INTRODUCTION

To tame the, sometimes, harmful power of enormous platforms, we need to reconsider the mathematics of regulation. The law tends to treat the growth of a company linearly, while the power and harm of online activity increases at a much faster rate. We need to scale up the mathematics of regulation to deal with many of the problems of massive digital platforms.

Most laws either treat all regulated actors the same or assume that twice as large means only twice as powerful and twice as harmful. For example, penalties for causing harm often multiply the number of individuals harmed by a set dollar figure, assessing $10,000 for each victim wiretapped or around $40,000 for each child monitored without parental consent.

Linearly bound regulation fails to reflect how the power and harm of some digital actors increase at much more than a linear, proportional rate. In at least three important ways, a platform with one billion users is more than one hundred times more powerful and potentially harmful than a company of ten million users.

First, a linear model fails to offer a proper moral accounting of the way human misery scales. We might feel more impelled to prevent a small harm affecting one million victims out of one billion users than we would to prevent the same harm affecting only ten victims out of ten thousand users, even though they reflect the same rate of injury with the only difference being the size of the injurer. Second, purely digital platforms expand automatically into any territory that the Internet touches, meaning platform providers need not attend to local regulators and regulations.
Third, size begets power, particularly for artificial intelligence, meaning we can expect more from globe-spanning digital platforms.

Massive digital platforms thus raise significant concerns of potential harm that calls for a regulatory response that accounts for effects of size. From privacy to tort to contract to consumer protection to intellectual property laws, we should better account for the power and potential harm of size.

This essay proceeds in three parts. Part I surveys the current approach to the mathematics of regulation. Part II presents the argument for why the power and harm of large digital platforms grows at faster than a linear rate. Finally, Part III begins to sketch what a new system for regulating at scale would require.

I. THE MATHEMATICS OF REGULATION

For the most part, regulations that turn on a mathematical standard or impose a numerical penalty relate their quantitative values to the size of the regulated entity according to only one of two mathematical relationships: constant value or linearity. Some rules apply a single constant value of burden or obligation to a regulated company, regardless of the size of that company. Every company, large or small, must meet some numerical threshold for, say, occupational safety, food safety, or minimum wage.

Other regulations account for the size of a corporation through a linear relationship. Corporate tax law imposes a simple percentage tax rate, meaning the tax owed grows linearly with the amount of corporate income. Laws that assign statutory damages for violations tend to multiply a set dollar value by the number of violations or damages. Digital platforms have been hit in recent years by these provisions—for example, by violating user privacy.

---

3 E.g., 29 C.F.R. § 1910.95 (2009) (requiring noise protection for work environments according to a precise numerical formula involving decibel levels and hours of exposure per day).
4 E.g., 21 C.F.R. § 101.17(g)(7) (2017) (requiring a warning label for any juice product that has not been processed to achieve at least a 100,000-fold reduction in pathogens through pasteurization).
7 See Bert I. Huang, Surprisingly Punitive Damages, 100 VA. L. REV. 1027, 1031–32 (2014).
8 See Petition for Writ of Certiorari, Google v. Joffe, 134 S. Ct. 2877 (2014) (No. 13–1181) at 25 (noting possible statutory damages of $10,000 per incident for alleged illegal wiretapping as part of Google’s Street View program); Jeffrey Lambe, Applying VPPA to Online Video Privacy, INT’L ASS’N PRIVACY PROF’LS,
There are limited, narrow exceptions. Our personal income tax system implements a progressive crude step function, one which subjects higher income to higher tax rates. Some laws provide relief at the “small” end of the spectrum, providing carve outs for businesses beneath some threshold size. Similarly, the Fair Housing Act exempts small, owner-occupied housing units from housing discrimination law—colloquially known as the “Mrs. Murphy exemption.” Finally, and perhaps most pertinently, antitrust laws focus intently on scale and size, triggering enforcement only at the very high end of market concentration.

