

2015

## Putting Disclosure to the Test: Toward Better Evidence-Based Policy

Talia B. Gillis  
*Harvard University*

Follow this and additional works at: <http://lawcommons.luc.edu/lclr>



Part of the [Consumer Protection Law Commons](#)

---

### Recommended Citation

Talia B. Gillis *Putting Disclosure to the Test: Toward Better Evidence-Based Policy*, 28 Loy. Consumer L. Rev. 31 (2015).  
Available at: <http://lawcommons.luc.edu/lclr/vol28/iss1/3>

This Feature Article is brought to you for free and open access by LAW eCommons. It has been accepted for inclusion in Loyola Consumer Law Review by an authorized administrator of LAW eCommons. For more information, please contact [law-library@luc.edu](mailto:law-library@luc.edu).

# PUTTING DISCLOSURE TO THE TEST: TOWARD BETTER EVIDENCE-BASED POLICY

*Talia B. Gillis\**

I. Introduction .....	33
II. Consumer Financial Protection and the Prominence of Disclosure.....	38
A. The Goals of Consumer Financial Regulation .....	39
B. The Role of Disclosure .....	43
1. The Prominence of Disclosure as a Regulatory Tool.....	44
2. Why Disclosure May Benefit Consumers .....	46
a. Consumer Responses to Disclosure .....	46
b. Supply Side Effects of Disclosure .....	48
II. Why Test Financial Disclosures? .....	49
A. Changes in Theory of Consumer Behavior.....	50
B. Increased Stringency of Regulatory Impact Assessments.....	53
III. Examples of Financial Disclosure Regulatory Testing.....	57
A. EU: Key Investor Information Document .....	57
1. Methodology of Testing .....	61
2. Testing Results .....	62
3. Subsequent Experiments .....	64
B. US: CFPB Testing of Mortgage Disclosure .....	65
1. Methodology of Testing .....	67
2. Results .....	68
IV. Concerns with the Way Disclosures are Tested.....	69
A. Missing Benchmarks for Evaluation.....	70

---

\* Doctoral student at Harvard University. I am grateful for the helpful comments provided by Alexei Alexandrov, Rachel Bayefsky, Howell Jackson, Yotam Kaplan, Louis Kaplow, David Laibson, Steven Shavell, Cass Sunstein and participants at the Annual Meeting of the American Law and Economics Association and the Law and Economics Seminar at Harvard Law School.

B. Isolated Environment .....	73
C. Experiment Effects .....	76
D. Question Answering Format .....	77
1. Saliency .....	78
2. Comprehension Tests as “Under-Conclusive” .....	79
3. Comprehension Tests as “Over-Conclusive” .....	83
4. Creating Documents that Score Well .....	87
V. Changing the Way We Test Disclosure .....	87
A. Redesigning Testing Within Current Framework .....	88
1. Focusing on Decisions .....	88
2. Directly Testing Decisions .....	89
3. Benchmarks for Determining Improved Decisions .....	91
B. Testing Alternative Mechanisms .....	95
C. Expanding Testing Efforts .....	96
1. Real-life Testing .....	96
2. Retrospective Analysis .....	99
3. Dynamic Testing .....	100
VI. Conclusion .....	104

**F**inancial disclosures no longer enjoy the immunity from criticism they once had. While disclosures remain the hallmark of numerous areas of regulation, there is increasing skepticism as to whether disclosures are understood by consumers and do in fact improve consumer welfare. Debates on the virtues of disclosures overlook the process by which regulators continue to mandate disclosures. This article fills this gap by analyzing the testing of proposed disclosures, which is an increasingly popular way for regulators to establish the benefits of disclosure. If the testing methodology is misguided then the premise on which disclosures are adopted is flawed, leaving consumers unprotected. This article focuses on two recent major testing efforts: the European Union’s testing of fund disclosure and the Consumer Financial Protection Bureau’s testing of the integrated mortgage disclosures, which went into effect on August 1, 2015.

Despite the substantial resources invested in these quantitative studies, regulation based on study results is unlikely to benefit consumers since the testing lacks both external and internal validity. The generalizability of the testing is called into question since the isolated conditions of testing overlook the reality

of financial transactions. Moreover, the testing method mistakenly assumes a direct link between comprehension and improved decisions, and so erroneously uses comprehension tests.

As disclosure becomes more central to people's daily lives, from medical decision aids to nutritional labels, greater attention should be given to the testing policies that justify their implementation. This article proposes several ways to improve the content and design of quantitative studies as we enter the era of testing.

## I. INTRODUCTION

Consumers are increasingly required to make financial decisions in a world in which financial products are becoming significantly more complex.<sup>1</sup> With the increased burden on consumers, an issue that received a considerable amount of attention following the global financial crisis,<sup>2</sup> there is strong evidence that people are far from being perfect savers, investors and planners.<sup>3</sup> Consumers are not always sufficiently numerate or financially literate to assess financial products and services.<sup>4</sup> Moreover, people face various self-control problems and time-inconsistencies that make long-term decision-making difficult.<sup>5</sup> It is clear that financial institutions and financial service providers are not indifferent to consumer shortcomings, which are often

---

<sup>1</sup> For a discussion of the difference between consumers and investors, see Niamh Moloney, *The Investor Model Underlying the EU's Investor Protection Regime: Consumers or Investors?*, 13 EUR. BUS. ORG. L. REV. 169 (2012).

<sup>2</sup> See, e.g., Lauren Willis, *Decision-making and the Limits of Disclosure: The Problem of Predatory Lending: Price*, 65 MD. L. REV. 707, 718 (2006) (discussing the increased burden on consumers).

<sup>3</sup> For summaries of research, see Shlomo Benartzi & Richard Thaler, *Heuristics and Biases in Retirement Savings Behavior*, 21 J. ECON. PERSP. 81 (2007); Nicholas Barberis & Richard Thaler, *A Survey of Behavioral Finance*, in *Handbook of Economics of Finance*, 1052-1121 (George M. Constantinides, Milton Harris & Rene M. Stulz eds., 2003); John Campbell, *Household Finance*, 61 J. FIN. 1553 (2006).

<sup>4</sup> See Annamaria Lusardi, *Financial Literacy: An Essential Tool for Informed Consumer Choice?* (Nat'l Bureau of Econ. Res., Working Paper No. 14084, 2008), available at <http://www.nber.org/papers/w14084>; Jill Fisch & Tess Wilkinson-Ryan, *Why Do Retail Investors Make Costly Mistakes? An Experiment on Mutual Fund Choice* (2014), available at [http://scholarship.law.upenn.edu/faculty\\_scholarship/415](http://scholarship.law.upenn.edu/faculty_scholarship/415).

<sup>5</sup> See David Laibson, *Golden Eggs and Hyperbolic Discounting*, 112 Q.J. ECON. 443 (1997); Ted O'Donoghue and Matthew Rabin, *Doing it Now or Later*, AM. ECON. REV. 103 (1999).

exploited.<sup>6</sup>

While consumers face many challenges when making financial decisions, the regulatory structure of consumer finance continues, controversially,<sup>7</sup> to rely heavily on enhancing decision-making by improving disclosures. Opponents of disclosure have argued that disclosure is “ineffective,”<sup>8</sup> does not “result in good deliberate decision-making,”<sup>9</sup> and that it “has failed time after time, in place after place, in area after area, in method after method, and in decade after decade.”<sup>10</sup> Yet disclosure continues to play a central role in many domains including the regulation of consumer finance.

Understanding how regulators decide to require disclosures, despite their potential shortcomings, is crucial to any policy recommendation regarding mandated disclosure. In recent years, the process of adopting disclosures has changed significantly as regulators are required to provide stronger evidence of the benefits and effectiveness of proposed disclosure. Regulators now tend to recognize the need to replace abstract notions about how people process and use disclosures with the empirical study of their impact. Opponents of disclosure have overlooked this change in policy, claiming that the persistence of disclosure is a result of its

---

<sup>6</sup> See generally Xavier Gabaix & David Laibson, *Shrouded Attributes, Consumer Myopia, and Information Suppression in Competitive Markets*, 121 Q.J. ECON. 505 (2006); Jennifer Brown, Tanjim Hossain, & John Morgan, *Shrouded Attributes and Information Suppression: Evidence from the Field*, 125 Q.J. ECON. 859 (2010); Victor Stango & Jonathan Zinman, *Fuzzy Math, Disclosure Regulation and Credit Market Outcomes: Evidence from Truth-in-Lending Reform*, 24 REV. FIN. STUD. 506 (2011).

<sup>7</sup> See, e.g., Omri Ben-Shahar & Carl E. Schneider, *The Failure of Mandated Disclosure*, 159 U. PA. L. REV. 647 (2011); OMRI BEN-SHAHAR & CARL SCHNEIDER, *MORE THAN YOU WANTED TO KNOW: THE FAILURE OF MANDATED DISCLOSURE* (2014) [hereinafter *MORE THAN YOU WANTED TO KNOW*]; Richard Craswell, *Static Versus Dynamic Disclosures, and How Not to Judge Their Success or Failure*, 88 WASH. L. REV. 333 (2013); Willis, *supra* note 2, at 789; Ryan Bubb, *TMI? Why the Optimal Architecture of Disclosure Remains TBD*, 113 MICH. L. REV. 6 (2015).

<sup>8</sup> In the context of online disclosures, see Florencia Marotta-Wurgler, *Does Disclosure Matter?*, 168 J. INST. & THEORETICAL ECON. 94 (2011).

<sup>9</sup> See Willis, *supra* note 2, at 712.

<sup>10</sup> See Omri Ben-Shahar & Carl Schneider, *The Failed Reign of Mandated Disclosure*, REG BLOG (June 15, 2015), <http://www.regblog.org/2015/06/15/ben-shahar-schneider-failed-disclosure/>.

political appeal and the excessive belief in “disclosurism.”<sup>11</sup> In reality, regulators often do not merely assert the benefits of disclosure but try to establish the effectiveness of disclosure by commissioning consumer testing of disclosures.<sup>12</sup> Previous writing on disclosure, however, has failed to engage with the critical issue of the process by which disclosures are adopted by regulators, causing inconsistency between the scholarly debate on disclosure and the rule-making reality.

This article contributes to our understanding of the prevalence of disclosure by analyzing the current testing methodology used to support the adoption of disclosure, showing that it is inadequate and how it might be improved. I evaluate two examples of extensive attempts to test disclosure documents prior to their adoption: the testing of the Integrated Mortgage disclosure by the Consumer Financial Protection Bureau in the United States, which went into effect on August 1, 2015, and the testing of the Key Investor Information Document by the European Union, which was implemented in 2012. The use of quantitative studies as a means to justify the adoption of disclosures is expected to expand, given other disclosures that are currently on the agenda of the Consumer Financial Protection Bureau and other consumer financial regulators, such as those for payday lending, credit cards, and insurance.<sup>13</sup>

As empirical testing of financial disclosures is becoming the norm, greater thought must be dedicated to the *design* and *purpose* of this testing if the goal of disclosures is to benefit consumers.

---

<sup>11</sup> See *infra*, Part II.B. Ben-Shahar often uses the term “disclosurites” as people who refuse to abandon their belief in disclosure even when faced with evidence of it being ineffective. See MORE THAN YOU WANTED TO KNOW, *supra* note 7, at 6.

<sup>12</sup> For a summary of research methods used by the Federal Reserve Board to test consumer disclosures, see Jeanne Hogarth & Ellen Merry, *Designing Disclosures to Inform Consumer Financial Decisionmaking: Lessons Learned from Consumer Testing*, 97 FED. RES. BULL. 3 (2011) [hereinafter FEDERAL RESERVE BULLETIN 2011].

<sup>13</sup> See, e.g., Know Before You Owe: Credit Cards, CONSUMER FIN. PROT. BUREAU, <http://www.consumerfinance.gov/credit-cards/knowbeforeyouowe/> (last visited Sept. 30, 2015); Prepaid Cards: Help Design a New Disclosure, CONSUMER FIN. PROT. BUREAU, <http://www.consumerfinance.gov/blog/prepaid-cards-help-design-a-new-disclosure/> (last visited Sept. 30, 2015); Know Before You Owe: Student Loans Project, CONSUMER FIN. PROT. BUREAU, <http://www.consumerfinance.gov/students/knowbeforeyouowe/> (last visited Sept. 30, 2015).

First, regulators need to consider the aim of financial disclosures, such as whether disclosures are intended to make consumers fully informed, or whether they are meant to influence consumers in a less deliberative manner.<sup>14</sup> I argue that the failure of consumer financial regulators to adequately articulate the mechanism through which disclosures assist consumers creates confusion as to how to judge the effectiveness of disclosures. Second, regulators must analyze whether current testing methodologies coincide with this aim. One concern is that comprehension tests, which are becoming the standard for testing disclosures, may neglect to test the actual effect of disclosures on financial decisions. While regulators assume that improving consumer comprehension in the narrow sense leads to better decisions, improved decision-making may require a number of steps beyond narrow comprehension. Third, regulators need to examine and justify the setting in which they test disclosures given concerns regarding the extent to which experimental settings can provide information on real-life impact of disclosures. This could mean either shifting testing methodologies to real-life testing, such as randomized controlled trials, or designing experiments that attempt to better capture real-life situations.

The first part of this article provides background on the way in which disclosure has become a central regulatory tool for improving consumer protection, and the various premises on which this strategy relies. I argue that regulators have not always explicitly recognized how disclosure's effectiveness relies both upon consumers' understanding of the disclosure *and* their ability to apply the information effectively to a financial decision. The conflation of comprehension and improved decision-making is behind many of the concerns with current testing practices.

