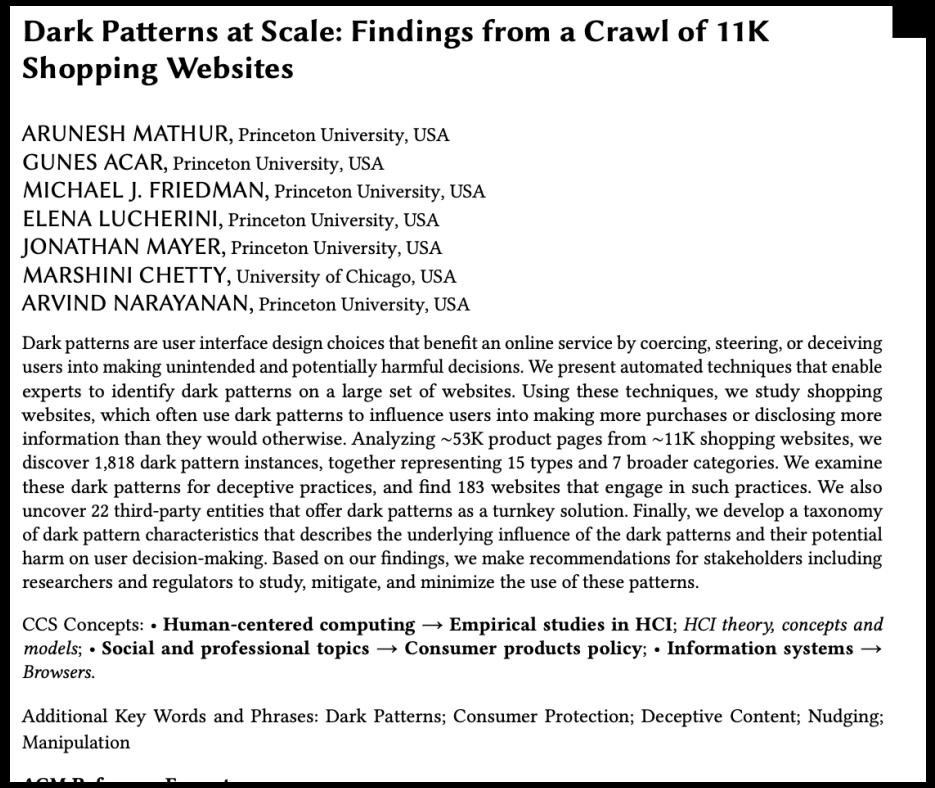
Shining a Light on Dark Patterns

Lior J. Strahilevitz
Sidley Austin Professor of Law
University of Chicago

Dark Patterns Are Widespread



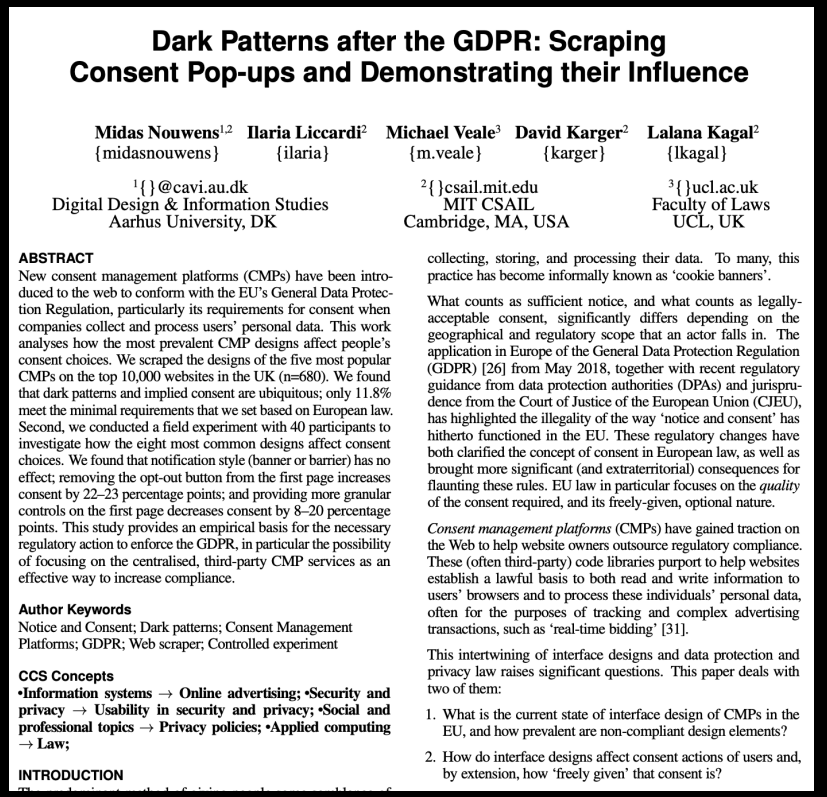
Dark Patterns at Scale: Findings from a Crawl of 11K Shopping Websites

ARUNESH MATHUR, Princeton University, USA
GUNES ACAR, Princeton University, USA
MICHAEL J. FRIEDMAN, Princeton University, USA
ELENA LUCHERINI, Princeton University, USA
JONATHAN MAYER, Princeton University, USA
MARSHINI CHETTY, University of Chicago, USA
ARVIND NARAYANAN, Princeton University, USA

Dark patterns are user interface design choices that benefit an online service by coercing, steering, or deceiving users into making unintended and potentially harmful decisions. We present automated techniques that enable experts to identify dark patterns on a large set of websites. Using these techniques, we study shopping websites, which often use dark patterns to influence users into making more purchases or disclosing more information than they would otherwise. Analyzing ~53K product pages from ~11K shopping websites, we discover 1,818 dark pattern instances, together representing 15 types and 7 broader categories. We examine these dark patterns for deceptive practices, and find 183 websites that engage in such practices. We also uncover 22 third-party entities that offer dark patterns as a turnkey solution. Finally, we develop a taxonomy of dark pattern characteristics that describes the underlying influence of the dark patterns and their potential harm on user decision-making. Based on our findings, we make recommendations for stakeholders including researchers and regulators to study, mitigate, and minimize the use of these patterns.