II. SCALE, PLATFORM POWER, AND HARM

This mathematical regulatory state of affairs fails to account for what we have learned about the power of size and scale of platforms. There are three reasons to turn to a greater than linear-bound regulation for massive platforms.

1. The accumulation of human misery. First, for regulations meant to protect individuals from harm, focusing only on an acceptable percentage rate of injury does not properly account for the way we ought to think about the accumulation of human misery. A proper moral accounting of human misery should understand what orders-of-magnitude differences in the size of the harmed population means.

If we are inattentive to scale, we might focus only on the rate of injury. Consider an action that injures only one out of every thousand users, meaning a rate of 0.1%. Is this the kind of injury that demands a response, perhaps civil liability or the creation of a new regulation to deter similar action? My claim is we cannot answer these questions without

---

considering the denominator used to calculate the rate. Injuring ten out of ten thousand users is not the same thing as injuring one million out of one billion. Ten victims, however regrettable, might be justified by the countervailing benefits of corporate activity. Ten victims might also be justified if the harm to each of them is slight. But those considerations ought to carry less weight when we are talking about subjecting one million victims to harm. Even if that 99.9% non-injury rate might be deemed reasonable at a smaller level, we might not tolerate allowing the infliction of harm on the equivalent of a city’s population.

2. Platform power. Second, before the rise of digital platforms, companies could not touch the lives of billions without fully incorporating themselves into local communities and subjecting themselves to local norms and regulations. To sell billions of burgers, bottles of cola, or barrels of oil, companies in the pre-platform economy had to hire a gigantic workforce of individuals spread across the globe. Physical presence required them to scale their regulatory obligations with their size. This is often not true with today’s platforms and would-be platforms because of the global reach of the Internet and the efficiencies of code. Platforms can grow to touch people around the world without any local, and maybe not even regional, corporate activity whatsoever. Instagram required only thirteen employees to reach thirty million users by the time Facebook acquired the service. Today, Facebook employs only one employee for every 85,000 users.

The ability of digital platforms to passively expand into a new country or community without studying, much less abiding by, local norms and regulations means that local values tend to get ignored or trampled. Facebook does not need to worry nearly as much about what the citizens and officials of Bangalore or Nairobi want and expect from a social media platform compared to Coca-Cola or Exxon. Tailoring regulation to the size of a platform is one way to force companies to attend to the local conditions of the places they enter through passive expansion.

This argument cuts against one-size-fits-all complaints that tend to get lodged against all large online companies. Some corporate giants traffic solely in bits—they sell purely online services with no connection to the physical world of atoms. Think Facebook, Google (for the most part), and Twitter. Other companies, such as Apple, Uber, and Amazon’s

---


e-commerce service, are companies that, for the most part, move atoms, even if they facilitate their activity through online services. The complaint that massive companies can act with impunity in the face of local regulators and regulation rings truer against purely digital companies than against the giants that traffic in atoms. Thus, my argument that regulation should scale at a supra-linear rate with size applies with a bit less force against Apple and Uber than against Facebook and Google.

To be clear, I am quite worried about the potentially abusive power wielded by companies like Apple, Amazon, and especially, Uber. I have been an outspoken critic of Uber, which I have likened to an organized crime syndicate, one which has poured enormous resources into evading local regulation at every turn. One reason I worry about atom-bound companies is because of two other factors I offer in this part: how scale implicates the moral accounting of harm and the power of artificial intelligence. But these companies seem in many ways to be nothing more than the digital scions of ethically compromised corporate giants from the past, from the robber baron railroad titans to more recent moves by big oil and big pharma. Like these companies, Uber has had to scratch and claw its way into every new jurisdiction. Similarly, every time Apple opens a store in a new country, it subjects itself and its employees to the power of a new sovereign, which forces it to think much more proactively about local law.