The second part of this article discusses the need to test proposed disclosures. It explains why regulators have only recently emphasized disclosure testing, despite the prevalence of financial disclosures for many decades. It argues that the mounting evidence on the shortcomings of disclosures and the incorporation of the insights of behavioral economics into policy, as well as the

---

<sup>14</sup> See, e.g., Bubb, *supra* note 7, at 115 (discussing the various ways disclosures may affect consumer behavior). Bubb uses the well-known distinction between System 1 and System 2 as a framework of discussing the two ways in which disclosure can affect behavior—either in a deliberative manner or in a fast and unconscious way. *Id.*

increased stringency of cost-benefit analysis and the recent emphasis of evidence-based policy, explain the rise of consumer testing.

The third part of this article explains the prevailing regulatory testing methodology through two recent examples of robust empirical testing of disclosure regulation. The first example is the European Union's testing of the summary prospectus for retail investment funds, Undertakings for the Collective Investment of Transferable Securities (UCITS). The second example is the Consumer Financial Protection Bureau's testing of integrated mortgage disclosures in the United States. In both cases, regulators sought to provide evidence to back their proposed financial disclosures through a quantitative study. This section provides background to the proposed policy, describes the methodology of the quantitative testing, and discusses the results. In both cases, the testing focused on consumers' comprehension of the information by asking questions about the content of the documents while overlooking the impact of the disclosures on actual decisions.

The fourth part of this article discusses concerns that current testing practices raise. Current testing has both external and internal validity problems. It lacks external validity because it overlooks the ways in which disclosure may have an effect in real life, and consequently the test results may not be generalizable. It does not test whether consumers will actually read the disclosure when not prompted by test questions. It does not test whether human interaction or other distractions impact the effect of disclosure. Furthermore, it does not test how the financial industry will respond to disclosure requirements, relying instead on the isolated and pro-consumer environment created by regulators testing the disclosure.

Current testing methodologies also raise internal validity concerns. Their current focus on narrow comprehension does not test the stated purpose of disclosure—to improve consumers' decisions. This section provides several reasons to believe that comprehension does not necessarily lead to improved decisions. Current testing efforts, which focus on consumer comprehension, do not therefore establish whether disclosures will of themselves improve consumer decisions. For example, relying on disclosures to make financial decisions requires financial literacy beyond the



narrow comprehension of the document as well as knowledge of the consumer's financial situation and how it affects the specific transactions being considered. Furthermore, the current testing design does not test additional mechanisms through which disclosure may assist consumers beyond making consumers fully informed, such as increasing financial awareness.

The article concludes by describing how regulators can improve disclosure testing, either by redesigning experimental testing or by using real-life testing through randomized control trials and retrospective analysis. While some of the proposals may be resource-intensive, requiring an expansion of testing efforts, many of the proposals require only a refocusing of the current effort levels.

Since regulators rely on empirical evidence in formulating consumer financial regulation, the appropriate focus of scholarly debate should be the nature of that evidence. As quantitative testing becomes the standard for good regulation, there is a greater need to articulate the mechanisms by which they assist consumers, as well as placing greater focus on the content and design of testing. By accurately aligning the testing methodology with the purpose of financial disclosures, as suggested in this article, regulators will then be able to adopt better-informed regulation that benefits consumers.

## II. CONSUMER FINANCIAL PROTECTION AND THE PROMINENCE OF DISCLOSURE

Consumer finance, where financial decision-making meets human psychology, covers many areas of daily life in which people are required to make complex choices. For example, the move from Defined Benefit pension plans to Defined Contribution saving plans requires consumers to estimate their future financial needs as well as the current amount they need to save to achieve their future financial goals. In addition, many savings plans require consumers to decide how to invest the assets that are saved. Consumers seeking credit have many more options than in the past, both in terms of types of credit issuers and the terms of credit. Many types of consumer loans have become far more complex, an issue that received a considerable amount of attention following

the global financial crisis.<sup>15</sup>

There is often a significant gap between the knowledge and planning abilities required for informed financial decisions, and those that people with limited time and financial literacy can reasonably be expected to possess. The existence of these market imperfections is largely recognized, so current debates focus on the extent to which they are prevalent and the appropriate regulatory response.

### A. *The Goals of Consumer Financial Regulation*

Defining the precise goal of consumer financial protection can be challenging given the departure from classic justifications for financial regulation in general. Traditional justifications for financial regulation relate to classic market failures that would cause financial market imperfections. Some of the traditional financial regulation justifications can also apply to consumer financial markets, such as the existence of information asymmetries in insurance markets and high searching costs of credit terms.<sup>16</sup> However, the classic framework seems insufficient in capturing the challenges facing unsophisticated consumers, especially when the mere *availability* of information itself is not the real source of concern for many consumers.<sup>17</sup>

If consumers with limited financial literacy face cognitive limitations and biases, what should regulation aim to achieve? If the goal is to improve consumer financial outcomes,<sup>18</sup> the efforts to increase financial literacy and awareness of the terms of products and services would be the means to achieving that end. Even if the end goal is to improve consumer welfare,<sup>19</sup> it may be that this is

---

<sup>15</sup> See, e.g., Willis, *supra* note 2, at 718.

<sup>16</sup> See John Campbell, *et al.*, *Consumer Financial Products*, 25 J. ECON. PERSP. 91 (2011).

<sup>17</sup> Some have articulated this type of regulation as regulating “internalities.” See, e.g., Hunt Allcott & Cass Sunstein, *Regulating Internalities*, J. POL’Y ANALYSIS & MGMT. (forthcoming 2015).

<sup>18</sup> This article does not deal directly with the question of what a “good” decision is. Rather, it assumes that a good decision is one that increases a person’s welfare. While there are many aspects of life in which there is significant uncertainty as to what increases one’s welfare, in the context of consumer finance, in which one’s objective is primarily pecuniary gain, these questions are less important. This approach is consistent with the broader context of practice and literature in consumer finance.

<sup>19</sup> There is a sense in which people making decisions increase their welfare.

best achieved through improving consumer decision-making rather than having regulators decide for people. Another possibility is that the concern is about lack of consumer agency caused by impaired decision-making. According to this view, regulation aims to empower consumers through making their decisions informed, regardless of whether consumer decisions actually improve their financial outcomes. If this were the goal of regulation, we would test the success of regulation using criteria such as the level of understanding of consumers or their subjective feeling of empowerment. The evaluation of the methodology of regulatory testing requires a clear understanding of how regulators view the goal of consumer financial regulation and the problem to which regulation needs to respond.

However, financial regulators have not been clear regarding the goal of consumer financial protection. In the past this was partially due to the fragmentation of consumer financial regulation, whereby consumer financial regulation often fell under the authority of the regulator of a particular industry.<sup>20</sup> These regulators adhered to traditional justifications for financial regulation, presuming a generally functioning financial market with rational players in which regulation is needed only for specific failures.<sup>21</sup>

---

For a full discussion on this topic, see CASS SUNSTEIN, *WHY NUDGE?: THE POLITICS OF LIBERTARIAN PATERNALISM* (2015).

<sup>20</sup> For a more comprehensive analysis of the structure of financial regulation in various European countries, see Eddy Wymeersch, *The Structure of Financial Supervision in Europe: About Single Financial Supervisors, Twin Peaks and Multiple Financial Supervisors*, 8 EUR. BUS. ORG. L. REV. 237 (2007). Some countries, like the UK, followed an integrated model of financial regulation in which the Financial Services Authority oversaw all financial regulation. In the US, regulation followed a fragmented model of regulation, primarily along institutional lines. Despite these differences, both models did not include a separation between prudential regulators and conduct regulators. See, e.g., MICHAEL TAYLOR, "TWIN PEAKS": A REGULATORY STRUCTURE FOR THE NEW CENTURY (1995). On the fragmentation of consumer financial protection, see Oren Bar-Gill & Elizabeth Warren, *Making Credit Safer*, 157 PENN. L. REV. 1 (2008).

<sup>21</sup> Traditional justifications for financial regulation can be more nuanced than presented here. For example, in the context of securities regulation Gilson and Kraakman recognize the existence of unsophisticated investors, but argue that the existence of informed investors provides protection of unsophisticated investors. See Ronald Gilson and Reinier Kraakman, *The Mechanisms of Market Efficiency*, 70 VA. L. REV. 549, 569 (1984). A more accurate characterization

The mission statements and mandates of consumer financial protection regulators since the financial crisis do little to resolve the uncertainties. According to the Consumer Financial Protection Bureau (“CFPB”),<sup>22</sup> its “mission is to make markets for consumer financial products and services work for Americans.”<sup>23</sup> While this may imply a focus on consumer welfare, the CFPB continues to explain that “[a]bove all, this means ensuring that consumers get the information they need to make the financial decisions *they believe* are best for themselves and their families,”<sup>24</sup> thereby emphasizing the subjective goal of empowering consumers. Does the CFPB hold that a decision one believes is best in fact increases one’s welfare? Or do they merely believe they should be concerned with whether people deem their decision as best, regardless of the actual welfare outcome? Despite the importance of the distinction, the CFPB fails to recognize the difference between the welfare outcomes and empowering decision-making, or even discuss the way in which they interact.<sup>25</sup> Mission statements are unlikely to act as a daily guide to the work

---

of traditional justifications of financial regulation is that the existence of unsophisticated investors was not perceived as fundamentally undermining the functioning of an efficient market. This too has been challenged in recent years. *See, e.g.*, Stephen M. Bainbridge, *Mandatory Disclosure: A Behavioral Analysis*, 68 U. CIN. L. REV. 1023 (2000); Stephen Choi & A. C. Pritchard, *Behavioral Economics and the SEC*, 56 STAN. L. REV. 1 (2003); Troy A. Paredes, *Blinded by the Light: Information Overload and Its Consequences for Securities Regulation*, 81 WASH. U. L.Q. 417 (2003).

<sup>22</sup> The CFPB was established under Title X of the Dodd-Frank Wall Street Reform and Consumer Protection Act as a centralized body for consumer protection, concerns that were often previously dealt with by several separate agencies. The powers of the CFPB include the authority to administer, enforce, and implement federal consumer financial laws. *See* 12 U.S.C. § 5511 (2010).

<sup>23</sup> The Bureau, CONSUMER FIN. PROT. BUREAU, <http://www.consumerfinance.gov/the-bureau/> (last visited Sept. 30, 2015).

<sup>24</sup> *Id.* (emphasis added).

<sup>25</sup> The CFPB’s 2013 Strategic Plan is also vague regarding these distinctions. *See* Strategic Plan FY 2013-FY 2017, CONSUMER FIN. PROT. BUREAU (2013), available at <http://files.consumerfinance.gov/f/strategic-plan.pdf> [hereinafter STRATEGIC PLAN]. Goal 1, to prevent financial harm to consumers while promoting good practices that benefit them, and Goal 2, to empower consumers to live better financial lives, focus on providing consumers with information and preventing abuse. “Preventing financial harm” seems to be primarily concerned with the welfare result for consumers, while the “outcomes” the CFPB defines as its goals focus mainly on providing consumers with information. Despite the importance of the distinction, the CFPB fails to recognize the difference between the welfare outcomes and decision-making aids or to discuss the way in which they interact.

of the CFPB; however, they do provide insight into the agency's strategic vision, which is used to define its tasks.

The distinction between the ends and means of regulation is recognized in the goals of the recently established Financial Conduct Authority (FCA) in the United Kingdom, responsible for regulating financial services for the protection of consumers. One of the goals of the FCA is to "promote good outcomes for consumers,"<sup>26</sup> where "[a] further important way to promote good outcomes for retail consumers is to equip them with information so that they can avoid risks and protect themselves,"<sup>27</sup> thereby recognizing improving decision-making as the means of achieving better financial outcomes.

A slightly different view is presented in the Financial Stability Board (FSB) survey of Consumer Credit Protection, which states that "[c]onsumer protection is not about protecting consumers from bad decisions but about enabling consumers to make informed decisions in a marketplace free of deception and abuse."<sup>28</sup> The FSB does not, however, explain how an uninformed decision differs from a bad decision.

Although I have only discussed how the FCA and the CFPB perceive the role of consumer financial regulation, similar vagueness about this role exists in other countries.<sup>29</sup> In general, regulators assume that providing consumers with information will lead them to make better decisions that will improve financial outcomes, a position not properly articulated or justified. This,

---

<sup>26</sup> The definition of the goals of the FCA can be found in a document of the Financial Services Authority, the centralized financial regulator in the UK dissolved in 2013, which sets out the goals of the newly created FCA. See *The Financial Conduct Authority: Approach to Regulation*, FIN. SERV. AUTH. (2011), available at [http://www.fsa.gov.uk/pubs/events/fca\\_approach.pdf](http://www.fsa.gov.uk/pubs/events/fca_approach.pdf) [hereinafter FSA DOCUMENT]. For a critical analysis of the abolition of the Financial Services Authority and creation of the Financial Conduct Authority, see Eilis Ferran, *The Break-up of the Financial Services Authority*, 31 OXFORD J. OF LEGAL STUD. 455 (2011).

<sup>27</sup> FSA DOCUMENT, *supra* note 26, at 17.

<sup>28</sup> *Consumer Finance Protection With Particular Focus on Credit*, FIN. STABILITY BD. (2011), available at [http://www.financialstabilityboard.org/wp-content/uploads/r\\_111026a.pdf](http://www.financialstabilityboard.org/wp-content/uploads/r_111026a.pdf) [hereinafter FSB REPORT].