CCS Concepts: • **Human-centered computing → Empirical studies in HCI**; *HCI theory, concepts and models*; • **Social and professional topics → Consumer products policy**; • **Information systems →** *Browsers.*

Additional Key Words and Phrases: Dark Patterns; Consumer Protection; Deceptive Content; Nudging; Manipulation



Dark Patterns after the GDPR: Scraping Consent Pop-ups and Demonstrating their Influence

Midas Nouwens[1,2] {midasnouwens}
Ilaria Liccardi[2] {ilaria}
Michael Veale[3] {m.veale}
David Karger[2] {karger}
Lalana Kagal[2] {lkagal}

[1]{}@cavi.au.dk
Digital Design & Information Studies
Aarhus University, DK

[2]{}csail.mit.edu
MIT CSAIL
Cambridge, MA, USA

[3]{}ucl.ac.uk
Faculty of Laws
UCL, UK

ABSTRACT
New consent management platforms (CMPs) have been introduced to the web to conform with the EU's General Data Protection Regulation, particularly its requirements for consent when companies collect and process users' personal data. This work analyses how the most prevalent CMP designs affect people's consent choices. We scraped the designs of the five most popular CMPs on the top 10,000 websites in the UK (n=680). We found that dark patterns and implied consent are ubiquitous; only 11.8% meet the minimal requirements that we set based on European law. Second, we conducted a field experiment with 40 participants to investigate how the eight most common designs affect consent choices. We found that notification style (banner or barrier) has no effect; removing the opt-out button from the first page increases consent by 22–23 percentage points; and providing more granular controls on the first page decreases consent by 8–20 percentage points. This study provides an empirical basis for the necessary regulatory action to enforce the GDPR, in particular the possibility of focusing on the centralised, third-party CMP services as an effective way to increase compliance.

Author Keywords
Notice and Consent; Dark patterns; Consent Management Platforms; GDPR; Web scraper; Controlled experiment

CCS Concepts
•**Information systems → Online advertising;** •**Security and privacy → Usability in security and privacy;** •**Social and professional topics → Privacy policies;** •**Applied computing → Law;**

INTRODUCTION
collecting, storing, and processing their data. To many, this practice has become informally known as 'cookie banners'.

What counts as sufficient notice, and what counts as legally-acceptable consent, significantly differs depending on the geographical and regulatory scope that an actor falls in. The application in Europe of the General Data Protection Regulation (GDPR) [26] from May 2018, together with recent regulatory guidance from data protection authorities (DPAs) and jurisprudence from the Court of Justice of the European Union (CJEU), has highlighted the illegality of the way 'notice and consent' has hitherto functioned in the EU. These regulatory changes have both clarified the concept of consent in European law, as well as brought more significant (and extraterritorial) consequences for flaunting these rules. EU law in particular focuses on the *quality* of the consent required, and its freely-given, optional nature.

Consent management platforms (CMPs) have gained traction on the Web to help website owners outsource regulatory compliance. These (often third-party) code libraries purport to help websites establish a lawful basis to both read and write information to users' browsers and to process these individuals' personal data, often for the purposes of tracking and complex advertising transactions, such as 'real-time bidding' [31].

This intertwining of interface designs and data protection and privacy law raises significant questions. This paper deals with two of them:

1. What is the current state of interface design of CMPs in the EU, and how prevalent are non-compliant design elements?
2. How do interface designs affect consent actions of users and, by extension, how 'freely given' that consent is?

- Teams of researchers have documented the prevalence of dark patterns, Arunesh Mathur and co-authors (2019) (US), and Midas Nouwens and co-authors (2020) (Europe)

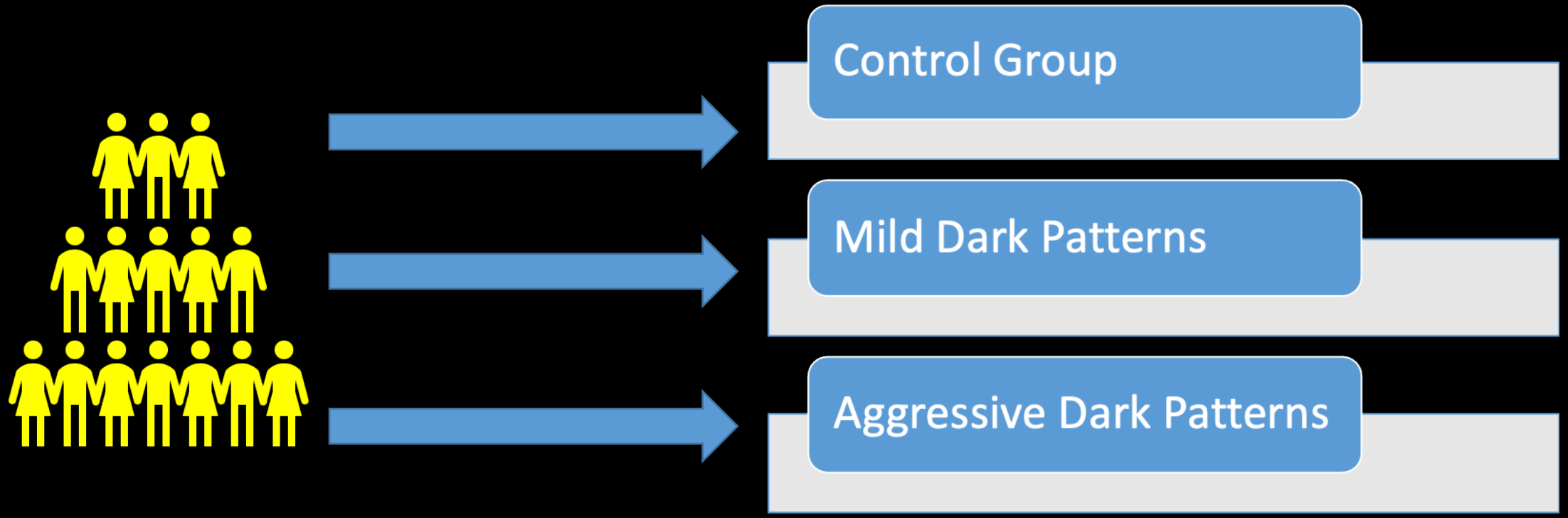
Knowing Dark Patterns are Prevalent Implies They Are Effective but Does Not Prove It

- Jamie Luguri and I tested dark patterns experimentally on large, census-weighted samples of American adults. Respondents look like the portion of the US adult population that is online (census weighted for race, age, gender, region, education).

