. Ct. 2391, 2429 (2018) (Kennedy, J., concurring) ("[Nationwide] injunctions are legally and historically dubious. If federal courts continue to issue them, this Court is dutybound to adjudicate their authority to do so.").

126 *See, e.g.*, Regents of Univ. of California v. United States Dep't of Homeland Sec., 279 F. Supp. 3d 1011, 1049 (N.D. Cal. 2018) ("[O]ur court of appeals considered this very issue . . . and upheld a nationwide injunction imposed by a single district court . . . . Indeed, the Fifth Circuit reached the same conclusion in determining the appropriate scope of an injunction over DAPA."); *Int'l Refugee Assistance Project*, 241 F. Supp. 3d at 565 (citing circuit precedent and the nature of the issue to determine whether a nationwide injunction is proper). District courts have broad discretion to grant nationwide injunctions, and appellate courts review them for abuse of discretion. Bray, *supra* note 5, at 466.

127 Bray, *supra* note 5, at 460.

128 Skinner, *supra* note 11, at 453.

and controversial political issues, applying the political question doctrine would alter the litigants' strategy in asking for them in the first place.[129]

Those who have criticized nationwide injunctions have offered various alternatives address the consequences of limiting nationwide injunctions.[130] The clear-cut solution to nationwide injunctions is that injunctions should be "plaintiff protective," as Samuel Bray put it, and should not afford relief any "broader than what the plaintiffs . . . should logically be able to bring contempt proceedings to enforce."[131] He gave the example of *Texas v. United States* and reasoned that a plaintiff-protective injunction would prohibit the Obama administration from requiring the plaintiff States to grant driver licenses on the basis of the DACA and DAPA program—the source of the plaintiff States' injury in fact.[132] A plaintiff-protective limitation is incompatible with the political question doctrine but it is tolerable. Although it would make no difference to the prudential propriety of a court to weigh in on a political question, a plaintiff-protective limitation would at least remove the overbroad effect of a nationwide injunction, which is part of what makes them so politically controversial and lucrative for politically motivated litigants. An injunction that affords relief only to the particular plaintiffs before the court would not halt an administration's policy wholesale, and thus could not be used as a political alternative to congressional gridlock.[133]

At the other end of the extreme is the idea to get rid of nationwide injunctions altogether and make plaintiffs resort to "[c]lass actions, associational standing, and third-party standing."[134] Like the plaintiff-protective approach, only named plaintiffs would get the benefit of the relief sought. Applying the political question doctrine to the matter would not entirely get rid of nationwide injunctions; there are conceivable circumstances in which a matter would not meet the threshold of being a political question. Therefore, this approach goes farther than the one I am proposing. However, this

---

129 Dishman, *supra* note 51, at 414 ("State litigants and courts will have greater incentive to seek more limited injunctions if they cannot get immediate relief with a nationwide injunction.").

130 Ezra Ishmael Young, *The Chancellors Are Alright: Nationwide Injunctions and an Abstention Doctrine to Salve What Ails Us*, 69 CLEV. ST. L. REV. 859, 905 (2021) (proposing a prudential, discretionary kind of abstention); Ryan Kirk, *A National Court for National Relief: Centralizing Requests for Nationwide Injunctions in the D.C. Circuit*, 88 TENN. L. REV. 515 (2021) (proposing making the D.C. Circuit the national court for all nationwide injunctions); Heavenrich, *supra* note 1, at 3 (proposing a change to the Federal Rules of Civil Procedure); Wasserman, *supra* note 10, at 386 (proposing getting rid of nationwide injunctions altogether and making plaintiffs resort to "[c]lass actions, associational standing, and third-party standing.").

131 Bray, *supra* note 5, at 469.

132 *Id.* at 470.

133 Charlton C. Copeland, *Seeing Beyond Courts: The Political Context of the Nationwide Injunction*, 91 U. COLO. L. REV. 789, 816–19 (2020) ("Congress's failure to legislate leaves a policy vacuum that puts the President at risk of not achieving important policy agenda items for his electoral coalition--thereby increasing the likelihood that the executive will undertake unilateral action . . . [and] trigger[ing] . . . nationwide injunction[s].").

134 Wasserman, *supra* note 10, at 386.

approach is the principled one to take if one believes that nationwide injunctions fall wholly outside the judiciary's Article III or equity powers.[135] This Comment has not emphasized these arguments because the focus of the Comment is to emphasize and address the politicization of the judiciary resulting from nationwide injunctions. Of course, a solution that entirely gets rid of nationwide injunctions would eliminate the problem as well.