<sup>29</sup> For further discussion of consumer protection in various countries, see *id.* See also ORG. FOR ECON. CO-OPERATION AND DEV., G20 HIGH LEVEL PRINCIPLES OF FINANCIAL CONSUMER PROTECTION (Oct. 2011), available at <http://www.oecd.org/daf/fin/financial-markets/48892010.pdf>.

however, is not true if people ignore the information disclosed,<sup>30</sup> do not understand the disclosure or do not use the information to make better decisions.<sup>31</sup> The assumed link between disclosure and improved outcomes is largely what causes misguided regulatory testing of proposed disclosures.

### B. The Role of Disclosure

There are various types of regulatory responses aimed at protecting financial consumers. Regulators may choose to regulate financial products directly. For example, regulators may ban certain financial products<sup>32</sup> while restricting access to others.<sup>33</sup> Another type of regulatory intervention is creating structural requirements that are meant to protect consumers. Rather than regulating the financial product, this regulatory strategy aims at regulating the financial sector players.<sup>34</sup>

---

<sup>30</sup> Carlos Jensen, Colin Potts & Christine Jensen, *Privacy Practices and Internet Users: Self-reports Versus Observed Behavior*, 203 INT'L J. HUMAN-COMPUTER STUD. 63 (2005).

<sup>31</sup> There is a large amount of literature on how people do not save, even when prompted to by their employer and when they believe they should be saving. See, e.g., Gabriel Carroll, et al., *Optimal Defaults and Active Decisions*, 124 Q. J. ECON. 1639 (2009); James Choi, David Laibson, & Brigitte Madrian, *\$100 Bills on the Sidewalk: Violations of No-Arbitrage in 401(k) Accounts*, 93 REV. OF ECON. & STAT. 748 (2011).

<sup>32</sup> Such regulation is often perceived as more legitimate when a product is deceptive or unfair. For example, in January 2013, the CFPB issued a rule banning balloon payments and repayment penalties on high-cost mortgages. See High-Cost Mortgage and Homeownership Counseling Amendments to the Truth in Lending Act (Regulation Z) and Homeownership Counseling Amendments to the Real Estate Settlement Procedures Act (Regulation X), 79 Fed. Reg. 6855 (Jan. 10, 2013), available at <http://www.consumerfinance.gov/regulations/high-cost-mortgage-and-homeownership-counseling-amendments-to-regulation-z-and-homeownership-counseling-amendments-to-regulation-x/>. Other examples include restrictions on assets held by pension funds and restrictions on retain investment funds.

<sup>33</sup> For example, according to Regulation D only investors who fulfill certain criteria can invest in hedge funds. See 17 CFR 230.506. Lighter versions of regulation of financial products include the use of defaults from which financial institutions may opt out of exposing them to tightened regulation.

<sup>34</sup> A classic example is the creation of fiduciary duties meant to align incentives of firms with those of the consumers. Another type of structural imposition is the requirement of gatekeepers, such as independent directors on mutual fund boards.

## 1. The Prominence of Disclosure as a Regulatory Tool

Although consumer financial regulators have employed several of the above strategies in various contexts, in many jurisdictions regulation relies primarily on empowering consumer decision-making. Various governmental and private efforts, both through formal education channels, such as seminars in schools, as well as informal education, have tried to increase consumer literacy. Results have been mixed.<sup>35</sup>

The most common strategy is the regulation of information given to consumers. This type of regulation can take many forms, from requiring information to be disclosed,<sup>36</sup> to prohibiting information that may be misleading.<sup>37</sup> Regulators often choose themselves to provide or sponsor information, allowing consumers to compare information in a credible and neutral setting.<sup>38</sup>

Regulatory efforts focused on enhancing consumer decision-making disclosure are perceived as autonomy-promoting as well as preserving the appearance that consumer finance is

---

<sup>35</sup> For a survey of consumer education efforts see Annamaria Lusardi & Olivia S. Mitchell, *The Economic Importance of Financial Literacy: Theory and Evidence*, 52 J. ECON. LITERATURE 5, 29-33 (2014). See also Justine S. Hastings, Brigitte C. Madrian & William L. Skimmyhorn, *Financial Literacy, Financial Education and Economic Outcomes*, 5 ANN. REV. ECON. 347 (2013).

<sup>36</sup> Pre-contractual requirements may oblige that consumers receive certain disclosures at certain times before entering a contract, like the TILA and RESPA disclosures discussed in Section III. Other requirements include post-contractual mandated disclosures and ongoing reporting, often required of financial institutions. In many cases the form, as well as content of disclosure is closely regulated. Other types of disclosure requirements include the disclosure of conflicts of interest.

<sup>37</sup> This could be done through the regulation of marketing and advertising, such as the limitations placed on mutual fund advertising.

<sup>38</sup> See, e.g. *Fund Analyzer*, FIN. INDUS. REGULATORY AUTH., <http://apps.finra.org/fundalyzer/1/fa.aspx> (last visited Sept. 30, 2015); *Paying for College: Compare Financial Aid*, CONSUMER FIN. PROT. BUREAU, <http://www.consumerfinance.gov/paying-for-college/compare-financial-aid-and-college-cost/> (last visited Sept. 30, 2015) (CFPB's tool to compare college costs and financial aid offers). For more details on the various types of consumer financial protection, see FSB REPORT, *supra* note 28, at 12-18 (survey of various country). On regulation in middle income countries, see Susan Rutledge, *Consumer Protection and Financial Literacy*, (World Bank Pol'y Res., Working Paper No. 5326, 2010), available at [http://sitere-sources.worldbank.org/EXTFINANCIALSECTOR/Resources/Consumer\\_Pr otection\\_and\\_Fin\\_LiteracyWPS5326.pdf](http://sitere-sources.worldbank.org/EXTFINANCIALSECTOR/Resources/Consumer_Pr otection_and_Fin_LiteracyWPS5326.pdf).

primarily based on fair market interactions.<sup>39</sup> In a context in which financial engagements are an exercise of free choice, it seems natural that disclosure becomes a central part of regulation.<sup>40</sup> First, disclosure is appealing since it is perceived as a light form of regulation, as it does not alter the market significantly. In fact, supporters of disclosure argue that informing consumers enhances the functioning of the free-market. Ben-Shahar and Schneider argue that disclosure is alluring because it appeals to fundamental American ideology of autonomy and that people “are entitled as a matter of moral right and of practical policy to make the decisions that shape their lives.”<sup>41</sup> Second, disclosure is often more politically practical,<sup>42</sup> and, as is argued in more detail below, it avoids solving complex regulatory questions, such as what decisions actually improve consumer welfare. The centrality of disclosure to consumer financial regulation can also be explained by the importance of disclosure to financial regulation and investor protection in general.<sup>43</sup>

---

<sup>39</sup> FSB REPORT, *supra* note 28, at 15. According to the FSB, “Most jurisdictions are working to enable consumers to make better informed consumer credit decisions in a safer marketplace. They are strengthening consumer education and consumer protection, and disclosure requirements for both basic and complex products.”

<sup>40</sup> For an overview of the philosophical underpinnings of disclosure, see Matthew Edwards, *The Virtue of Mandatory Disclosure*, 28 NOTRE DAME J.L. ETHICS & PUB. POL’Y 47 (2014).

<sup>41</sup> MORE THAN YOU WANTED TO KNOW, *supra* note 7, at 5.

<sup>42</sup> Omri Ben Shachar & Carl Schneider, *The Futility of Cost Benefit Analysis in Financial Disclosure Regulation* (U. Mich. L. Sch. L. & Econ., Working Paper No. 100, 2014) [hereinafter *Futility of CBA*].

<sup>43</sup> The Securities Act of 1933 and the creation of the SEC with the passing of the Securities Act of 1934 were premised on a market failure of insufficient information in the market. See, e.g., Frank H. Easterbrook & Daniel R. Fischel, *Mandatory Disclosure and Protection of Investors*, 70 VA. L. REV. 669 (1984); John C. Coffee, *Market Failure and the Economic Case for Mandatory Disclosure System*, 70 VA. L. REV. 717 (1984). It is unclear to what degree such justifications for disclosure should hold for other consumer financial transactions, particularly when they involve privately negotiated agreements. See, e.g., Howell Jackson, *Regulation in Multisector Financial Services Industry: An Exploratory Essay*, 77 WASH. U. L.Q. 319 (1999). In addition, many investment vehicles, such as mutual funds, are retail vehicles in which sophisticated investors do not participate. See John Coates & Glenn Hubbard, *Competition in the Mutual Fund Industry: Evidence and Implications for Policy*, 33 J. CORP. L. 151 (2007) (discussing the degree to which sophisticated investors may protect the unsophisticated). An additional concern is that in markets in which knowledgeable and non-knowledgeable consumers can be distinguished. For example, the sub-prime market being for poorer and less sophisticated households. See Willis, *supra* note 2, at 762.



## 2. Why Disclosure May Benefit Consumers

A prerequisite for regulators to adequately test their proposed disclosure is an explicit theory of the way in which disclosures benefit consumers. Similar to the existing vagueness on the role of consumer protection, regulators have not been explicit as to how disclosure assists consumers.

This ambiguity is apparent from the early days of consumer financial regulation. Edward Rubin discusses the adoption of the Truth in Lending Act (“TILA”) in 1968, the first consumer financial protection legislation. Edward Rubin asks with respect to TILA: “what was the goal of legislation, consumer protection or disclosure?”<sup>44</sup> Rubin argues that the way in which the Act was adopted assumes the solution of disclosure “precludes an explicit inquiry into the bedrock goals of the legislative effort.”<sup>45</sup> Rubin argues that TILA involved a conflation between instrumentalities and goals, also apparent in today’s regulation.

### a. Consumer Responses to Disclosure

The claim that disclosure improves consumer welfare, based on consumer responses to disclosures, relies on establishing a number of links. For example, there is the link between receiving financial disclosures and reading and then understanding their content. This is the connection between exposure to disclosure and comprehension. Second, there must be a connection between understanding the information and making an informed decision. In other words, comprehension needs to lead to improved decision-making.<sup>46</sup>

---

<sup>44</sup> Edward Rubin, *Legislative Methodology: Some Lessons from the Truth-in-Lending Act*, 80 GEO. L.J. 233, 283 (1991).

<sup>45</sup> *Id.* at 285.

<sup>46</sup> This list is not exhaustive. For example, there also must be a connection between whether a decision an individual believes is best for them, and whether it truly is a good decision, if regulation is concerned with consumer welfare. Also, there must be a link between a good financial decision (at the time of the decision) and financial outcomes that may materialize far in the future and are shaped by economic realities unknown at the time of the decision. Ben Shachar and Schneider discuss further links that are required for disclosure to succeed including that the lawmakers impose the right mandates and that disclosers obey the disclosure requirement. See MORE THAN YOU WANTED TO KNOW, *supra* note 7, at 34.

However, there is a significant amount of evidence that the first of these links fails, in that people do not understand financial disclosures, or even attempt to read them.<sup>47</sup> This has led many authors to doubt the effectiveness of disclosures in helping consumers. Other authors, however, have advanced a more nuanced perspective, looking to define when disclosures are more likely to succeed.<sup>48</sup> In a recent article by Loewenstein, Sunstein, and Golman,<sup>49</sup> the authors state that “important and reasonable questions have been raised about the efficacy of disclosure requirements,”<sup>50</sup> and based on research in the field of behavioral economics, they suggest a few guidelines as to how disclosure may be improved. The ways to improve disclosure include information simplification, standardization, social comparison information, and vividness.

Regulatory testing can be seen as the response to much of this literature, as it recognizes the possible pitfalls of disclosure and the difficulty in determining its comprehensibility in the abstract. As the literature on the ways in which disclosure might succeed or fail continues to grow, it is clear that the terms of this debate are insufficiently defined. Most importantly, it is unclear what the criterion is for the success or failure of a disclosure. For example, is it necessary that the document in its entirety be understood? Do all readers have to understand disclosures for them to be a “success”?<sup>51</sup> This is particularly apparent in the examples discussed below, where the EU and CFPB adopt disclosures based partially on low comprehension results without providing any relevant benchmark.

---

<sup>47</sup> *Id.* at 33-54; see further discussion *infra*, Part II.1.

<sup>48</sup> See ARCHON FUNG, MARY GRAHAM, & DAVID WEIL, *FULL DISCLOSURE: THE PERILS AND PROMISE OF TRANSPARENCY* (2008) [hereinafter *FULL DISCLOSURE*]; Lauren Willis, *The Consumer Financial Protection Bureau and the Quest for Consumer Comprehension* (May 14, 2015), available at [http://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=2606373](http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2606373); MORE THAN YOU WANTED TO KNOW, *supra* note 7, at 121-137; Bubb, *supra* note 7, at 107.

<sup>49</sup> See George Loewenstein, Cass Sunstein & Russell Golman, *Disclosure: Psychology Changes Everything*, 6 ANN. REV. ECON. 391 (2014).

<sup>50</sup> *Id.* at 398.

<sup>51</sup> Richard Craswell has suggested a distinction between static disclosure, relating to the specific attributes of a product, and dynamic disclosure, relating to the directional impact of attributes, giving sellers incentives to improve their products. Which of these two types of disclosure should be the focus of a comprehensible disclosure? See generally Craswell, *supra* note 7.