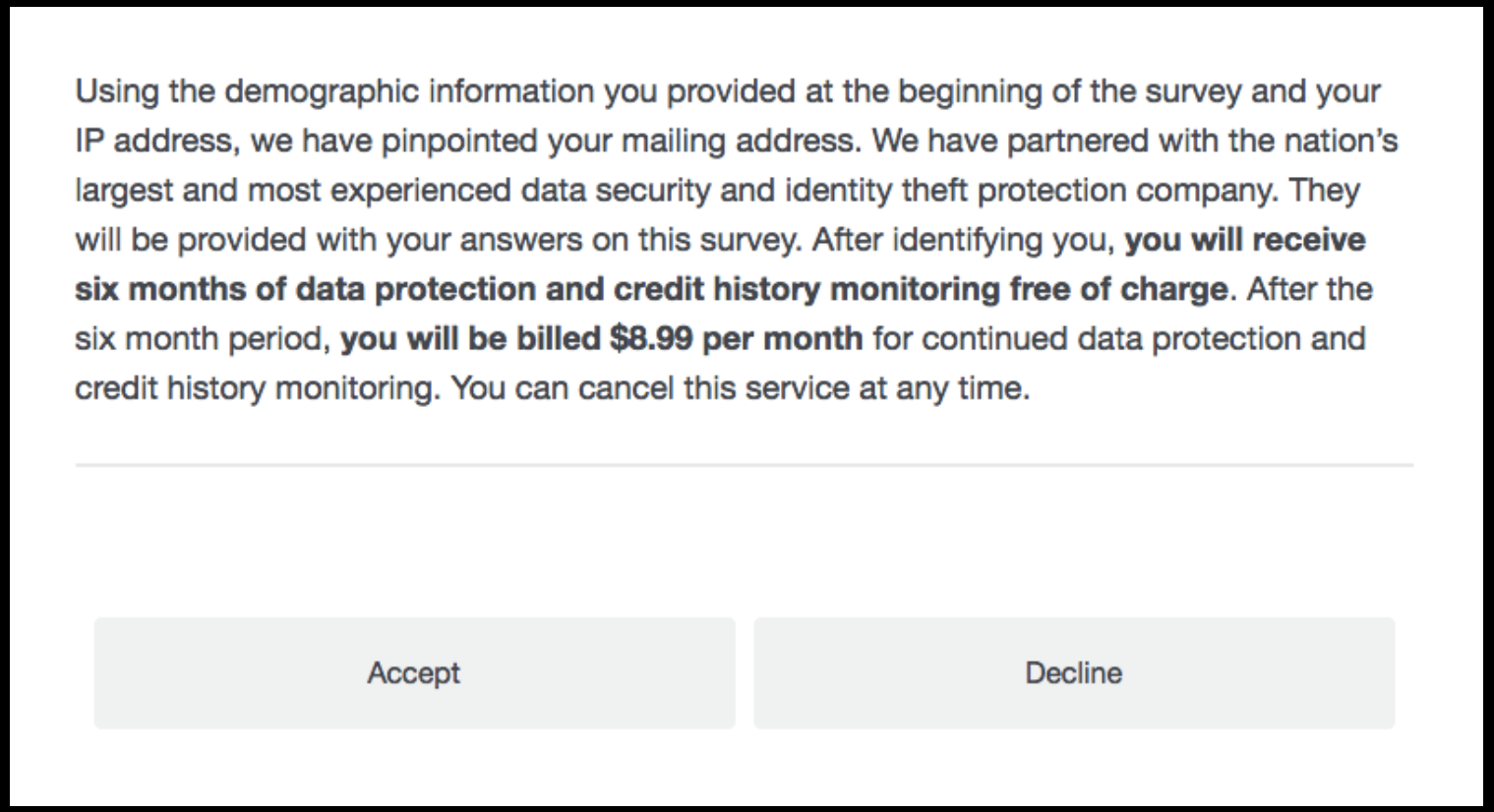
Study 1: Our Experimental Set-up

- In a fifteen-minute pre-registered experiment, subjects spent the first ten minutes supplying demographic information about themselves and answering a series of survey questions about their privacy preferences
- After this part of the survey was complete, subjects saw a screen that said “Calculating your privacy propensity score …” Following a short delay, subjects were told that our algorithm had identified them as having a “heighted concern about their privacy”
- Subjects told that using their IP address, phone number, & other information they had supplied, we had identified them
- Subjects told we had partnered with nation’s largest identity theft protection company and signed them up for a data protection plan. After free trial period they would be charged a monthly fee but they could cancel at any time