Facebook, Google, and Twitter rarely feel the same level of pressure from local regulators. They need not engage at the same level with local regulation, and they can often avoid paying little more than lip service to national and regional regulation. Only when the number of users in a given country becomes economically significant to the company, and only when a local regulator decides to regulate these actors, do these companies need to pay serious attention to the law of a place.

3. The power of AI. Third, scale powers artificial intelligence (AI). Research suggests that the power of AI techniques, such as neural

---

15 Although my most pointed critiques have not been documented, I am on the record referring to Uber’s “fundamentally corrupt” nature. Paul Ohm, The Microsoft Design Decisions that Caused this Mess, JUST SECURITY (Feb. 21, 2018), https://www.justsecurity.org/52805/microsoft-design-decisions-caused-mess/ [https://perma.cc/QBR4-C6HU].


17 Id.

18 See Paul Ohm, We Couldn’t Kill the Internet if We Tried, 130 HARV. L. REV. F. 79 (2016).
networks and other forms of machine learning, do not proceed linearly.\(^{19}\) An algorithm trained on a million data points beats the accuracy of one trained on only one thousand data points by much more than three orders of magnitude. The power of a harmful AI can increase at a much higher rate than linear growth.

Unlike the first two reasons, the effect of scale on the power of AI provides a positive rather than negative reason to regulate more aggressively at scale. We can, and thus should, expect more and better regulatory compliance from companies with massive pools of data on which to train their AI.

In fact, this essay was first motivated by a Twitter exchange I had with Facebook’s then-Chief Security Officer, Alex Stamos. Stamos bemoaned the fact that Facebook’s efforts to automate the detection of child pornography for take down had suffered a public relations firestorm after some process automatically removed a picture of the iconic “napalm girl” photograph taken during the Vietnam War by Nick Ut.\(^{20}\) Stamos—without naming the napalm girl example directly—complained to me about the “difficulties of content moderation where the 99% of correct decisions go unnoticed but the 1% of mistakes create multi-week press cycles and calls for regulation.”\(^{21}\) Google’s then-copyright expert Fred Von Lohmann, who no doubt had given a lot of thought to the difficulty in automating takedown, clicked the like button on this tweet.\(^{22}\)

I am not criticizing Facebook for failing to develop an automated solution that could identify child pornography generally without flagging a false positive for this particular image, which after all does depict a full-frontal image of a naked child. It cannot be easy to train an AI for the very few exceptions that should be triggered because of historical significance.


\(^{22}\) Id.
What struck me about the complaint was the 99% figure. I understand that this was something spontaneously uttered on Twitter and not necessarily a precisely measured figure. But for the sake of argument, assume Stamos meant what he tweeted. Given the scale of Facebook’s user base and the massive number of photos that Facebook has to use as training data, 99% accuracy seems wholly insufficient. It seems unacceptable for Facebook to miss the mark on 1% of photos, measuring either false positives or false negatives. One percent inaccuracy also seems insufficient for Google’s processes for automatically flagging allegedly infringed copyrighted materials.

Companies like Google and Facebook are able to train automated detection algorithms on massive amounts of training data. We should expect the models that result to be much better than 99% accurate. Because of the way AI power increases with scale, we might expect companies like these to be accurate 99.99% or 99.999% of the time. Our expectations of accuracy might grow with the scale of the training data. For every order-of-magnitude increase in size, we might expect a correlative order of magnitude improvement in the accuracy of the detection.

The idea that giant platforms will soon use powerful AI to solve monumentally vexing problems has been embraced by another Facebook employee, the ur-Facebook employee, CEO and founder Mark Zuckerberg. In the wake of the Cambridge Analytica fiasco, Zuckerberg testified to committees of both chambers of Congress. He returned repeatedly to AI as the savior technology that would help Facebook solve some bafflingly difficult problems. For example, he predicted to Congress that “[h]ate speech—I am optimistic that over a five-to-ten-year period we will have AI tools that can get into some of the linguistic nuances of different types of content to be more accurate, to be flagging things to our systems . . . .” We should take Zuckerberg at his word, but we should require his savior AI to exhibit extremely high degrees of accuracy given the benefits of scale.