### b. Supply Side Effects of Disclosure

Another way in which disclosure can increase consumer financial protection is through the effect of mandatory disclosures on those that supply them. Requiring loan originators to disclose their terms and mutual funds to disclose their fees may make financial institutions alter their provisions to avoid adverse consequences of harmful terms being public,<sup>52</sup> thereby not relying directly on consumer comprehension. Classic examples of disclosure systems that have altered the behavior of disclosers, while it is highly questionable whether consumers notice this information when making a purchase, include food nutritional labels<sup>53</sup> and restaurant hygiene disclosures.<sup>54</sup>

While this theory provides an explanation for how disclosure can be beneficial even when many or even most consumers do not understand the disclosure, it nevertheless does rely on a group of users being sensitive to the information being disclosed. For example, it relies on the existence of an informed minority who understands the disclosure<sup>55</sup> and can protect uninformed consumers or consumers partially understanding the disclosure.<sup>56</sup> The theory could also rely on an intermediary who is able to interpret the information in a way that can create a

---

<sup>52</sup> See David Weil, *The Benefits and Costs of Transparency: A Model of Disclosure Based Regulation* (2002).

<sup>53</sup> See FULL DISCLOSURE 2007, *supra* note 48, at Chapter 4. For a nuanced analysis of the cases in which disclosure of nutritional value led to improved nutrition of products. See Christine Moorman, Rosellina Ferraro & Joel Huber, *Unintended Nutrition Consequences: Firm Responses to the Nutrition Labeling and Education Act*, 31 *MARKETING SCI.* 717 (2012).

<sup>54</sup> See FULL DISCLOSURE, *supra* note 48, at Chapter 4. For a skeptical account of the impact of restaurant hygiene grading. See Daniel Ho, *Fudging the Nudge: Information Disclosure and Restaurant Grading*, 122 *YALE L.J.* 574 (2012).

<sup>55</sup> See Alan Schwartz & Louis Wilde, *Intervening in Markets on the Basis of Imperfect Information: A Legal and Economic Analysis*, 127 *U. PA. L. REV.* 630 (1979). In general, the more private consumer interactions are the less one can rely on an informed minority. See, e.g., Florencia Marotta-Wurgler, *Will Increased Disclosure Help? Evaluating Recommendations of the ALI's "Principles of the Law of Software Contracts"*, 78 *U. CHI. L. REV.* 165 (2011).

<sup>56</sup> See Craswell, *supra* note 7, at 342 for example Craswell's characterization of "dynamic disclosures" in which sellers alter their products because consumers respond to the levels of an attribute even when they fall short of being fully informed.

consumer response.<sup>57</sup> However, it is questionable whether such mechanisms exist for many consumer financial products that deal with private contracts, and not uniform products, or products that are primarily for unsophisticated consumers.

To summarize this section, since regulatory testing is meant to confirm whether the proposed rule will achieve its goal, laying out the goals of regulation is essential in assessing testing efforts. While financial regulators ultimately wish to improve consumer welfare, they focus on decision-making and providing information to consumers, confusing the ends and means of regulation.

The current philosophy of consumer financial regulation focuses on enhancing consumer autonomy, in which disclosure plays a dominant role. As financial disclosures have increasingly been subject to criticism, their effectiveness needs to be established by empirical testing. This point is further developed in the next section.

## II. WHY TEST FINANCIAL DISCLOSURES?

Rigorous testing of proposed regulation, particularly quantitative studies, is a fairly recent development. In the context of financial disclosures, The Board of Governors of the Federal Reserve System (the “Federal Reserve Board”) has sponsored a number of quantitative studies since 2007.<sup>58</sup> Regulatory testing of proposed financial disclosures has become more prevalent, particularly after the financial crisis, for two main reasons. First, changes that have taken place regarding perceptions of the average consumer based on evidence that people do not understand or read disclosures, as well theoretical changes in economic theory, have increased disclosure skepticism necessitating verification of its effectiveness.

The second explanation is the increased rigor of cost-benefit analysis and data-driven analysis that requires regulators to present concrete evidence as to the benefit of proposed regulation. In the context of disclosure, this has led to quantitative testing of financial disclosures.

---

<sup>57</sup> See FULL DISCLOSURE, *supra* note 48, at 61.

<sup>58</sup> See FEDERAL RESERVE BULLETIN 2011, *supra* note 12.

*A. Changes in Theory of Consumer Behavior*

When mandated financial disclosures became an important component of financial regulation in the 1930s, little attention was given to the importance of the presentation and format of the information. Instead, the focus was on the mere release of information into the public domain.<sup>59</sup> As the use of financial disclosures increased, there was greater concern regarding the ability of consumers to understand many pages of technical and legalistic language.<sup>60</sup> Moreover, disclosure was increasingly used in the context of individual consumer financial contracts, like insurance and mortgage contracts, in addition to capital markets. While the government can reduce information costs for consumers, some costs may persist, particularly when shopping for credit, reducing the impact of disclosure.<sup>61</sup>

However, evidence of actual impact of disclosures, as well as broader research in the field of behavioral economics, undermined the theory behind disclosure whereby as long as the effort is worth consumers' while, they will invest in understanding disclosures.<sup>62</sup> Research showed that people had difficulty understanding disclosures in a way not limited to lack of expertise, and that people did not read disclosures, even when making significant financial decisions or when disclosures were not significantly long.<sup>63</sup> Furthermore, it was shown that people do not

---

<sup>59</sup> See, e.g. LOUIS BRANDEIS, OTHER PEOPLE'S MONEY AND HOW THE BANKERS USE IT (2009). Brandeis says: "Publicity is justly commended as a remedy for social and industrial diseases." *Id.* at 62. The economic formalization of the benefit of financial disclosures began to be articulated in the late 70s and early 80s, while also focusing on the connection between publicly available information and market efficiency. See, e.g., Eugene Fama, *Efficient Capital Markets: A Review of Theory and Empirical Work*, 25 J. FIN. 383 (1970).

<sup>60</sup> For a further account of the increased use of disclosures and emphasis on transparency from the 1980s, see Archon Fung, et al., *The Political Economy of Transparency: What makes disclosure policies effective?*, HARVARD U. JOHN F. KENNEDY SCH. OF GOV'T, 8 (2004).

<sup>61</sup> For such an analysis in the context of TILA and RESPA disclosures, see Jinkook Lee & Jeanne Hogarth, *Consumer Information Search for Home Mortgages: Who, What, How Much, and What Else?*, 9 FIN. SERVICES REV. 277 (2000).

<sup>62</sup> See *supra*, Part I.B.2.a.

<sup>63</sup> See ANJAN V. THAKORE, JESS C. BELTZ, & JO ANN S. BAREFOOT, COMMON GROUND: INCREASING CONSUMER BENEFITS AND REDUCING REGULATORY COSTS IN BANKING (1993); Willis, *supra* note 2, at 766; Rubin, *supra* note 44, at 235; Matthew Edwards, *Empirical and Behavioral Critiques*

understand basic financial concepts, and they overlook information that is not easily understood.<sup>64</sup> Indications that disclosure may be harder to understand than thought at the time mandated went beyond financial disclosures to other disclosures such as nutrition labels<sup>65</sup> and medical informed consent.<sup>66</sup>

Regulators mandating financial disclosures have been slow in addressing the full impact of this evidence. It is clear that regulators see the need to engage with the users of disclosure when developing disclosure requirements; however, they often fall short of truly questioning their effectiveness. Prior to quantitative testing, the basis for the adoption of disclosure requirements was either the regulator's theory as to what information would be important and useful for consumers or even the stated preferences of consumers or the regulated industry.<sup>67</sup> Testing efforts by the Federal Reserve Board in the 1990s mostly included focus groups in which consumers were asked to comment on proposed disclosures.<sup>68</sup> This form of consumer engagement, through eliciting their preferences for the disclosure document, has serious limitations as consumers are often confused themselves about what information is important in comparing financial offers.

---

*of Mandatory Disclosure: Socio-Economics and the Quest for Truth in Lending*, 14 CORNELL J.L. & PUB. POL'Y 199, 229 (2005).

<sup>64</sup> See Rubin, *supra* note 44, at 236; William Whitford, *The Functions of Disclosure Regulations in Consumer Transactions*, 2 WISC. L. REV. 400 (1973). For a comprehensive survey of the various challenges facing consumers, see Angelo Capuano & Ian Ramsay, *What Causes Suboptimal Financial Behaviour? An Exploration of Financial Literacy, Social Influences and Behavioural Economics*, CENTRE FOR CORP. L. AND SEC. REG., U. MELBOURNE, (2011).

<sup>65</sup> Alan Levy, Sara Fein & Raymond Schucker, *Performance Characteristics of Seven Nutrition Label Formats*, 15 J. OF PUB. POL'Y & MARKETING 1 (1996).

<sup>66</sup> Barrie Cassileth, *Informed Consent—Why are Its Goals Imperfectly Realized?*, 302 N. ENGL. J. MED. 896 (1980). See MORE THAN YOU WANTED TO KNOW, *supra* note 7, at 43.

<sup>67</sup> See Rubin, *supra* note 44, at 300 (“[T]heir efforts to determine the statute’s actual effects were largely limited to speculation.”). See also SEC Enhanced Disclosure and New Prospectus Delivery Option For Registered Open-End Management Investment Companies, 74 Fed. Reg. 4546 (Jan. 26, 2009), <http://www.sec.gov/rules/final/2009/33-8998fr.pdf> (describing the steps taken before the SEC adopted the Summary Prospectus for mutual funds) [hereinafter SEC SUMMARY PROSPECTUS]. Prior to the adoption of the SEC SUMMARY PROSPECTUS, the SEC hosted a roundtable with investor group representatives and held investor focus group testing.

<sup>68</sup> FEDERAL RESERVE BULLETIN 2011, *supra* note 12, at 4.

In some areas the recognition that disclosures often failed to assist people in making financial decisions led regulators to adopt new disclosures, meant to address behavioral concerns, only to later reject them for being ineffective.<sup>69</sup> One possible reason for the repeated shortcomings of updated disclosures is that they rely on abstract notions of how disclosure formats help consumers. Thus, the effect of disclosure could be determined only through the testing of disclosure documents prior to their adoption.<sup>70</sup>

Recent years have marked a change as to how financial regulators adopt disclosures. In 2004, the Federal Reserve Board created a more rigorous testing program to inform the development of financial disclosures, adding elements beyond focus groups, such as usability testing and quantitative validation surveys. This is clearly reflected in the Federal Reserve Board's summary of consumer testing of disclosures from 2004 to 2011, emphasizing the learning process of disclosure testing.<sup>71</sup> In a staff report of the Federal Trade Commission from 2007 that tested the effectiveness of mortgage disclosures, the authors state: "testing is

---

<sup>69</sup> For example, as will be discussed in Part III, the European Commission replaced the simplified prospectus, meant to provide a short standardized document in simple English, with the Key Investor Information Document, meant to provide a similar function.

Another example is the 1968 Truth in Lending Act mandated disclosures, which have been reviewed and changed several times. See Part III.B. The process through which TILA disclosures have been amended within short periods of time, despite the persistence of the overall goals of TILA, raises questions as to how TILA disclosures are adopted in the first place. While many changes relate to the increased complexity of loan terms, for example the proliferation of adjustable rate mortgages, documents are often deemed non-user friendly sometimes within only years after they are implemented. For an in depth analysis of the situation of mortgage consumers in 2007. See JAMES LACKO & JANIS PAPPALARDO, FED. TRADE COMM'N, IMPROVING CONSUMER MORTGAGE DISCLOSURES: AN EMPIRICAL ASSESSMENT OF CURRENT PROTOTYPE DISCLOSURE FORMS (2007) [hereinafter FTC REPORT 2007].

<sup>70</sup> Others have argued that this is partially because the Federal Reserve Board lacks expertise and motivation to focus on consumer financial protection. See, e.g., Jeff Sovern, *Fixing Consumer Protection Laws So Borrowers Understand Their Payment Obligations*, 48 J. OF CONSUMER AFF. 17 (2014). See also FTC REPORT 2007, *supra* note 69, at 7 ("Despite the importance of consumer mortgage decisions, decades of experience with federal mortgage disclosures, a general perception that these disclosures are inadequate and confusing, and repeated calls for improvement, little empirical evidence exists to document the effect of the current disclosures on consumer understanding of mortgage costs, consumer mortgage shopping, or consumer mortgage choice.").

<sup>71</sup> FEDERAL RESERVE BULLETIN 2011, *supra* note 12.

essential to ensure that the disclosures effectively convey the desired information to consumers,<sup>72</sup> and that “[w]ithout testing, it is difficult to know what information is conveyed to consumers.”<sup>73</sup>

The creation of the CFPB may also mark the increased incorporation of behavioral economics research into regulatory policy.<sup>74</sup> Furthermore, the CFPB’s Strategic Plan includes as one of its goals the conduction of “qualitative and quantitative research to deepen understanding of consumer decision making.”<sup>75</sup> The next section discusses the CFPB’s most extensive consumer study to date, testing mortgage disclosures, reflecting the Strategic Plans’ goal of research-driven, evidence-based perspectives on consumer behavior.

Although the creation of the CFPB and the Board of Governors of the Federal Reserve System (FRB) testing marks recognition of the importance of consumer testing, current financial regulation has not fully abandoned abstract notions of reasonable consumer behavior. The basic regulatory structure has remained intact – consumer financial protection relies heavily on consumer decision-making, yet testing overlooks decision-making and focuses on comprehension of disclosures instead. As discussed in Part IV, it is theories of consumer rationality that assume the connection between comprehension and decision-making, and therefore regulators that rely on this connection have not fully incorporated behavioral economics research.