A more modest approach is to geographically limit an injunction's scope.[136] The benefit of this approach compared to applying the political question doctrine or any other proposing a balancing test of sorts[137] is that a geographical limitation "is not subject to any broad form of interpretation."[138] It also benefits from the fact that it is a solution that Congress can implement via legislation. Applying the political question doctrine, on the other hand, would require circuit courts and eventually the Supreme Court to endorse and establish. Applying the political question doctrine to nationwide injunctions is a very novel idea, so there is the concern that district courts will be hesitant to consider it and run the risk of being overruled. Currently, district courts are governed by their circuit's precedent in deciding how and when nationwide injunctions are appropriate. A similar precedent setting would be needed to consider the political question doctrine to nationwide injunctions.

B. *Advantages of Applying the Doctrine*

When confronted with a prayer for nationwide injunctive relief, a district court applying the political question doctrine changes the incentive structures behind politically motivated litigants seeking nationwide injunctions in the first place. Namely, applying the political question doctrine would limit the success of forum and judge shopping, as litigants would not be able to rely on favorable judges if they are bound to apply the political question doctrine and finding that issuing a nationwide injunction is a nonstarter. This change in the incentive structure for litigants would depoliticize the courts because courts would be less likely to issue nationwide injunctions under a political question paradigm and avoid wading into highly sensitive

---

135 Trump v. Hawaii, 138 S. Ct. 2392, 2425 (2018) (Thomas, J., concurring) ("[Universal injunctions] appear to be inconsistent with longstanding limits on equitable relief and the power of Article III courts."); Wasserman, *supra* note 10, at 340 ("[U]niversal injunctions are inappropriate as a matter of equitable principle, judicial decisionmaking, and Article III of the Constitution.").

136 Joseph D. Kmak, *Abusing the Judicial Power: A Geographic Approach to Address Nationwide Injunctions and State Standing*, 70 EMORY L.J. 1325, 1363 (2021); Michael T. Morley, *Nationwide Injunctions, Rule 23(b)(2), and the Remedial Powers of the Lower Courts*, 97 B.U. L. REV. 615, 656 (2017); Getzel Berger, *Nationwide Injunctions Against the Federal Government: A Structural Approach*, 92 N.Y.U. L. REV. 1068, 1100 (2017).

137 *See, e.g.*, Erickson, *supra* note 45 (proposing a multifactor balancing test for courts considering nationwide injunctions).

138 Kmak, *supra* note 136, at 1365.

and controversial political issues in the first place. Lastly, litigants' strategy of challenging executive actions to put them on immediate halt would be frustrated. This would offer more legal and practical certainty to the effectiveness of executive actions.

Applying the political question doctrine to the matter of nationwide injunctions runs up against criticism of the doctrine itself and arguments that it is confusing,[139] irrelevant,[140] and even dead.[141] The Supreme Court has inconsistently applied and sent mixed signals on the political question doctrine, leaving lower courts to sort and create caselaw around it. The Court's refusal to even consider the political question doctrine in a case like *Bush v. Gore*,[142] for example, indicated a decline in deference overall as the Court has become more assertive with its abilities to address complex and political issues.[143] The prudential factors themselves have also come under attack, particularly after *Zivotofsky v. Clinton*[144] where the majority largely ignored the prudential factors.[145] Concurring in *Zivotofsky,* Justice Sotomayor cautioned against adjudicating a dispute on prudential bases alone. She warned lower courts not to "refuse to adjudicate a dispute merely because a decision 'may have significant political overtones' . . . . [n]or . . . because the question is difficult, the consequences weighty, or the potential real for conflict with the policy preferences of the political branches."[146] Still, the alarming rise of nationwide injunction and its attendant problems of politicizing the judiciary is something that the Court needs to confront, and the political question doctrine currently exists to address those types of concerns. The concerns that animated

---

139 *See, e.g.*, Elizabeth Earle Beske, *Political Question Disconnects*, 67 AM. U.L. REV. F. 35, 36 (2018).

140 *See, e.g.*, Shemtob, *supra* note 11, at 1026 ("[P]olitical questions themselves are increasingly ignored. . . . Thus, . . . perhaps a future Court will abolish political questions altogether."); Barkow, *supra* note 11, at 300 ("The Supreme Court's failure even to consider the political question doctrine reflects a broader trend in which the Court overestimates its own powers and prowess vis-à-vis the political branches."); Henkin, *supra* note 11, at 598 (1976) ("One needs no special doctrine to describe the ordinary respect of the courts for the political domain.").

141 *See, e.g.*, Skinner, *supra* note 11, at 459 ("The fact that the Court rejected the 'political question doctrine' in these cases raises significant questions about whether the doctrine continues to exist at all.").

142 531 U.S. 98 (2000).