---


24 Id.


III. THE NEW MATHEMATICS OF REGULATING AT SCALE

For all of these reasons, platform power and platform harm grow more than linearly with platform size. We need new approaches to regulations that address the power of platforms that touch billions of human beings, approaches that scale regulatory obligations with size on a polynomial \( n^2 \) or even exponential \( 2^n \) basis.\(^{27}\) Vitally, we should embrace supra-linear regulation throughout the lifecycle of a company, not only at the rough “go or don’t go” circuit-breaker style regulation of antitrust law. From consumer protection to information privacy to tort and contract law, the larger a platform becomes, the stricter the standards it faces should be.

Most fundamentally, some standards of care ought to be subjected to an “orders of magnitude” rule. For every order of magnitude growth in the number of users, the standard of care to which you are subjected should grow at a faster-than-linear rate. We should consider holding a company like Equifax to a higher standard of care for securing its networks than we would apply to a smaller company.\(^{28}\) The Equifax data breach exposed a staggering number of individuals, perhaps as many as 150 million Americans, to the risk of identity theft and other information harm.\(^{29}\) Even if litigation reveals that Equifax’s security practices lived up to what we might expect from a mid-sized company, it should be unacceptable for a company to harm so many people.

For similar reasons, Google/YouTube should labor under a much higher accuracy standard for how it applies its Content ID copyright infringement detection system or its Right to be Forgotten system in Europe than a much smaller search engine or video service.\(^{30}\) All three

\(^{27}\) To give the reader a sense of the growth represented by these mathematical expressions: A tenfold increase in size means a ten times increase in a linearly bound value, a one hundred times increase for \( n^2 \), and more than one thousand times increase for \( 2^n \). At one hundred times increase in size, the difference is even greater, with a 10,000 increase for \( n^2 \) and more than \( 1 \times 10^{30} \) (one nonillion) increase for \( 2^n \).


\(^{30}\) MEG LETA JONES, CTRL+Z: THE RIGHT TO BE FORGOTTEN (2016); Matthew Sag, Internet Safe Harbors and the Transformation of Copyright Law, 93 NOTRE DAME L. REV. 499, 540–41 (discussing the Content ID program).
reasons offered in Part II support this conclusion: Google’s platform implicates billions of users, meaning we should worry about copyright infringement or privacy harm at scale; Google can amass millions of users in faraway lands without proactively thinking about local norms and regulations; and the company uses the benefits of scale to build some of the most sophisticated AI systems ever developed. For all of these reasons, we should hold Google to a much higher standard of care and accountability than we would most other companies.

One way to implement an orders-of-magnitude rule would be to borrow a concept from the field of reliability engineering—the discipline that studies how to design systems for mission critical situations that remain available in the face of complex risks. Reliability engineers who focus on services, like data centers, measure the availability of those services using what is known as the “nines” approach. Data centers are buffeted by hardware failures, weather conditions, and other risks that knock them offline. Systems that are engineered with enough redundancy to remain available despite frequent hard disk failures and winter storms are often marketed as “high availability,” and the level of availability is often touted (and promised via contract) using a “nines” shorthand. A system promised to remain up 99.999% of the time is marketed as “five nines.” A five nines system will be down no more than 5.26 minutes each year. Those who purchase high availability data center services must calculate the tradeoffs between four nines, five nines, six nines, etc. service, with each additional nine adding an order-of-magnitude to the promise of reliability for a higher price.

We could imagine a similar order-of-magnitude obligation increase that is connected to order-of-magnitude growth in scale. A small startup that has not yet signed up one million users might be obligated to maintain industry-standard but not world-class data security. But a global platform with more than one billion users might be held to the strictest data security standards money can buy.