### *B. Increased Stringency of Regulatory Impact Assessments*

The increased use of consumer testing to validate financial disclosures can also be explained by the more general trend of evidence-based policy and regulation.

Cost-benefit analysis (“CBA”) is a fundamental part of regulation in the United States.<sup>76</sup> Although CBA has many

---

<sup>72</sup> FTC REPORT 2007, *supra* note 69, at 13.

<sup>73</sup> *Id.* at 128.

<sup>74</sup> See Dee Prigden, *Sea Changes in Consumer Financial Protection: Stronger Agency and Stronger Laws*, 13 WYO. L. REV. 405, 406 (2013) (claiming that the birth of the CFPB is a shift “to a system of regulation-based on the more realistic view of consumer decision-making as revealed by behavioral economics”).

<sup>75</sup> STRATEGIC PLAN, *supra* note 25, at 27.

<sup>76</sup> It dates back to 1974 when Council on Wage and Price Stability



opponents,<sup>77</sup> it seems clear that CBA is fairly entrenched in the current regulatory regime. Agencies are producing increasingly longer and more sophisticated analyses, which are then reviewed by the Office of Information and Regulatory Affairs (“OIRA”) and sometimes courts, while investing considerable resources in their production.<sup>78</sup> The increased stringency of CBA means that agencies are now required to produce more concrete evidence on the benefits and costs of regulation, as well as a monetized estimate of the benefits and costs.<sup>79</sup>

Even for agencies that did engage in CBA, this analysis was rarely applied in the case of mandated disclosures.<sup>80</sup> Although it has been argued that “informal cost-benefit analysis” constantly takes place under the assumption that disclosure of information has potentially significant benefits while its costs are low,<sup>81</sup> there has been little formal analysis, particularly of alternative regulatory measures or even alternative disclosure options. This is likely to change given the growing skepticism of the effectiveness

---

(“COWPS”) created in 1974 assumed the role of CBA of regulation. Its successor, the Regulatory Analysis Review Group (“RARG”), was set up during the Carter Administration. In 1981 during the Regan Administration, the Office of Information and Regulatory Affairs (“OIRA”) was created within the White House’s Office of Management Budget (“OMB”), and was responsible for regulatory oversight of all major federal agency rules. Although at the time OIRA was considered by many as a mechanism to facilitate deregulation by creating significant hurdles to rule-making, its functioning was continued by the Clinton Administration. Executive Order 12866, issued by President Clinton in 1993, remains the fundamental guidelines as to OIRA’s authority and the principles of rulemaking. In 2011 President Obama issued Executive Order 13565, reaffirming Executive Order 12866 as well as adding additional principles, such as the need to examine existing regulation. For an overview of the history of CBA and its prominence in American regulatory decision-making, see CASS R. SUNSTEIN, *THE COST-BENEFIT STATE: THE FUTURE OF REGULATORY PROTECTION* (2002).

<sup>77</sup> See, e.g., FRANK ACKERMAN ACKERMAN & LISA HEIZERLING, *PRICELESS: ON KNOWING THE PRICE OF EVERYTHING AND THE VALUE OF NOTHING* (2005); PETER DORMAN, *MARKETS AND MORTALITY: ECONOMICS, DANGEROUS WORK, AND THE VALUE OF HUMAN LIFE* (2009).

<sup>78</sup> See Stuart Shapiro, *The Evolution of Cost-Benefit Analysis in U.S. Regulatory Decisionmaking*, in *HANDBOOK ON THE POLITICS OF REGULATION* 385 (David Levi-Faur ed., 2011).

<sup>79</sup> In the context of privacy notices, see Loretta Garrison, et al., *Designing Evidence-based Disclosures: A Case Study of Financial Privacy Notices*, J. CONSUMER AFF. 204 (2012).

<sup>80</sup> *Futility of CBA*, *supra* note 42, at 3.

<sup>81</sup> *Id.* at 4.

of financial disclosures. When the benefits are no longer assumed, they will have to convincingly be established, which will necessitate consumer testing.

The same is likely to be true for independent agencies, primarily responsible for financial disclosures, which are not subject to OIRA's review.<sup>82</sup> For example, the Securities Exchange Commission ("SEC"), responsible for mutual fund oversight including mutual fund disclosure, has not emphasized CBA in its rulemaking in the past.<sup>83</sup> However, in March 2012 the SEC circulated the Guidance on Economic Analysis in SEC Rulemaking,<sup>84</sup> which requires that the benefits and costs of rules be clearly identified and quantified, likely to increase future empirical testing of regulation.<sup>85</sup>

---

<sup>82</sup> They are also not covered by the requirements of Executive Order 12866 and OMB Circular A-4. Although many rules of independent agencies contain an impact analysis, they have been considered less robust than the analyses of agencies reviewed by OIRA. For example the National Securities Market Improvement Act, 15 U.S.C. §77b(b), requires the SEC to consider effects of its regulation, and the Securities Exchange Act of 1934 itself requires the SEC to consider the impact of its regulation on competition. On the lower level of rigor of these analyses see the OMB Annual Report from 2012, which questions the rigor of independent agencies' cost benefit analyses. OFF. MGMT & BUDGET, OFF. INFO. AND REG. AFF. 2013 REPORT TO CONGRESS ON THE BENEFITS AND COSTS OF FEDERAL REGULATIONS AND UNFUNDED MANDATES ON STATE, LOCAL, AND TRIBAL ENTITIES (2013), available at [https://www.whitehouse.gov/sites/default/files/omb/inforeg/2013\\_cb/2013\\_cost\\_benefit\\_report-updated.pdf](https://www.whitehouse.gov/sites/default/files/omb/inforeg/2013_cb/2013_cost_benefit_report-updated.pdf).

<sup>83</sup> This is clear from reviewing the Final Rule of the Summary Prospectus (the example discussed in the introduction to this paper) that contains little analysis on concrete benefits to consumers. See SEC SUMMARY PROSPECTUS, *supra* note 67.

<sup>84</sup> Memorandum from the RSFI & OGC, Secs. & Exch. Comm'n, to Staff of the Rulewriting Divisions and Offices (Mar. 16, 2012), [http://www.sec.gov/divisions/riskfin/rsfi\\_guidance\\_econ\\_analy\\_secrulemaking.pdf](http://www.sec.gov/divisions/riskfin/rsfi_guidance_econ_analy_secrulemaking.pdf). This followed a report of the Government Accountability Office and recent court decisions challenging the SEC's cost-benefit analysis. See *Chamber of Commerce v. SEC*, 412 F.3d 133 (D.C. Cir. 2005) (challenging the SEC rule that boards mutual funds have at least 75% independent directors and an independent chairman); *Am. Equity Inv. Life Ins. Co. v. SEC*, 613 F.3d 166 (D.C. Cir. 2010) (challenging the SEC rule that classified fixed indexed annuities be subject to federal securities regulation); *Bus. Roundtable & U.S. Chamber of Commerce v. SEC*, 647 F.3d 1144 (D.C. Cir. 2011) (challenging a rule regarding companies' proxy materials to shareholders).

<sup>85</sup> See Eric Posner & Glen Weyl, *Benefit-Cost Paradigms in Financial Regulation* (Coase-Sandor Inst. for L. and Econ., Working Paper No. 660, 2013); See also Maeve Carey, CONG. RESEARCH SERV., RL77775, COST-BENEFIT AND

Similarly, the CFPB, also an independent agency,<sup>86</sup> sees evidence-based policy and data-driven regulation as a fundamental part of its Strategic Plan. Moreover, Dodd-Frank, establishing the authority and responsibilities of the CFPB, requires that the CFPB conduct a CBA on its proposed rules.<sup>87</sup> Dodd-Frank also explicitly requires that disclosures mandated by the CFPB be validated through consumer testing.<sup>88</sup>

A similar trend can be seen outside the US. Since 2002, any new initiative of the European Commission requires an Impact Assessment. Such an assessment includes the consideration of economic, social and environmental consequences of new rules. The Commission has published various guidelines on preparing impact assessments<sup>89</sup> and created the Impact Assessment Board in 2006, which advises and scrutinizes impact assessments.<sup>90</sup> This likely indicates the increased use of empirical testing of regulation, which is seen in the example discussed next, the Key Investor Information Document. The basic similarities between the structure of impact assessments and CBA, as well as the oversight of the Impact Assessment Board and OIRA, have led to claims of convergence on the analytic basis for regulation across the Atlantic.<sup>91</sup>

---

OTHER ANALYSIS REQUIREMENTS IN THE RULEMAKING PROCESS (2014), available at <http://fas.org/sgp/crs/misc/R41974.pdf> (quoting the statement of the SEC Chairman from 2011).

<sup>86</sup> The CFPB is also not subject to OIRA's review and Executive Orders 12866 and 13563.

<sup>87</sup> Dodd-Frank Wall Street Reform and Consumer Protection Act (hereinafter as Dodd-Frank), Pub. L. No. 111-203, §1022(b), 124 Stat. 1376 (2010).

<sup>88</sup> Dodd-Frank §1032(b)(3).

<sup>89</sup> *Impact Assessment Guidelines*, SECURITIES EXCH. COMM'N 92 (Jan. 15, 2009), available at [http://ec.europa.eu/governance/impact/commission\\_guidelines/docs/iag\\_2009\\_en.pdf](http://ec.europa.eu/governance/impact/commission_guidelines/docs/iag_2009_en.pdf).

<sup>90</sup> Similarly, with regards to lower level rule-making, the Committee of European Securities Regulators ("CESR") along with the Committee of European Banking Supervisors and the Committee of European Insurance and Occupational Pensions Supervisor, created Impact Assessment Guidelines for level 3 committees emphasizing the need for "evidence-based policy making." *Impact Assessment Guidelines for EU Level 3 Committees* (May 2007), Committee of European Insurance and Occupational Pensions Supervisors, available at [http://www.esma.europa.eu/system/files/07\\_089.pdf](http://www.esma.europa.eu/system/files/07_089.pdf).

<sup>91</sup> Jonathan Wiener & Alberto Alemanno, *Comparing Regulatory Oversight Bodies Across the Atlantic: The Office of Information and Regulatory Affairs in*

These developments explain the increased use of consumer testing of proposed financial disclosures and the expected future occurrences of consumer testing. They demonstrate how a behavioral approach necessitates greater caution in determining the impact of regulation in abstract. The next section discusses two examples of extensive testing efforts. I utilize these examples to discuss how these efforts reflect the need for consumer testing but are not designed to allow testing of what really matters for financial disclosures—namely, that they improve consumer decision-making.

### III. EXAMPLES OF FINANCIAL DISCLOSURE REGULATORY TESTING

In two recent quantitative studies, proposed disclosures were evaluated through comprehension questionnaires. In the next section, I discuss the various issues these testing methodologies raise, namely that comprehension questionnaires fail to measure the relevant outcome, which is the decision made based on the disclosure, and that the setup of the testing does not allow for inference to real-life situations.

#### *A. EU: Key Investor Information Document*

Collective Investment in Transferable Securities (UCITS) are the European equivalent to US mutual funds; they are the predominant type of fund that allows retail investors to pool resources for investments that are more complicated, costly or sometimes impossible to conduct without such pooling. In 1985 the European Council and Parliament adopted Directive 85/611/EEC, which established UCITS as vehicles for collective investments for retail investors. UCITS are a unique type of fund as they can be authorized by a single Member State but operate freely throughout the EU. Therefore, establishing such a fund was an important step for creating a single market for financial services. UCITS are open-

---

*the US and the Impact Assessment Board in the EU*, in *COMPARATIVE ADMINISTRATIVE LAW 309* (Susan Rose-Ackerman & Peter Lindseth eds., 2011). For discussion of Regulatory Impact Assessments addressing concerns of sustainable development in different OECD jurisdictions, see *OECD, SUSTAINABILITY IN IMPACT ASSESSMENTS: A REVIEW OF IMPACT ASSESSMENT SYSTEMS IN SELECTED OECD COUNTRIES AND THE EUROPEAN COMMISSION* (Feb. 21, 2012), available at [http://www.oecd.org/gov/regulatory-policy/Sustainability%20in%20impact%20assessment%20SG-SD\(2011\)6-FINAL.pdf](http://www.oecd.org/gov/regulatory-policy/Sustainability%20in%20impact%20assessment%20SG-SD(2011)6-FINAL.pdf).

ended funds that invest in transferable securities and are similar to the US mutual fund but are limited in a number of ways.<sup>92</sup> While the current consumer participation in the UCITS market is narrower than US participation in mutual funds,<sup>93</sup> the trend seems toward broader participation in collective investment schemes in Europe,<sup>94</sup> requiring disclosure regimes to consider future investors and current investors.

Similar to the regulation of mutual funds in the US,<sup>95</sup> a major concern with the regulation of UCITS is how to allow consumers to benefit from the investments in these funds while addressing concerns regarding the protection of unsophisticated consumers. As a means to protect investors, UCITS were initially required to submit statutory prospectuses to the Member State regulator, which approved the prospectus as a safeguard for investors. The disclosure requirement was amended in 2001 as part of a larger regulatory reform. The 2001 amendment required UCITS to provide a simplified prospectus to investors.<sup>96</sup> The

---

<sup>92</sup> For example, the UCITS Directive imposes rules relating to fund diversification, liquidity and use of leverage. See Directive 2009/65/EC of the European Parliament and of the Council of 13 July 2009 on the Coordination of Laws, Regulations and Administrative Provisions Relating to Undertakings for Collective Investment in Transferable Securities (UCITS), O.J. OF THE EUROPEAN UNION L 302/32 (2009).