143 Barkow, *supra* note 11, at 275 ("The Court did not pause for even a sentence in *Bush* I to explain why the Article II question was within its jurisdiction and why it did not present a political question for resolution by Congress.").

144 566 U.S. 189 (2012).

145 Samantha Goldstein, *The Real Meaning of* Zivotofsky *and Its Impact on Targeted Killings Cases*, 2 NAT'L SEC. L.J. 147, 162–63 (2014) ("The Court only mentioned the two constitution-based Baker factors in analyzing the justiciability question in Zivotofsky. It ignored *Baker's* other considerations . . . ."); Szurkowski, *supra* note 104, at 348 ("The Court seized the opportunity presented by Zivotofsky to reassert the classical, pre-Baker interpretation of the political question doctrine, implicitly but conspicuously disavowing the prudential theory even in foreign affairs cases.").

146 Zivotofsky, 566 U.S. at 204–05 (Sotomayor, J., concurring).

the creation of a unified political question doctrine in the first place exist with even greater force today in our polarized political landscape.

1. Reducing Forum/Judge Shopping

Forum and judge shopping are in a way inevitable and inherent to the fact that plaintiffs have the advantage of choosing where they file suit.[147] Indeed, Samuel Bray considered the "multiple-chancellor system", where different judges could rule differently, a "necessary precondition" to forum shopping.[148] But when plaintiffs are asking courts for an extraordinary form of relief like a nationwide injunction on a matter of high political saliency, the risk of politicizing the courts compounds the dangers in forum and judge shopping. The trend of litigants turning to red or blue states depending on the administration they are challenging indicates the partisan nature of nationwide injunction requests.[149] This is also the way the media report it: "Judge blocks Trump" or "Judge blocks Obama" are common headlines nowadays.[150]

The political question doctrine would act as a formidable bar to the likelihood of achieving a nationwide injunction, thus disincentivizing litigants from filing lawsuits in favorable jurisdictions and with favorable judges in the first place. Applying the political question doctrine would have the most impact and relevancy in cases like *Texas v. United States,* where a partisan coalition of States joined together. State attorneys general have "unique political incentives to sue the federal government" generally, but especially to "seek nationwide injunctions."[151] There, the political question is clearer, and the court has more prudential reasons to deny a request for a nationwide injunction.

Judges too would be bound by a coherent, existing doctrine in the decision to issue a nationwide injunction and in what is currently an unbounded

---

147 Markus Petsche, Wh*at's Wrong with Forum Shopping? An Attempt to Identify and Assess the Real Issues of A Controversial Practice*, 45 INT'L LAW. 1005, 1028 (2011) ("[F]orum shopping is inevitable as long as litigants are offered jurisdictional options or alternatives.").

148 Bray, *supra* note 5, at 448.

149 Bray, *supra* note 5, at 459–60.

150 William P. Barr, *End Nationwide Injunctions*, WALL ST. J. (Sept. 5, 2019), https://www.wsj.com/articles/end-nationwide-injunctions-11567723072.

151 Dishman, supra note 51, at 396. National injunctions allow state attorneys general to affect policy outside their state's borders, gaining notoriety and attention to their causes. *Id.* "[S]tate attorneys general [act] as partisan warriors against presidential administrations" and especially "Blue-state attorneys general" suits against Trump administration. Lisa Friedman & John Schwartz, *Borrowing G.O.P. Playbook, Democratic States Sue the Government and Rack Up Wins*, N.Y. TIMES (Mar. 21, 2018), https://www.nytimes.com/2018/03/21/climate/attorneys-general-trump-environment-lawsuits.html. For more details into this matter of state attorneys general and what can be done to limit their access and attraction to nationwide injunctions, *see* Jonathan Remy Nash*, State Standing for Nationwide Injunctions Against the Federal Government*, 94 NOTRE DAME L. REV. 1985 (2019).

and inconsistent field. The inconsistency in how nationwide injunctions are granted and the possibility of conflicting injunctions calls for a doctrine that could provide clarity.[152] Forum and judge shopping create the appearance of unequal treatment of the laws and precedent. This issue has important ramifications for the rule of law,[153] and upholding the public's confidence in the rule of law would itself mitigate the dangers to judicial confidence that nationwide injunctions pose.

2. Depoliticizing the Courts

Applying the political question doctrine to nationwide injunctions would act as a check for the court to avoid appearing politicized. The politically salient issues that often generate requests for nationwide injunctions are already highly publicized and watched;[154] when a court enjoins that policy, it risks contributing to the political debate and drawing the ire of the political supporters of that policy. The threat to the judiciary's apolitical integrity is especially threatened when nationwide injunctions cause the judiciary to face the criticism of *both* sides of the political aisle depending on the current political circumstances. Indeed, both Democrats and Republicans despise the nationwide injunction only when used against their own presidents, creating equal opportunity for both parties and one-half of the country to always be suspicious of the judiciary's political role. Nationwide injunctions are inherently highly political weapons; limiting courts from issuing them helps alleviate some of the political backlash the judiciary gets.