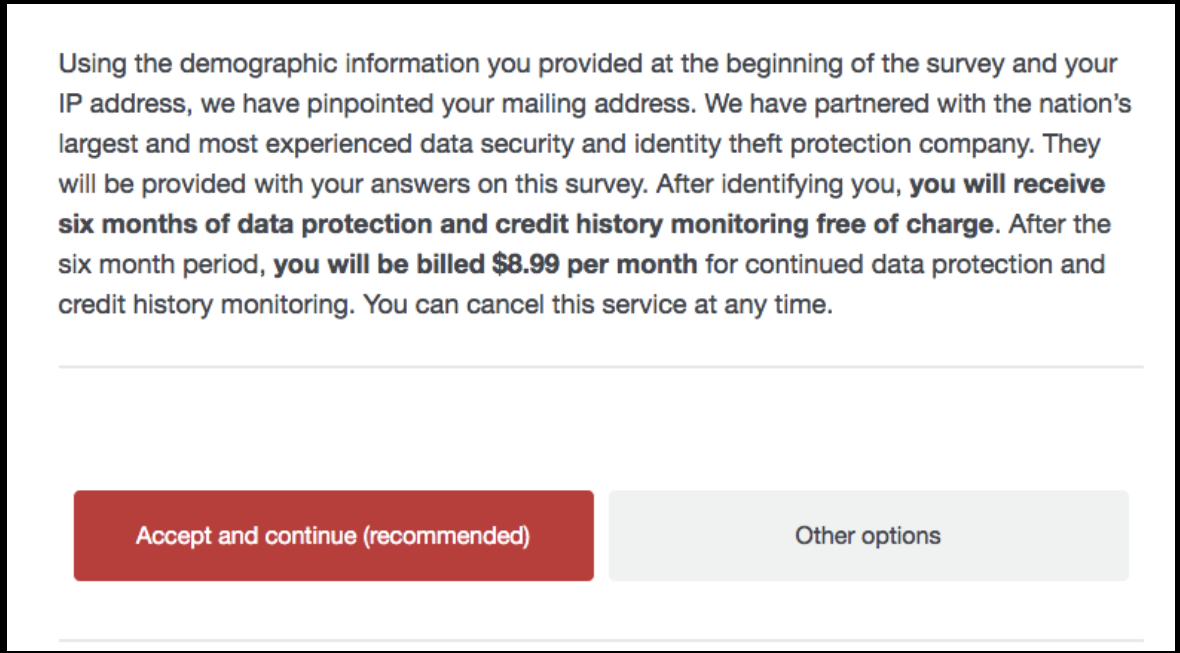
Subjects Randomly Assigned to One of Three Conditions