<sup>93</sup> See for example in the US: Investment Company Institute, *Worldwide Mutual Fund Assets and Flows, Third Quarter 2011*. At the time the size of the US mutual fund industry was \$12.9 trillion whereas the size of the UCITS industry was \$7.3 trillion; For data on Europe see the *Commission Staff Working Document, Impact Assessment for Amendment of Directive, 2009/65/EC* (2012), §3, available at [http://ec.europa.eu/internal\\_market/investment/docs/ucits-directive/20120703-impact-assessment\\_en.pdf](http://ec.europa.eu/internal_market/investment/docs/ucits-directive/20120703-impact-assessment_en.pdf), in which it suggested that based on the statistics of six Member States 10% of EU households are invested in mutual funds. Compare to 44% of US households that own mutual funds. See INVESTMENT COMPANY INSTITUTE, 2013 INVESTMENT COMPANY FACT BOOK 90, available at [http://www.ici.org/pdf/2013\\_factbook.pdf](http://www.ici.org/pdf/2013_factbook.pdf).

<sup>94</sup> *Impact Assessment Accompanying White Paper on Enhancing the Single Market Framework for Investment Funds*, COM(2006) 686 Final, at 16 (Nov. 15, 2006), available at [http://ec.europa.eu/internal\\_market/investment/docs/legal\\_texts/whitepaper/impact\\_assessment\\_en.pdf](http://ec.europa.eu/internal_market/investment/docs/legal_texts/whitepaper/impact_assessment_en.pdf).

<sup>95</sup> For the US attempt at creating a summary prospectus for mutual funds, see SEC SUMMARY PROSPECTUS, *supra* note 67.

<sup>96</sup> The actual Directive was adopted in 2001. Directive 2001/107/EC of the European Parliament and of the Council of 21 January 2002 Amending Council Directive 85/611/EEC on the Coordination of Laws, Regulations and Adminis-

simplified prospectus, which came into full use in late 2005, was intended to include only the key features of the fund, such as the investment aims, fee structure and risks, in simple English for consumers.

In 2006, two workshops about the simplified prospectus revealed that the simplified disclosure was not serving its purpose.<sup>97</sup> The workshops determined that the simplified prospectus was ineffective because it was implemented differently by Member States and often contained too much information to be a concise document.<sup>98</sup> Information about risks, costs, and performance was calculated and presented inconsistently making it difficult to use the simplified prospectus and compare across funds.<sup>99</sup> In general, retail investors did not easily understand the information in the simplified prospectus and producing the simplified prospectus was costly for UCITS. Ultimately, the simplified disclosure was deemed to have “failed to become the clear, concise, and meaningful information document that it was intended to be”<sup>100</sup> and was therefore “a regulatory failure”.<sup>101</sup>

This prompted an extensive search for an alternative to the simplified prospectus, which led to the proposal of requiring funds to provide a Key Investor Information Document (“KIID”).<sup>102</sup> The

---

trative Provisions Relating to Undertakings for Collective Investment in Transferable Securities (UCITS) With a View to Regulating Management Companies and Simplified Prospectuses, O.J. OF THE EURO. UNION L 41/20 (2002).

<sup>97</sup> For full reports of the workshops see Key Investor Information, EUROPEAN COMM’N, *available at* [http://ec.europa.eu/internal\\_market/investment/investor\\_information/index\\_en.htm](http://ec.europa.eu/internal_market/investment/investor_information/index_en.htm).

<sup>98</sup> Consultation Paper on Content and Form of Key Investor Information Disclosure for UCITS 10, THE COMMITTEE OF EUROPEAN SECURITIES REGULATORS (CESR), *available at* [http://www.esma.europa.eu/system/files/07\\_669.pdf](http://www.esma.europa.eu/system/files/07_669.pdf) [hereinafter CESR CONSULTATION PAPER 2007].

<sup>99</sup> See also *Impact Assessment Accompanying Draft Commission Regulation Implementing Directive*, 2009/65/EC (2010) at 16, *available at* [http://ec.europa.eu/internal\\_market/investment/docs/ucits-directive/100521-impact\\_assessment\\_en.pdf](http://ec.europa.eu/internal_market/investment/docs/ucits-directive/100521-impact_assessment_en.pdf) [hereinafter *Impact Assessment* 2010].

<sup>100</sup> *Simplified Prospectus Workshops, Issues Paper*, EUROPEAN COMM’N at 13 (May 2006), [http://ec.europa.eu/finance/investment/docs/other\\_docs/prospectus/bp-1st-sp-workshop\\_en.pdf](http://ec.europa.eu/finance/investment/docs/other_docs/prospectus/bp-1st-sp-workshop_en.pdf).

<sup>101</sup> CESR CONSULTATION PAPER 2007, *supra* note 98, at 10.

<sup>102</sup> For the various documents related to the adoption see *UCITS – Undertakings For the Collective Investment in Transferable Securities*, EUROPEAN COMM’N, [http://ec.europa.eu/internal\\_market/investment/ucits-directive/in](http://ec.europa.eu/internal_market/investment/ucits-directive/in)

development of recommendations for the technical details and definitions of this document was assigned to the European Securities Committee and the Committee of European Securities Regulators (“CESR”), the predecessor of European Securities and Markets Authority (“ESMA”).<sup>103</sup> CESR’s initial set of advice<sup>104</sup> used the consumer testing research of the KIID proposed format conducted by IFF Research and YouGov.<sup>105</sup>

The UCITS IV Directive was finally adopted in June 2009.<sup>106</sup> In April 2010, the European Securities Committee voted in favor of the implementing measures based on the CESR’s advice, which was accompanied by an impact assessment.<sup>107</sup> The Commission, in July 2010, finally adopted the implementing measures containing the details of the KIID requirement. UCITS IV came into effect on July 1, 2012.<sup>108</sup>

---

dex\_en.htm. The regulatory process leading to the adoption of the KIID is extensive including many stages such as bilateral informal meetings with stakeholders, studies, expert groups, open hearings, public consultations, workshops and impact assessments as well as multiple level rule making.

<sup>103</sup> See *Lamfalussy Report*, EUROPEAN COMM’N, [http://ec.europa.eu/internal\\_market/securities/lamfalussy/report/index\\_en.htm](http://ec.europa.eu/internal_market/securities/lamfalussy/report/index_en.htm).

<sup>104</sup> In 2007 the Commission submitted a request for assistance to CESR in developing the implementing measures of the KIID. CESR then engaged in a number of rounds of consultation papers and comments, eventually producing its initial advice.

<sup>105</sup> IFF RESEARCH & YOUGov, UCITS DISCLOSURE TESTING RESEARCH REPORT (2009), available at [http://ec.europa.eu/internal\\_market/investment/docs/other\\_docs/research\\_report\\_en.pdf](http://ec.europa.eu/internal_market/investment/docs/other_docs/research_report_en.pdf) [hereinafter KIID DISCLOSURE RESEARCH REPORT].

<sup>106</sup> Undertakings for Collective Investment in Transferable Securities Directive 85/611/EEC of 20 December 1985 amended, inter alia, by Directives 2001/107/EC (“Management company Directive”) of 21 January 2002 and 2001/108/EC (“Product Directive”) of 21 January 2002 (UCITS III directives). UCITS IV included many aspects of which the adoption of the KIID was only one. An important part of the adoption process was the impact assessment, which discussed core issues, such as whether the KIID should prioritize harmonization and standardization over flexibility. See *Commission Staff Working Document Accompanying the Proposal for a Directive of the European Parliament and of the Council*, SEC(2008) 2263 (2008), available at [http://ec.europa.eu/internal\\_market/investment/docs/ucits-directive/ia\\_report\\_en.pdf](http://ec.europa.eu/internal_market/investment/docs/ucits-directive/ia_report_en.pdf) [hereinafter *Impact Assessment 2008*]. *Impact Assessment 2010* dealt with Level 2. *Impact Assessment 2010*, *supra* note 99. In addition, not all Level 2 issues were considered in depth. *Id.* The issue of consumer response to KIID was considered.

<sup>107</sup> *Impact Assessment 2010*, *supra* note 99 (dealing with Level 2 regulation).

<sup>108</sup> The advice on the content of the KIID was later supplemented with

## 1. Methodology of Testing

A major criticism of the simplified prospectus (the disclosure requirement prior to the adoption of the KIID) was that the Commission did not carry out a detailed analysis of alternative regulatory tools and the expected effect of the proposed regulation. In one discussion of the failure of the simplified prospectus, the Commission points out that the disclosure document was developed without consumer testing of the document. Therefore, the document lacked evidence that the information in the simplified prospectus was information that consumers needed or understood.<sup>109</sup> The need to carefully consider the impact of regulation and consider alternatives is more than apparent in the hundreds of pages of assessments and reports produced by the various European institutions.

The study by IFF Research and YouGov was intended to determine which standardized format of the KIID would be most comprehensible to consumers. The five elements of the KIID tested were: 1) strategy and objectives; 2) risk and reward profile; 3) past performance; 4) charges; and 5) additional information. Unlike the CFPB's testing, discussed below, the KIID testing was not meant to determine whether the KIID performed better than the simplified prospectus, but rather to test different potential KIID formats.

The Study was designed to provide information on consumer preferences as well as information regarding consumers' comprehension of information in the KIID.<sup>110</sup> The study included a variety of research methods such as online testing, telephone interviews and focus groups. In addition to quantitative research that involved online testing of consumers,<sup>111</sup> the study included a

---

technical advice regarding the calculation of the synthetic risk and reward indicator, discussed below, and the calculation of ongoing charges. See *Guideline on the Methodology for the Calculation of the Synthetic Risk and Reward Indicator in the Key Investor Information Document*, CESR/10-673 (2010), available at [http://www.esma.europa.eu/system/files/10\\_673.pdf](http://www.esma.europa.eu/system/files/10_673.pdf) [hereinafter *Risk Indicator Calculation Guidelines*]; *Guideline on the Methodology for Calculation of Ongoing Charges Figure in the Key Investor Information Document*, CESR/10-674 (2010), available at [http://www.esma.europa.eu/system/files/10\\_674.pdf](http://www.esma.europa.eu/system/files/10_674.pdf).

<sup>109</sup> *Impact Assessment 2010*, *supra* note 99, at 133.

<sup>110</sup> CESR CONSULTATION PAPER 2007, *supra* note 98, at 54.

<sup>111</sup> For a more detailed description of the study see KIID DISCLOSURE RESEARCH REPORT, *supra* note 105, at 6.



qualitative stage in which various consumers and intermediaries were interviewed. The quantitative study, on which I focus, was divided into two phases. The first stage explored presentation options for individual key elements of the KIID with over 500 participants from each of the 7 member states that were included in the study. Disclosure presentations that were more successful were combined into two variants of KIID, which were then tested in the second stage with 600 participants from each of the 7 member states. The study took place online. Participants were divided into two groups, each receiving one of the two KIID variants for two mock UCITS funds. One of the funds was an Emerging Market Fund while the other was an Absolute Return Fund with a lower risk profile.

Participants were then asked questions about the content of the documents. Many of the questions related to information in the KIID or questions that required basic interpretation. For example, participants were asked whether it is true or false that “[b]ecause the fund invests in bonds 40% of my capital is protected whatever happens”.<sup>112</sup> Other questions focused on the comparison between the two types of funds. For example, participants were asked to choose for which of the two funds was it more likely to receive much less than originally invested<sup>113</sup> and to select the fund with the higher ongoing charge.<sup>114</sup> Thus, the study compared the answers between each group of participants to reveal the variant for which participants were able to provide more accurate answers.

## 2. Testing Results

The Commission used three dimensions to evaluate

---

<sup>112</sup> This question was on the “Strategy and Objectives” part of the mock KIID of an Emerging Market Fund. Other questions include whether it is true or false that “The fund can invest in bonds from any country” and “The majority of the fund’s assets must be invested in stocks but the remaining part may be invested in bonds.” *See id.* at 42. For the Absolute Return Fund, participants were asked questions like whether the statements “The fund intends to lessen the effects of sharp rises and falls in value” and “The value of the fund will closely follow market trends” were true or false. *See id.* at 46.

<sup>113</sup> *See id.* at 54. Participants were also asked to select which of the two funds has the “greater chance of achieving a more stable growth without too many sharp ups or downs”.

<sup>114</sup> *See id.* at 112. Other questions include selecting the fund that “incurs additional charges if the fund achieves its objective”. For a full list of questions see *id.*

information in the KIID. The first dimension was whether the information is engaging and this dimension was closely related to consumer preferences.<sup>115</sup> The second dimension was whether retail investors understood information. The last dimension was whether information was presented in a way that facilitates comparison with other funds.

While results varied across categories and question types, overall comprehension levels were not very high. A high percentage of participants (around 15%-20%) responded, “don’t know,” to questions. In some categories the mean of correct answers was 1.4 out of 5 questions.<sup>116</sup>

The purpose of using a full mock-up KIID was to allow the study to test overall interaction with the document. When asked what section participants would pay the “most attention” to, the risk and reward section was of most interest followed by the Objectives and Investment Strategy section. Only 7% would pay “most attention” to the charges section. When asked what section was easiest to understand, participants were “most likely” to select the past performance section and “least likely” to select the Additional Information and charges sections. Many participants also selected the charges section as the hardest to understand particularly with regards to the variant containing the illustration of charges.