We could apply this logic to penalties or fines for statutory violations as well. Platforms of a billion users that process information about children under age thirteen should owe more than the paltry $41,484

---


32 *Id.*

33 A year has 60 minutes/hour x 24 hours/day x 365 days/year = 525,600 minutes per year. A five nines promise guarantees no more than 525,600 x 0.00001 = 5.256 minutes per year of downtime.

per COPPA violation to account for the power of scale.\footnote{Press Release, FTC, FTC Publishes Inflation-Adjusted Civil Penalty Amounts (Jan. 23, 2018), https://www.ftc.gov/news-events/press-releases/2018/01/ftc-publishes-inflation-adjusted-civil-penalty-amounts [https://perma.cc/AC4G-2S96] (adjusting maximum fining authority for FTC-enforced statutory violations).} We, of course, should worry about companies of any size that build dossiers about vulnerable children, but we should do more to discipline companies that can collect such information about tens or hundreds of millions of children, to account for the way harms accrue with scale. Similarly, platforms found to violate the Wiretap Act should be fined more than the statutorily prescribed $10,000 per violation.\footnote{18 U.S.C. § 2520(c)(2) (stating statutory damages for Wiretap Act civil violations of “whichever is greater of $100 for each day of violation or $10,000”).} For example, we might apply such small fines to companies under ten million users, but for companies with more than one billion users, we might gradually increase the fines per violation to, perhaps, $500,000 per violation for the largest platforms.\footnote{This argument cuts against an emerging argument that statutory damages were never anticipated to cover such large classes of victims. Geoffrey S. Stewart, Emerging Issues in Statutory Damages, JONES DAY PUBS (July 2011), [https://perma.cc/N8HL-BGUE] (“The expansion of class action litigation and the advent of the internet, in particular, have led to abuses of statutory damages remedies.”). Some have argued that courts should decrease the net award for large classes, perhaps by pooling some of their damages concurrently. Huang, supra note 7.}

These orders of magnitude effects need not be so quantitatively focused. We might simply trigger qualitatively different regulatory obligations on massive platforms. New legislation could require a “comprehensive data privacy and security program,” akin to what the FTC extracts in settlements for privacy or security violations, for any platform that reaches the ten million user level.\footnote{Daniel J. Solove & Woodrow Hartzog, The FTC and the New Common Law of Privacy, 114 COLUM. L. REV. 583, 617-18 (2014) (describing FTC’s consent decree provisions requiring comprehensive security and privacy programs).} Under the GDPR, the EU could withhold the “one-stop shop” model of a single “lead supervisory authority” for massive platforms, freeing each individual country’s supervisory authority to regulate data protection in ways that reflect each nation’s individual context.\footnote{Commission Regulation 2016/679, General Data Protection Regulation, art. 25 2016 O.J. (L 127) (EU) (“Competence of the lead supervisory authority”). See also Art. 29 Data Protection Working Party, Guidelines for Identifying a Controller or Processor’s Lead Supervisory Authority (revised and adopted Apr. 5, 2017).}

Finally, consider how thinking about regulating at scale would enrich the emerging debate around using antitrust law as a way to address some of the problems of platforms, including by other contributors to this
Antitrust law has a long pedigree, and its gatekeepers—scholars, practitioners in the antitrust bar, and policymakers who care about the subject—tend to be firmly entrenched in economic models. Additionally, antitrust law seems focused on avoiding only the most extreme examples of a narrowly defined class of harms. Regulating at scale can happen much more broadly and far beyond antitrust law because new regulations of large platforms can be calibrated to be less drastic than those from antitrust.

CONCLUSION

A regulation that grows only at the rate of a company’s user base or revenue cannot always serve its purposes given the way platform power grows. The time has come to ramp up regulation on platforms that grow to sizes once unimaginable in order to help stem the worst effects of the platform economy.

---

See Maurice E. Stucke, Should We Be Concerned about Data-opolies?, 2 GEO. L. TECH. REV. 275 (2018); Lina Khan, Sources of Tech Platform Power, 2 GEO. L. TECH. REV. 325 (2018).