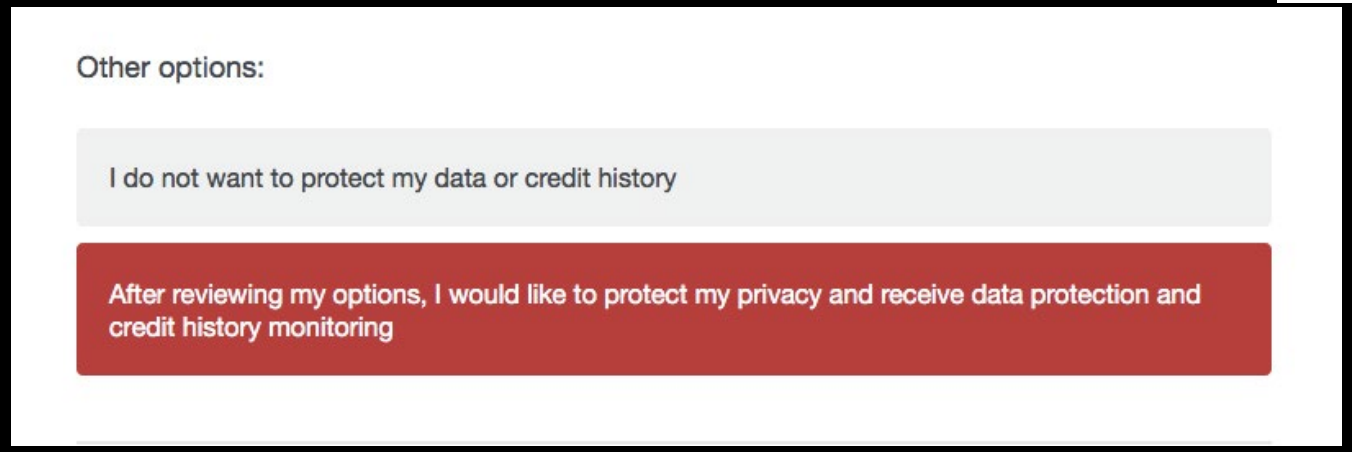


Control Group Condition



Mild Dark Patterns Condition

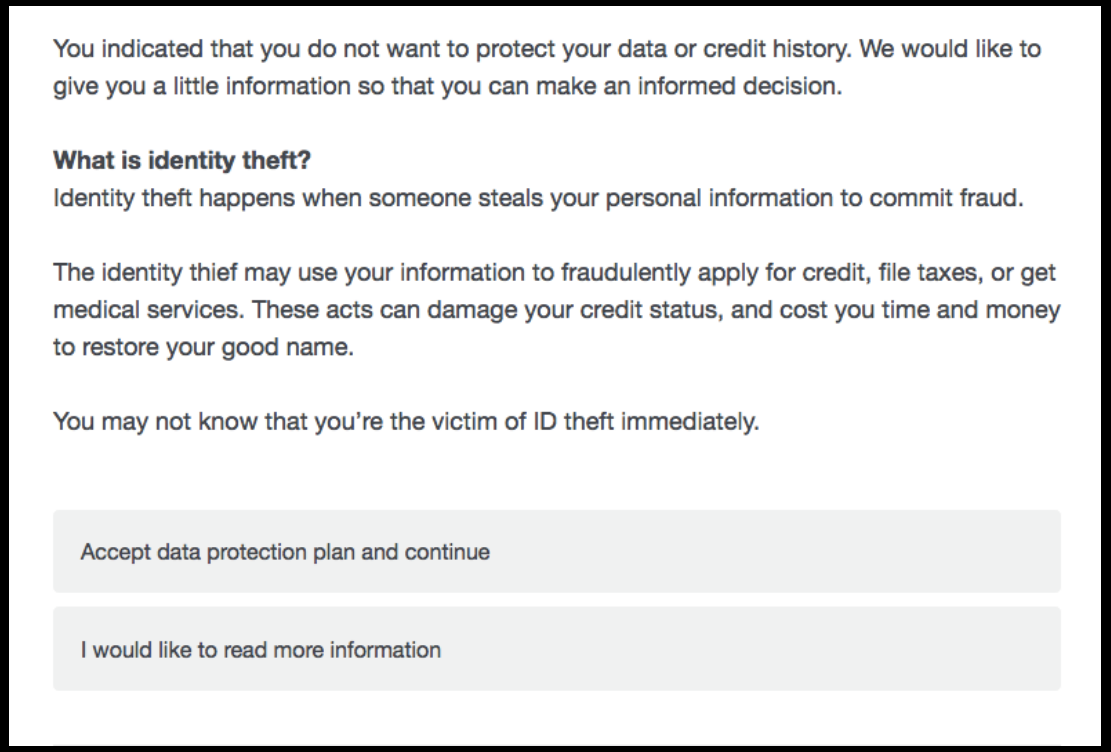




Aggressive Dark Patterns Condition – First Two Screens Identical to Mild Dark Pattern

- Do you wish to accept or decline the data protection plan?
  - ☑ Accept and Continue (Recommended)
  - ❑ Other options
- Selecting “Other options” led to this prompt:
  - ❑ I do not wish to protect my data or credit history
  - ❑ After reviewing my options, I would like to protect my privacy and receive data protection and credit history monitoring

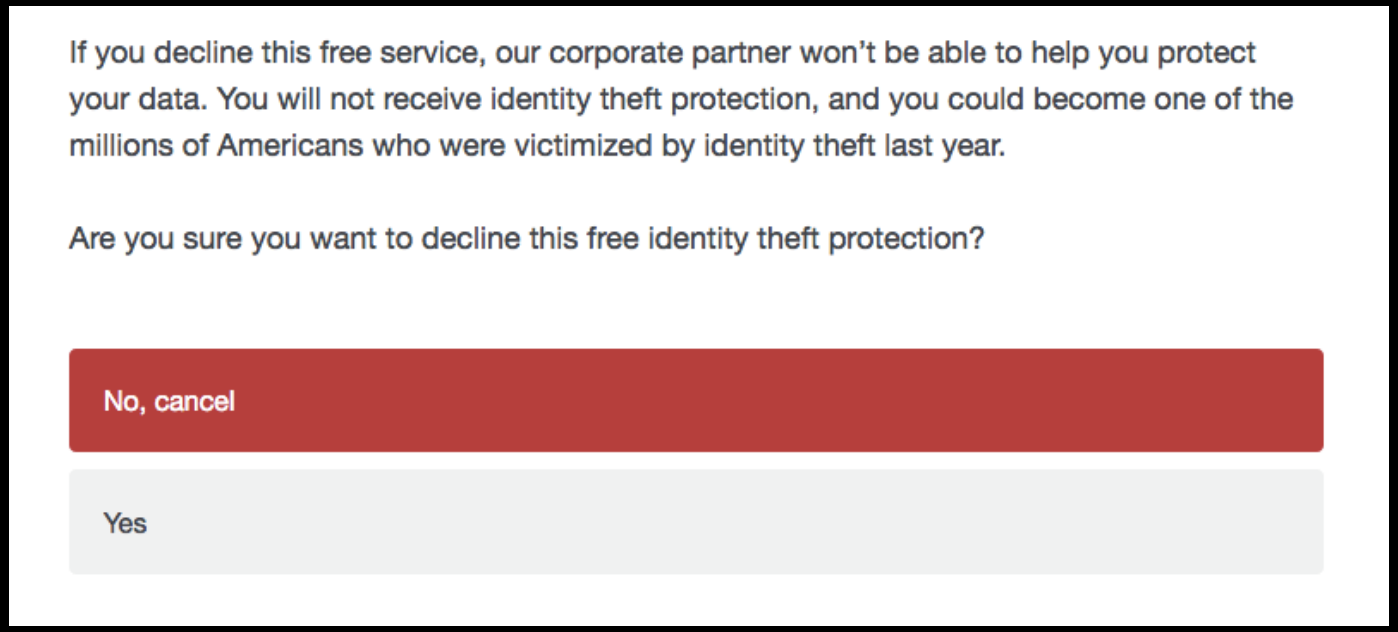
Aggressive Dark Patterns – Screens 3 to 5



- (selecting lower box meant subjects would see two additional similar screens with more text, the same options & 10-second countdown timer)

Aggressive Dark Patterns – Screen 6

- Subjects who declined on the third delay screen saw this prompt:



- Respondents who selected Yes advanced to a final (tell us why you declined) screen; Respondents who selected "No, cancel" were treated as having accepted the data protection plan

Final Screens Across All Conditions

- Please describe your current mood (using a 7-point Likert scale)
- ”Some survey participants may be contacted to do a follow up survey by the same researchers. Are you interested in potentially participating?” (7-point scale of responses)
- “How free did you feel to refuse the offered data protection and identity theft plan?” (7-point scale)
- ”Do you have any questions or comments for the researchers?”
- After this screen, all participants were fully debriefed on the experiment, and the purpose of the deception was explained

How Effective Were the Dark Patterns?

Condition	Acceptance Rate	Adjusted Acceptance Rate (treats experiment drop-outs as declines)
Control Group	11%	11%
Mild Dark Pattern	26%	25%
Aggressive Dark Pattern	42%	37%

Equity Concerns: Less Educated Americans Are More Vulnerable to Dark Patterns

- We analyzed a host of demographic factors to discover which ones are associated with vulnerability to dark patterns
- We predicted that less educated subjects would be more readily manipulated by dark patterns. The data supports that hypothesis
  - In the control group condition, education is not significantly correlated with accept / decline decisions
  - In the dark pattern conditions, less educated subjects were significantly more likely to accept the plan
  - Controlling for income and other demographics, less educated subjects are significantly more likely to accept in mild dark pattern condition but not in aggressive dark pattern condition

Data on Respondents' Mood & Sentiment

	Control	Mild	Aggressive
Mood (1-5)	M = 2.96; SD = 1.61	M = 3.05; SD = 1.73	M = 3.94; SD = 2.06***
Anger Expressed %	5.70%	6.09%	12.82%***
Exit the Survey %	N/A	1.5%	11.1%***

Do Market Forces Deter the Use of Dark Patterns?