The study’s report carefully concludes that the proposed KIID promotes consumer understanding. However, the study is interpreted and summarized in a way that lacks nuance in the Commission’s Impact Assessment of the UCITS IV implementing measures: “Research shows that standardisation in content and presentation of information aids investors in comparing between funds and, if standardisation focuses on delivery of key information in an investor-friendly manner, it can aid comprehension of messages and engagement with disclosures.”<sup>117</sup> Ultimately, the scope of the quantitative study is quite narrow. The qualitative stage of the study involved comparison of KIID to pre-sales information available at the time. However, the quantitative stage was restricted to comparing proposed KIID formats.

---

<sup>115</sup> The wording of the Commission is “is the information presented so as to engage with retail investors?” *Impact Assessment 2010, supra* note 99, at 132.

<sup>116</sup> KIID DISCLOSURE RESEARCH REPORT, *supra* note 105, at 36.

<sup>117</sup> *Impact Assessment 2010, supra* note 99, at 17.

Broader inferences may be made from the European study on the use of disclosure. The disclosure presentations tested in the study follow many of the insights suggested by authors as to when disclosure might be effective. The disclosure presentations are written in simplified and plain English<sup>118</sup> and it is a standardized document that facilitates comparison.<sup>119</sup> In one variant, it used a scale to present risk,<sup>120</sup> and it presented costs using a specific illustration.<sup>121</sup> Despite the incorporation of many guidelines as to how to create effective disclosures, comprehension levels were low.

### 3. Subsequent Experiments

Following the adoption of the KIID, an online experiment sought to test its effectiveness in meeting its regulatory goals.<sup>122</sup> In the experiment, 137 participants were randomly divided into a treatment group receiving the KIID and a control group receiving the full prospectus. Participants were first asked to evaluate the documents, in which the KIID was perceived as of better quality.<sup>123</sup> Participants were then asked to make a hypothetical investment decision on an endowment of €10,000 for an investment horizon of 15-20 years. The amount invested in the risk fee asset was similar for the group receiving the KIID and the group receiving the full prospectus. However, the treatment group receiving the KIID was less likely to engage in naïve diversification<sup>124</sup> of dividing assets equally amongst offered funds.<sup>125</sup>

---

<sup>118</sup> On simplification of disclosure, see Loewenstein, Sunstein & Golman, *supra* note 49, at 21; FEDERAL RESERVE BULLETIN 2011, *supra* note 12, at 10; Willis, *supra* note 2, at 821. See study and conclusions of FTC REPORT 2007, *supra* note 69. -

<sup>119</sup> Loewenstein, Sunstein & Golman, *supra* note 49, at 22.

<sup>120</sup> FEDERAL RESERVE BULLETIN 2011, *supra* note 12, at 12.

<sup>121</sup> *Id.* at 15.

<sup>122</sup> Torsten Walther, *Key Investor Documents and Their Consequences on Investment Behavior*, 85 J. BUS. ECON. 129 (2014).

<sup>123</sup> On all ten aspects of quality, including comprehensibility and helpfulness, the KIID was perceived as of better quality. However, half the subjects receiving the KIID state that they are unable to understand the information. The study also showed that those receiving the KIID are far less likely to claim that the document contains too much information. *Id.* at 8.

<sup>124</sup> Shlomo Benartzi & Richard Thaler, *Naïve Diversification Strategies in Defined Contribution Savings Plans*, 91 AM. ECON. REV. 79 (2001).

<sup>125</sup> The author suggests that the lack of information overload for participants receiving the KIID decreases the probability of naïve diversification, however makes no additional evaluation of investment decisions. For example, it is

The study provides an important extension to the EU study in that it tests the KIID relative to the full prospectus, thereby comparing the new regulatory regime to the existing full prospectus. However, a more appropriate comparison would be between the KIID and the simplified prospectus, the previously existing regulatory regime.

Moreover, the study appears to have limited value in developing out understanding of the impact of KIID. Firstly, the extent to which the subjective comprehensibility of a document is correlated with the objective ability to understand a document is unclear. Therefore, the experiment's focus on subjective dimensions is possibly of little importance to broader issues of consumer protection. Secondly, the experiment does not address other aspects of investment decisions beyond naïve diversification.

In another study by Oehler *et al.*,<sup>126</sup> finance students were surveyed regarding the benefits of the KIID.<sup>127</sup> The study found that students deemed the KIID developed by the researchers, called the “neutral benchmark” KIID, more effective than KIIDs developed by suppliers and issuers.

The next section discusses the various methodological concerns raised by the testing of the KIID.

### B. US: CFPB Testing of Mortgage Disclosure

The Dodd-Frank Wall Street Reform and Consumer Protection Act created the Consumer Financial Protection Bureau (CFPB). The CFPB was directed to propose rules and disclosures combining the disclosures under the Truth in Lending Act (TILA)

---

unclear whether those that did not engage in naïve diversification underutilized the opportunity to diversify their portfolio, and so it is possible those receiving the KIID made worse investment decisions. Walther, *supra* note 122, at 15.

<sup>126</sup> Andreas Oehler, Andreas Hofer & Stefan Wendt, *Do Key Investor Information Documents Enhance Retail Investors' Understanding of Financial Products? Empirical Evidence*, J. FIN. REG. & COMPLIANCE [forthcoming 2015].

<sup>127</sup> The choice to have finance students surveyed was because the researchers considered them knowledgeable enough to assess financial documents. *Id.* at 8. However, it is not clear how to generalize the assessments of finance students to all consumers. For example, data that may be considered more accurate to a finance student may be confusing to a consumer.

and the Real Estate Settlement Procedures Act (RESPA).<sup>128</sup>

TILA was enacted in 1968 and was meant to provide consumers with awareness of the cost of credit, including mortgages.<sup>129</sup> TILA requires, among other things, that consumers receive disclosure of credit terms within three days of applying for a mortgage loan. If certain terms change, another disclosure is required three days before closing. RESPA was enacted in 1974 to provide disclosure of settlement charges.<sup>130</sup> RESPA requires that consumers receive a Good Faith Estimate (GFE) of the closing costs also within three days of an application for a mortgage and that they receive a uniform settlement statement at closing (HUD-1). In 1996, Congress directed the Federal Reserve Board, in charge of implementing TILA, and the Department of Housing and Urban Development (HUD), in charge of implementing RESPA, to create a single disclosure fulfilling the requirements of both Acts. After identifying the problems with the existing disclosures,<sup>131</sup> they sought to address the issues through implementing regulations. In 2010, the responsibility to create an integrated disclosure was transferred to the CFPB with the passage of Dodd-Frank.

To develop a proposal for the new integrated disclosures, the CFPB embarked on an extensive project in 2011. The CFPB's Mortgage Disclosure Project involved contracting with a professional research company to test the different types of mortgage disclosures. The report on the qualitative stage was published in July 2012.<sup>132</sup> Following the qualitative study, the CFPB proposed an integrated disclosure to be provided after consumers apply for the mortgage (the Loan Estimate), replacing the initial TILA disclosure and GFE, and an integrated disclosure to be provided before the loan closing (the Closing Disclosure), replacing the final TILA disclosure and the HUD-1 under RESPA.

---

<sup>128</sup> §§1098, 1100A of Dodd-Frank.

<sup>129</sup> 15 U.S.C. § 1601(a) (2015). For a critical analysis of the adoption process of TILA see Rubin, *supra* note 44, at 242.

<sup>130</sup> 12 U.S.C. § 2601(a), (b)(1) (2015).

<sup>131</sup> See, e.g., Jinkook Lee & Jeanne Hogarth, *The Price of Money: Consumers' Understanding of APRs and Contract Interest Rates*, 18 J. PUB. POL'Y & MARKETING 66 (1999).

<sup>132</sup> CONSUMER FIN. PROT. BUREAU, KNOW BEFORE YOU OWE: EVOLUTION OF THE INTEGRATED TILA-RESPA DISCLOSURES (July 9, 2012), available at [http://files.consumerfinance.gov/f/201207\\_cfpb\\_report\\_tila-respa-testing.pdf](http://files.consumerfinance.gov/f/201207_cfpb_report_tila-respa-testing.pdf).

In its quantitative study, the CFPB used the proposal of the new integrated disclosures to validate their effectiveness and to compare their performance to the previously existing disclosures. My analysis focuses on the quantitative stage, the report of which was published in November 2013, although many of my comments also refer to the qualitative stage.<sup>133</sup>

The CFPB's Mortgage Disclosure Project stated that its three main objectives are to create disclosures that: 1) are comprehensible so that consumers know the basic terms of the loan and its cost, 2) facilitate comparison, and 3) as a result of consumers comprehending the disclosure and comparing disclosures, allow them to make informed decisions. The efforts eventually produced a 3-page document replacing longer and overlapping disclosures.

The CFPB's disclosure project reflects an attempt at robust empirical testing of disclosure materials before their adoption and would seem like the ultimate application of the notion of evidence-based regulation. It seems unlikely that the resources and time investment in developing the integrated disclosures will be applied frequently in the future process of adopting regulation. However, it is clear that the case of mortgage disclosure is meant to be a proto-type of what robust evidence of the effectiveness of regulation is meant to look like.

### 1. Methodology of Testing

The quantitative testing compared the previously mandated disclosures with the disclosures proposed after the qualitative testing was completed in 2012. The study used a controlled experiment to assess understanding with a between-subjects factorial design. The disclosures presented to subjects varied along 3 different categories: (1) disclosure type (proposed/current); (2) loan type (fixed rate/adjustable rate); and (3) complexity of loan (easy/more challenging – such as loans with negative amortization and interest only loans).

The experiment also included two types of consumers; half of the subjects were experienced consumers while the other half

---

<sup>133</sup> For the full report see, Know Before You Owe, Quantitative Study, CONSUMER FIN. PROT. BUREAU, at xiv, [http://files.consumerfinance.gov/f/201311\\_cfpb\\_study\\_tila-respa\\_disclosure-comparison.pdf](http://files.consumerfinance.gov/f/201311_cfpb_study_tila-respa_disclosure-comparison.pdf) [hereinafter CFPB MORTGAGE DISCLOSURE STUDY].

were inexperienced consumers. A total of 858 participated in the study in 20 different locations.

Each subject received initial and final disclosures and was then asked 48 questions. The session in which subjects participated was divided into several parts, each part containing a different task, such as comparing the terms of two initial disclosures or comparing an initial and final disclosure. In each part, subjects received a questionnaire related to the disclosures.

Participants were asked standard comprehension questions relating either to the understanding of a specific document or to the comparison of two documents. For example, participants were asked: “How much principal will you pay in 5 years?” (question 24), “How does APR compare?” (question 7) and “Does this loan have mortgage insurance?” (question 25).<sup>134</sup>

The CFPB identified nine different financial concepts that they considered essential for informed consumers, which they tested through various questions. The accuracy of the answers to the questions was then compared between subjects receiving the former disclosures and the proposed disclosures.

Task 1 of the study was different than the comprehension questions presented as part of the other tasks. For Task 1, subjects were presented with their initial disclosures and asked to choose between two mortgages and justify their choice. The analysis focused on subjects’ ability to articulate their choice. The CFPB considered the fulfillment of the task based on the number of comments. In the words of the study itself, Task 1 was “not the central focus of our analyses”.<sup>135</sup>

## 2. Results

In analyzing the results, the CFPB cuts the information in various ways including groupings along parts, tasks and concepts with varying results. According to most data analyses, the percentage of subjects receiving the proposed disclosures who answered correctly is 16 percentage points more than those receiving the current (now former) disclosures.

---

<sup>134</sup> For a full list of questions, see *id.* at Appendix C.

<sup>135</sup> *Id.* at xiv.

Using the index that includes all 39 core questions according to the CFPB,<sup>136</sup> 59.3% of those that received the current disclosure answered correctly while 76.2% who received the proposed disclosures answered correctly. Similar to the European testing, the CFPB does not provide guidelines as to how to evaluate its results. The CFPB concludes in the CBA in its Final Rule: “Since the Bureau’s quantitative testing revealed that the Loan Estimate is substantially more understandable for consumers than the current early TILA disclosure and RESPA GFE, the Bureau therefore believes that the new form will enable consumers to make more informed choices when they are considering a mortgage.”<sup>137</sup>

#### IV. CONCERNS WITH THE WAY DISCLOSURES ARE TESTED

Despite the resources invested in consumer financial disclosure testing in the US and Europe, their methodology and results raise several concerns. The structure of the tests is at odds with the regulators’ stated purpose of financial disclosures. Moreover, the adoption of financial disclosures based on seemingly poor comprehension levels raises questions as to the goal of disclosures.

In this section, I will discuss the various concerns with current testing methodologies. I first discuss how the structure of consumer testing in experimental and isolated conditions question the external validity of the testing results. I then discuss the issue of internal validity and suggest that the current use of comprehension tests fails to test what regulators are concerned about, namely improved consumer decision-making. Furthermore, comprehension tests fail to capture potential benefits of disclosures that fall short of full comprehension.

---

<sup>136</sup> The core questions include all the questions other than Task 1 (choosing among two different loans) and excluding question 42 (“Do you have any comments about the final loan terms and costs?”) and questions 43–48 that were scale rating questions about the disclosures. For the reasons they were excluded, see *id.* at 39, note 21.