The benefit of applying the political question doctrine to a district court's decision of whether or not to issue a nationwide injunction is that it can help to preserve the separation of powers between the three branches of government. By deferring to the political process to resolve certain issues, the courts can avoid overstepping their jurisdiction and infringing on the authority of the other branches. It is important to note that the political question doctrine is not an absolute bar to judicial review, and it is not always clear

---

[152] One scholar has proposed centralizing requests for nationwide injunctions in the D.C. Circuit. Kirk, *supra* note 130. This is an intriguing prospect, as the D.C. Circuit already exists as a specialized court in many respects. The proposal would certainly address the forum and judge shopping issue but centralizing nationwide injunction requests in one court would not entirely eliminate the powerful incentive to seek nationwide injunctions in the first place or the political blowback to the judiciary for engaging in political questions.

[153] Cass, *supra* note 50, at 50 ("[T]he evidence of diverging expected outcomes for specific judges or courts in itself suggests a gap between current reality and important rule of law ideals.").

[154] Charlton C. Copeland, *Seeing Beyond Courts: The Political Context of the Nationwide Injunction*, 91 U. COLO. L. REV. 789, 796 (2020) ("[A] key component of the recent increase in nationwide injunction deployment likely was increased partisan polarization in Congress that led to increasingly gridlocked legislative processes, which in turn led to increased presidential unilateral action.").

when it applies. The doctrine is typically applied on a case-by-case basis,[155] and courts will consider a number of factors in determining whether an issue is properly resolved through the political process or whether it is appropriate for the courts to address.

3. Certainty of Executive Policies

If district courts are limited in their ability to issue nationwide injunctions, it could make it more difficult for individual plaintiffs to block or overturn federal policies or actions. This could provide greater certainty for the administration by reducing the risk that its policies will be disrupted or overturned by judicial action. Applying the political question doctrine to a district court's decision of whether or not to issue a nationwide injunction could result in greater certainty of an administration's policies in several ways.

Using the political question doctrine as a guide would provide a more consistent and transparent approach to nationwide injunction cases. The political question doctrine establishes clear criteria for determining when an issue is beyond the purview of the courts and should be left to the political branches of government to resolve. By using these criteria to evaluate nationwide injunction cases, district courts can make more consistent and predictable decisions about whether to issue a nationwide injunction. The political question doctrine recognizes that there are certain issues that are better left to the political branches of government to resolve, and that the resolution of these issues by the courts could lead to conflicting or inconsistent decisions. By leaving these issues to the other branches of government, the courts can help to avoid creating confusion and uncertainty about the law.

Using the political question doctrine would promote judicial restraint and avoid overreach by the judiciary. By recognizing that there are certain issues that are beyond the proper purview of the courts, the political question doctrine allows the courts to defer to the other branches of government in matters that are more appropriately within their jurisdiction. This can help to avoid instances where the courts are perceived as overstepping their authority and interfering with the policymaking process. By using the political question doctrine to determine whether or not to issue a nationwide injunction, district courts can provide a more consistent, transparent, and predictable approach to these cases, resulting in greater certainty of an administration's policies.

---

[155] *Baker*, 369 U.S. at 217 (1962) ("The cases we have reviewed show the necessity for discriminating inquiry into the precise facts and posture of the particular case, and the impossibility of resolution by any semantic cataloguing.").

CONCLUSION

District courts should apply the political question doctrine to determine whether or not to issue a nationwide injunction. By using the political question doctrine as a guide, courts can provide a more consistent and transparent approach to nationwide injunction cases, promoting judicial restraint and avoiding overreach by the judiciary. While the political question doctrine is a flexible principle, the *Baker*[156] criteria provide clear standards for determining when an issue presents a political question, and courts have developed a body of case law interpreting and applying these criteria. By using these criteria to evaluate nationwide injunction cases, courts can determine whether an issue is beyond their purview and should be left to the other branches of government to resolve.

Applying the political question doctrine to nationwide injunction cases would also have several additional benefits, including reducing forum shopping and judge shopping, promoting stability and continuity in the legal system, and providing greater certainty of an administration's policies. By using the political question doctrine as a guide, district courts can play a more appropriate and effective role in the policymaking process, balancing the authority of the other branches of government and ensuring that the law is applied fairly and consistently.

---

156 369 U.S. at 210.