- There was no meaningful backlash from our subjects when we employed mild dark patterns; aggressive dark patterns generated a strong negative response
- Subjects who accepted data protection plan did not have moods, willingness to repeat affected by dark pattern treatment; mood effects generated entirely by subjects who declined the plan
- Means also varied significantly on freedom to refuse question (6.2 control; 5.8 mild; 4.7 aggressive) (7 = perfectly free to refuse).

Study 2 – Format

- 3,777 Experimental subjects – once again we used a census weighted sample based on race, age, region, education, and gender
- Each subject sees zero, one, two, or three dark patterns – order randomly varied. This allows us to see which dark patterns are especially potent
- Testing other prevalent dark patterns that we did not use in Experiment 1 (urgency, double negatives, small and harder to see print, social proof)
- We randomly varied the cost of the service between $8.99 and $38.99 per month, with a one-month free trial period

Study 2 – 4 Form, 5 Content Conditions

	Control	Recommended	Default	Obstruction
Control				
Scarcity				
Confirmshaming				
Social Proof				
Hidden Information				

Content Conditions – Study 2

- **Hidden Information.** Participants told that they would "receive one month of data protection and credit history monitoring free of charge*", and that "[t]erms and conditions apply." At the bottom of the page, the price information was included in small, grey font.
- **Social Proof.** "1,657 other participants have accepted this free month of data protection and credit history monitoring in the last three weeks. Would you like to accept the program and join them?"
- **Scarcity.** "Congratulations! You have been selected to receive one month of free data protection and credit history monitoring. But you must **ACT FAST**! We only have three trial memberships left and this offer will expire in the next 60 seconds."
- **Confirmshaming.** Option to decline the program was phrased as "I don't care about protecting my data or credit history."

Form Conditions – Study 2

- Participants in the control condition could either choose “Accept” or “Decline.”
- Those in the default condition had the “accept” answer preselected.
- Those in the recommendation condition could chose either “Accept (recommended)” or “Decline.”
- Those in the obstruction condition saw the choices as “Accept and continue” or “Other options.”
  - Those selecting ”Other options” were randomly assigned to short or long obstruction. Short = Yes / No choice. Long = Two Identity Theft information screens with 15 second countdown timers.

Double Negative Question – Half of Our Experimental Subjects Were Shown This

- After making their selection, some respondents were asked: “Would you prefer not to decline this free data protection and credit history monitoring?”
  - Respondents who wish to reject the data protection plan should select “No”
  - Question employs a confusing double negative

Acceptance Rate by Content Condition

Condition	Acceptance Rate (%)	Number of Respondents Accepting
Control Group	14.8%	191 (out of 1289)
Scarcity	14.3% $p = .78$	91 (out of 635)
Confirmshaming	19.6% $p = .008$	120 (out of 612)
Social Proof	22.1% $p < .001$	140 (out of 634)
Hidden Information	30.1% $p < .001$	183 (out of 607)

Acceptance Rate by Form Condition

Condition	Acceptance Rate (%)	Number of Respondents Accepting
Control Group	16.7%	216 (out of 1294)
Recommended	18.1% $p = .39$	156 (out of 861)
Default	20.1% $p = .045$	171 (out of 851)
Obstruction	23.6% $p < .001$	182 (out of 771)

Study 2 – Acceptance Rates by Condition, *p*-value Compared to Control / Control

	Control	Recommended	Default	Obstruction
Control	13.2%	15.1% $p = .46$	15.0% $p = .49$	19.5% $p = .03$
Scarcity	10.6% $p = .39$	10.8% $p = .41$	18.9% $p = .061$	17.4% $p = .19$
Confirm-shaming	20.5% $p = .02$	16.4% $p = .29$	21.0% $p = .012$	20.4% $p = .03$
Social Proof	19.0% $p = .053$	21.0% $p = .01$	21.4% $p = .009$	27.9% $p < .001$
Hidden Information	30.8% $p < .001$	28.7% $p < .001$	26.7 $p < .001$	34.5% $p < .001$

Effects of Double Negative Question

- Percentage of respondents who answered yes in response to trick / double negative question (thereby accepting the service) – 33.4%
  - But only half of these respondents (16.7% of total) told us they had accepted the program. So half the acceptances were from people who didn't realize what they had done.
- Highly significant increase in acceptance rate ($p<.001$)
- Respondents who spent more time on double negative question screen were significantly less interested in participating in follow-up research with us

Study 2 – Costs of Service Didn't Matter

Acceptance Rate in High Stakes ($38.99 / month) Condition:

17.3%

Acceptance Rate in Low Stakes ($8.99 / month) Condition:

19.8%

- This difference is not statistically significant ($p = .09$)
- Increased price had no significant effect on acceptance rates
- 73.7% of respondents said they were at least somewhat likely to cancel plan after the first month; 21.1% said they definitely would cancel. This probably significantly overstates cancellation levels.

Again, No Evidence of Consumer Backlash to Mild Dark Patterns

- Confirmshaming and social proof had no significant effect on mood, compared to control
- Hidden information and scarcity conditions significantly improved the mood of subjects compared to control
- Form condition (control, recommended, default, obstruction) had no significant effect on mood or willingness to participate in future research

Study 2 – Education Level Again Predicts Susceptibility to Dark Patterns

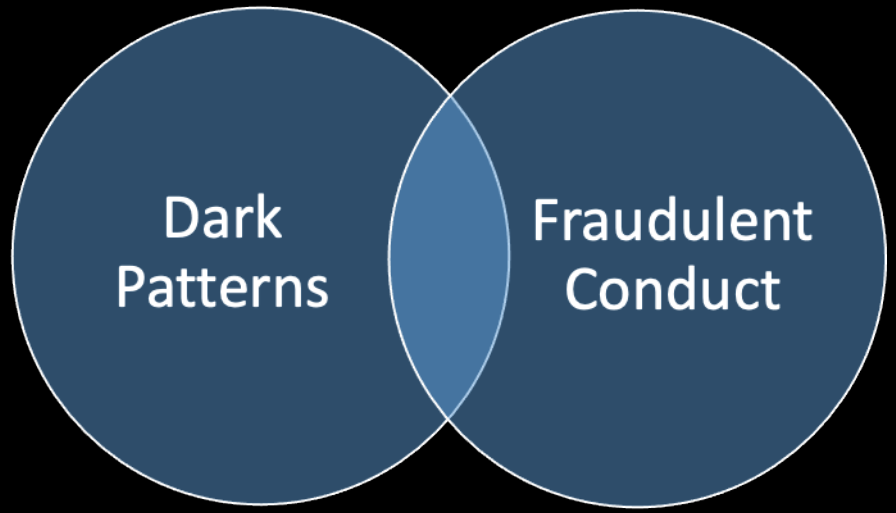
Education Level	Acceptance Rate in Control Condition (%)	Acceptance Rate in Treatment / Dark Pattern Condition (%)
High school diploma or yes	7.2%	17.8%
Some college or associate's degree	16.3%	22.2%
Bachelor's degree or graduate degree	17.8%	22%

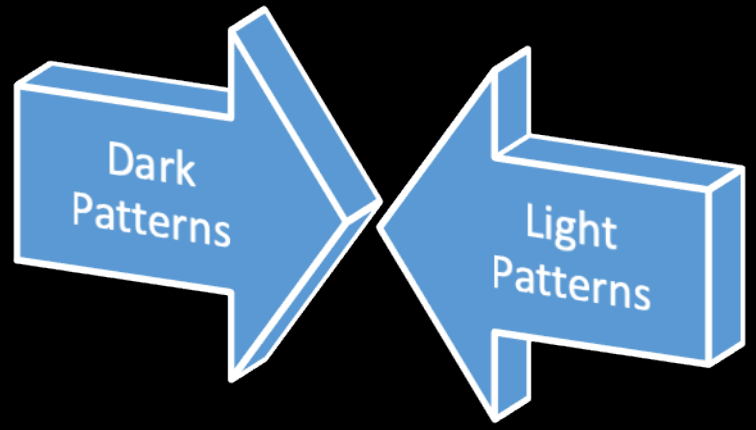
Key Take-aways

- It's the mild dark patterns that are most insidious. They significantly increased acceptance of a program with dubious benefits without alienating consumers or causing large numbers of them to log off
- Less educated subjects were particularly vulnerable to dark patterns
- Effects of dark patterns swamp effects of price changes
- Dark patterns vary substantially in terms of potency
- Dark patterns evidently have proliferated because they work
  - We're playing catch up with in-house social scientists doing A/B testing

Putting on My Legal Scholar Hat for a Moment….

- Two Points in Conclusion
  - Dark Patterns are not completely coterminous with fraud
  - A symmetry principle provides a transparent approach to regulation





Dark Patterns Are Not Inevitably Fraudulent

Usually entail fraud	Sometimes entail fraud	Rarely or never entail fraud
Sneaking items into cart	Visual interference	Obstruction
Hidden information	Unpopular defaults	Nagging
Bait & Switch		Confirmshaming
Disguised ads		
Trick questions		
False scarcity		
False social proof		

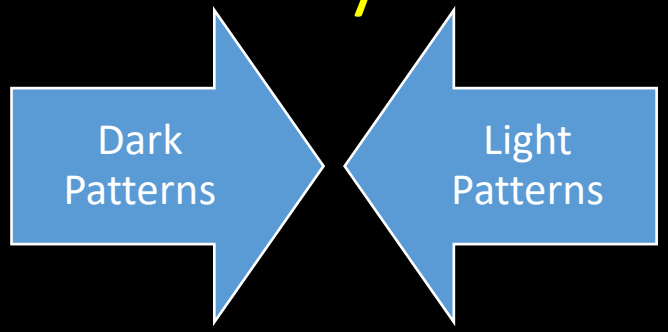
Potent Dark Patterns Are Not Inevitably Fraudulent

Usually entail fraud	Sometimes entail fraud	Rarely or never entail fraud
Sneaking items into cart	Visual interference	Obstruction
Hidden information	Unpopular defaults	Nagging
Bait & Switch		Confirmshaming
Disguised ads		
Trick questions		
False scarcity		
False social proof		

CPRA Language Does Not Require Fraud

- Section 1798.140(h) provides that “agreement obtained through use of dark patterns does not constitute consent.”
- Section 1798.140(l) defines “dark patterns” as “a user interface designed or manipulated with the substantial effect of subverting or impairing user autonomy, decisionmaking, or choice, as further defined by regulation.”
- Bottom line: Implementing a definition of dark patterns that does not include these manipulative non-fraudulent techniques would harm consumers and is inconsistent with the statutory definition provided in section 1798.140(l)

Towards a Symmetry Principle for Dark Patterns



- CCPA Regulations, § 999.315(h)(1), – "The business's process for submitting a request to opt-out shall not require more steps than that business's process for a consumer to opt-in to the sale of personal information after having previously opted out."
- Federal Trade Commission's October 22, 2021 Statement on Negative Option Marketing (e.g., a free trial that converts to a paid subscription at the conclusion of the trial period). FTC requires "cancellation mechanisms that are at least as easy to use as the method the consumer used to initiate the negative option feature"

Symmetry Principle Is a Strategy for Transparent, Workable Regulation

- Symmetrical Obstruction: “Are you sure?” prompts shown to those who want to decline a service would need to be presented to consumers who want to accept that service as well
- Symmetrical Nagging: If a company repeatedly asks consumers who initially disable location tracking to enable it, it must prompt consumers who initially enable it to consider disabling it with the same frequency
- Symmetrical Confirmshaming: One permissible option cannot be presented in a manner that employs more negative emotional language than the alternative permissible option. Compare to ballot initiative rules in elections

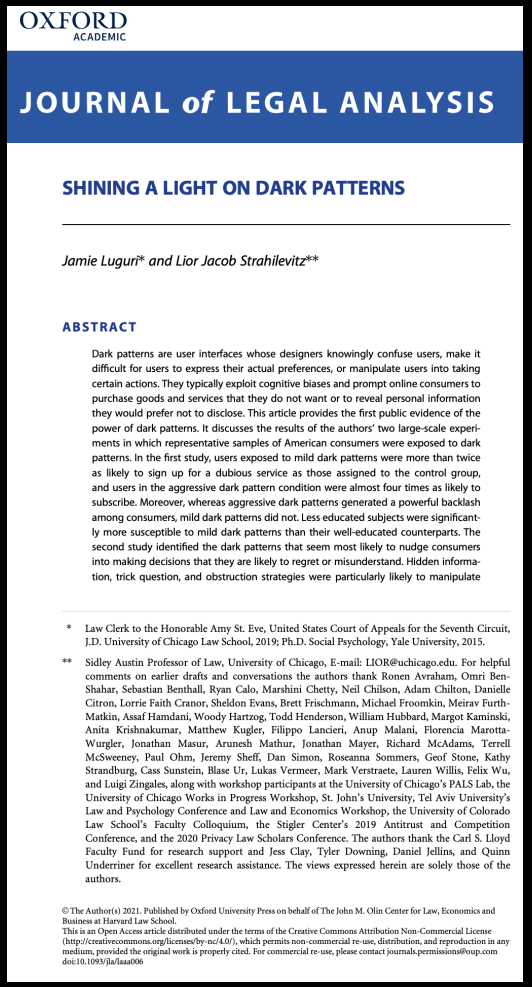
Symmetry Principle Is a Strategy for Transparent, Workable Regulation (continued)

- Symmetrical Social Proof: Data on number of consumers who accept offer must be accompanied by prominent data about number of consumers who decline the offer
- Symmetrical Information: Material information that makes consumers less likely to accept must be as visually prominent as material information that makes them more likely to accept
- Not every design needs to be symmetrical. It is ok to obstruct options that few consumers would prefer, or to bury information that few consumers deem relevant. But UX designers shouldn’t impede selection of popular options or the discovery of info that will be material to many consumers

For More Data & Detail

- Jamie Luguri & Lior Jacob Strahilevitz, *Shining a Light on Dark Patterns*, 13 Journal of Legal Analysis 43 (2021).
- You can download the article for free here:
  - bit.ly/darkpatternsarticle (all lower-case)

Or just type “Shining a Light on Dark Patterns” into your favorite search engine and download the paper from the Journal of Legal Analysis website. It’s an open-access journal.



OXFORD ACADEMIC

JOURNAL *of* LEGAL ANALYSIS

SHINING A LIGHT ON DARK PATTERNS

Jamie Luguri and Lior Jacob Strahilevitz***

ABSTRACT

Dark patterns are user interfaces whose designers knowingly confuse users, make it difficult for users to express their actual preferences, or manipulate users into taking certain actions. They typically exploit cognitive biases and prompt online consumers to purchase goods and services that they do not want or to reveal personal information they would prefer not to disclose. This article provides the first public evidence of the power of dark patterns. It discusses the results of the authors’ two large-scale experiments in which representative samples of American consumers were exposed to dark patterns. In the first study, users exposed to mild dark patterns were more than twice as likely to sign up for a dubious service as those assigned to the control group, and users in the aggressive dark pattern condition were almost four times as likely to subscribe. Moreover, whereas aggressive dark patterns generated a powerful backlash among consumers, mild dark patterns did not. Less educated subjects were significantly more susceptible to mild dark patterns than their well-educated counterparts. The second study identified the dark patterns that seem most likely to nudge consumers into making decisions that they are likely to regret or misunderstand. Hidden information, trick question, and obstruction strategies were particularly likely to manipulate

* Law Clerk to the Honorable Amy St. Eve, United States Court of Appeals for the Seventh Circuit, J.D. University of Chicago Law School, 2019; Ph.D. Social Psychology, Yale University, 2015.

** Sidley Austin Professor of Law, University of Chicago, E-mail: LIOR@uchicago.edu. For helpful comments on earlier drafts and conversations the authors thank Ronen Avraham, Omri Ben-Shahar, Sebastian Benthall, Ryan Calo, Marshini Chetty, Neil Chilson, Adam Chilton, Danielle Citron, Lorrie Faith Cranor, Sheldon Evans, Brett Frischmann, Michael Froomkin, Meirav Furth-Matkin, Assaf Hamdani, Woody Hartzog, Todd Henderson, William Hubbard, Margot Kaminski, Anita Krishnakumar, Matthew Kugler, Filippo Lancieri, Anup Malani, Florencia Marotta-Wurgler, Jonathan Masur, Arunesh Mathur, Jonathan Mayer, Richard McAdams, Terrell McSweeney, Paul Ohm, Jeremy Sheff, Dan Simon, Roseanna Sommers, Geof Stone, Kathy Strandburg, Cass Sunstein, Blase Ur, Lukas Vermeer, Mark Verstraete, Lauren Willis, Felix Wu, and Luigi Zingales, along with workshop participants at the University of Chicago’s PALS Lab, the University of Chicago Works in Progress Workshop, St. John’s University, Tel Aviv University’s Law and Psychology Conference and Law and Economics Workshop, the University of Colorado Law School’s Faculty Colloquium, the Stigler Center’s 2019 Antitrust and Competition Conference, and the 2020 Privacy Law Scholars Conference. The authors thank the Carl S. Lloyd Faculty Fund for research support and Jess Clay, Tyler Downing, Daniel Jellins, and Quinn Underriner for excellent research assistance. The views expressed herein are solely those of the authors.

doi:10.1093/jla/laaa006