<sup>137</sup> Integrated Mortgage Disclosures Under the Real Estate Settlement Procedure Act (Regulation X) and the Truth In Lending Act (Regulation Z), Final Rule, 78 Fed. Reg. 79730, 80079 (Dec. 31, 2013) [hereinafter *Integrated Mortgage Disclosure Final Rule*].



In the next section, I discuss several ways in which to improve regulatory testing, addressing the concerns raised.

#### A. *Missing Benchmarks for Evaluation*

Both the CFPB and EU tests do not provide a benchmark for the required level of comprehension needed to justify the adoption of the proposed disclosures. One aspect in which the KIID and RESPA-TILA testing differ is what their tests are comparing. The EU tests different potential variants for the new disclosure, while the CFPB tests comprehension of the proposed disclosure relative to the existing disclosures. Since the CFPB and EU do not explicitly state the level of comprehension needed for the adoption of regulation, it is possibly the mere fact that the adopted disclosure performed better, even if only marginally.

However, the lack of a defined and justified benchmark seems problematic given the overall low levels of comprehension. While the final KIID format contained elements that were deemed superior by the study, overall levels of accurate answers were low in many cases, including for questions essential for informed investments. In the past performance, category participants exposed to the presentation of information over a 10-year period did slightly better in the study than those exposed to a 5-year period bar graph. In general, however, the level of understanding was not high. For example, the mean of correct answers was 1.5 out of 5 questions on the understanding of the Emerging Market Fund mock KIID graph.<sup>138</sup> Another issue was that a large number of participants answered “don’t know” (around 15%-20%).<sup>139</sup> For example, only 35% of participants correctly answered whether the following statement is true or false: “If you had invested in 2005, the fund would not have achieved its growth objective by the end of 2006” is true or false.<sup>140</sup> 45% of participants who were asked to

---

<sup>138</sup> KIID DISCLOSURE RESEARCH REPORT, *supra* note 105, at 93. This is true for the variant select- 10-year period bar graph. The mean was even lower for the 5-year period graph.

<sup>139</sup> *Id.* at 107. The questions in this section focused on the understanding of the past performance graphs, such as asking whether the statement “If you had invested in 2005, the fund would not have achieved its growth objective by the end of 2006” is true or false (only 35% of participants answered this question correctly, see *id.* at 99), or the comparison of two funds, such as asking which of two funds is more likely to have positive performance over the next three years. *Id.* at 104.

<sup>140</sup> *Id.* at 99.

compare two funds' past performance and determine which of two funds is more likely to have positive performance over the next three years, correctly answered "neither".<sup>141</sup>

For the charges and costs category, subjects either received a mock KIID with charges presented only in the form of a table, showing charges as percentages, or received a mock KIID that also included an illustration of the charges for a 10,000 euros investment over a period of 1, 5 and 10 years assuming an annual growth of 5%. The findings of the study were that consumers preferred to receive the variant with the illustration of charges but found it difficult to use the information to estimate charges.<sup>142</sup> Participants were particularly confused regarding entry fees and their impact on an investment.<sup>143</sup> For example, over 40% cannot accurately answer which of two funds had a higher entry charge, and over 50% cannot accurately answer which fund is likely to have the higher level of transaction costs,<sup>144</sup> key factors directly effecting investment returns. Similar to the interpretation of the past performance information, there was consistently a high percentage of participants (around 15%-20%) who responded, "don't know" to questions.

In general, superiority of one variant over another was primarily on the dimension of engagement. In other words, participants often preferred one variant to another even though there were no significant differences in comprehension levels. For example, in the risk category, the study concluded that the overall understanding was similar for a narrative description of the risks and a synthetic risk indicator, however consumers preferred the synthetic indicator.<sup>145</sup> Although it may seem inconsequential whether consumers prefer a document when it does not promote understanding, the authors of the study suggest that engagement is important in real world contexts in determining whether consumers attempt to even read the disclosure.<sup>146</sup>

Ultimately, the EU study seems to support the variant that performed better. However, it is unclear why a better performing

---

<sup>141</sup> *Id.* at 104.

<sup>142</sup> *Id.* at 123.

<sup>143</sup> *Id.* at 120.

<sup>144</sup> *Id.* at 112.

<sup>145</sup> *Id.* at 52.

<sup>146</sup> *Id.* at 152.

mock variant in a particular test is the criteria for adopting a disclosure format, rather than, for example, further developing more effective disclosures.

Regarding the CFPB's results, they seem grim in terms of the expectation that consumers understand the disclosures, or that they be fully informed. As mentioned above, the percentage of questions answered correctly by participants receiving the proposed disclosure was 76.2%, which seems quite low given that it includes both experienced and inexperienced subjects. More importantly, this average seems to be driven by some questions in which participants score particularly high. Many of these questions merely required subjects to copy information clearly labeled in the disclosure. For example, 99.1% of those receiving the proposed disclosure answered the question "What is the loan amount?" correctly;<sup>147</sup> the disclosure contained a large rubric at the top labeled "Loan Amount".

In many cases the percentage of people answering the question correctly, even when presented with the proposed disclosures, remains low.<sup>148</sup> For example, question 22 asked subjects "Do you have a penalty if you want to refinance?", an important aspect of a mortgage loan. Only 54.9% of subjects answered correctly, an increase of 1.6 percentage points from the percentage of those receiving the current disclosure that answered correctly.<sup>149</sup>

The percentage point difference between the proposed and existing disclosure was 16.9%,<sup>150</sup> however this may be overstating the effect of the proposed disclosure. Partially, this is driven by abnormally high differences for particular questions. For example, participants were asked, "How much is the first monthly payment for mortgage insurance?",<sup>151</sup> with only 2.3% of those receiving the existing disclosures answering correctly and 65.1% of those

---

<sup>147</sup> CFPB MORTGAGE DISCLOSURE STUDY, *supra* note 133, at Question 11. See also *id.* at Appendix C, H.

<sup>148</sup> See for example question 18: "When you make your first total monthly payment, for how much will you write the check?" Interestingly this question slightly breaks the cycle of asking people to copy information that can be identified by the heading in the disclosure, as it makes the connection between the information in the disclosure and an actual action. *Id.*

<sup>149</sup> See *id.* at Appendix C for full list of questions.

<sup>150</sup> *Id.* at 41.

<sup>151</sup> *Id.* at Question 26.

receiving the proposed disclosure answering correctly. This is less striking when considering that the proposed disclosures had a separate amount titled “Mortgage Insurance”, whereas the existing disclosures simply provided one amount titled “Taxes and Insurance” noting below “Includes Private Mortgage Insurance.” Furthermore, for over 40% of the questions, the percentage point difference was under 10%.<sup>152</sup> For some questions, the percentage of those answering correctly with the proposed disclosures was *lower* than those with current disclosure.<sup>153</sup>

While the CFPB seems pleased with its own results, they too do not provide a criterion to judge whether the improvement in results justifies the adoption of the proposed disclosures. It would seem that any increase, for at least most of the questions would qualify as a “success.” With such a low threshold for success it is hard to imagine a scenario in which the study does not confirm the CFPB’s belief that the proposed disclosures should be adopted, particularly considering the type of questions the CFPB asked.

### *B. Isolated Environment*

Current testing practices raise several concerns regarding their generalizability, raising the question of whether their results can be extrapolated to real-life situations.

One way in which comprehension tests may not lead to improved financial decisions in the real world relates to the controlled environments in which the testing takes place. This controlled environment where people are isolated from distractions and focus on the task at hand can differ greatly from the real life context of exposure to financial disclosures, questioning the external validity of the testing results.

For example, in the case of mortgage disclosure, the CFPB’s project was focused on creating a concise disclosure fulfilling the requirements of TILA and RESPA, which was subsequently tested. However, real-life mortgage transactions

---

<sup>152</sup> See for example, question 30 on the interest rate in year 1: “How do the application disclosures and the final disclosures compare in terms of the interest rate in year 1?” *Id.* at Question 30.

<sup>153</sup> See for example, question 16 on Settlement Charges: “How much are your estimated settlement charges?” for which there was a *decrease* of 39.9%. *Id.* at Question 16.

often involve many other disclosures mandated by state legislatures, municipal law and courts as well as other documents presented to consumers at the time of signing a mortgage agreement. Ben Shahar and Schneider suggest that mortgage transactions can be accompanied by as many as fifty separate disclosures.<sup>154</sup> These documents may include documents regarding “the loans tax consequences; the property appraisal; the lender’s credit reporting practices; agents’ conflicts of interest; the right to cancel the transaction; compliance with non-discrimination statutes; privacy and data collection; payment options; escrow choices; and much more”.<sup>155</sup> The amount of disclosures may dwarf the impact of the TILA and RESPA disclosure, and therefore the context in which participants answer comprehension questions about their mortgage based on one document seems at odds with the real life context in which consumers receive the integrated disclosure. A major focus of the effort to create an integrated disclosure was to allow key information to be easily identified by consumers.<sup>156</sup> However, even if consumers can identify the key information when presented with the integrated disclosures in the experimental context, there is little reason to believe this will also be the case when the integrated disclosure is one of fifty documents.

Another way the testing fails to reflect the real life context of financial disclosures is through the lack of human interaction in the testing. In many of the contexts where consumers receive disclosures, they are also interacting with financial advisors or brokers, so that the real life impact of disclosure depends on its interaction with other factors competing with consumer’s attention, like human communication. The presence of a broker will often override what a consumer reads in a disclosure, or make the consumer believe they do not need to read the disclosure, as the broker provides the essential information.<sup>157</sup> Despite frequent

---

<sup>154</sup> *Futility of CBA*, *supra* note 42, at 12. See also Willis, *supra* note 2, at 790.

<sup>155</sup> *Futility of CBA*, *supra* note 42, at 13.

<sup>156</sup> CFPB MORTGAGE DISCLOSURE STUDY, *supra* note 133, at xv.

<sup>157</sup> Willis, *supra* note 2, at 798; Debra Pogrand Stark, Jessica M. Choplin & Mark A. Leoboef, *Ineffective in Any Form: How Confirmation Bias and Distractions Undermine Improved Home-Loan Disclosures*, 122 YALE L.J. ONLINE 377 (2013), available at [http://canhr.org/abuse/PDFs/YLJ\\_on\\_ineffective\\_disclosures.pdf](http://canhr.org/abuse/PDFs/YLJ_on_ineffective_disclosures.pdf); Financial Services Authority *Key Facts Quick Guide: Research Findings*. *Financial Services Authority*, CONSUMER RESEARCH 41 (July 2005); Debra Stark & Jessica

misalignment of the brokers' and consumers' incentives, consumers often misunderstand this conflict of interest and are motivated to trust brokers.<sup>158</sup> The testing design of the CFPB and the EU does not include this aspect and therefore is far from reflecting real-life impact.

In 2008 the Federal Reserve Board considered the broader impact of disclosures combined with the interaction with a broker, rather than the narrow effect of disclosure alone. The Federal Reserve Board tested its proposed amendments to Regulation Z that would prohibit creditors paying mortgage brokers unless certain information regarding the broker's compensation was disclosed.<sup>159</sup> One element of the disclosure addresses the possibility that brokers offer consumers loans that are not in their interest given the payment structure. Cognitive interviews were used to test whether consumers understood this potential conflict of interest. Interviewees were asked "to imagine that they had met with a broker who had given them this agreement to read and sign".<sup>160</sup> Although this is an interesting attempt to have participants consider the situation in which they receive the disclosure, it is surprising that the interviews only looked at the impact of the document itself given how the broker-consumer relationship is built on human interaction.

The real life circumstances in which consumers receive financial disclosures include other distractions, beyond other disclosures and human interaction that compete for consumers' attention. Although it is difficult to replicate such distractions in an experimental context, some studies have created an environment with distractions, such as distractions created by television.<sup>161</sup> Some experiments have engaged participants viewing

---

Choplin, *A Cognitive and Social Psychological Analysis of Disclosure Laws and Call for Mortgage Counseling to Prevent Predatory Lending*, 16 PSYCH. PUB. POL'Y & L. 85 (2010).

<sup>158</sup> Stark & Choplin, *supra* note 157; Roman Inderst & Macro Ottaviani, *Financial Advice*, 50 J. OF ECON. LITERATURE 494 (2012); Sunita Sah, George Loewenstein & Daylian Cain, *The Burden of Disclosure: Increased Compliance with Distrusted Advice*, 104 J. PERSONALITY & SOC. PSYCH.: INTERPERSONAL REL. & GRP. PROCESSES 289 (2013).

<sup>159</sup> FED. RES. BD., SUMMARY OF FINDINGS: CONSUMER TESTING OF MORTGAGE BROKER DISCLOSURES (2008), <http://www.federalreserve.gov/newsevents/press/bcreg/20080714regzconstest.pdf>.

<sup>160</sup> *Id.* at i.

<sup>161</sup> Georganne Ylias & Patrick Heaven, *The Influence of Distraction on*

disclosures in distracting conversation, showing that they are less likely to overcome confirmation bias.<sup>162</sup>

### C. *Experiment Effects*

Concerns regarding external validity, which relates to the ability to infer from the experimental results to other relevant settings, are partially recognized in reports of regulatory testing. The EU report explains, for example, that:

“It is important to appreciate that all of our testing is subject to a research effect in that our respondents read the documents put in front of them because that is what they had been recruited for and agreed to do. It would be a mistake to assume that in reality all respondents would actually read the documents. Therefore any document testing in a research environment can be regarded as a “best-case” scenario.”<sup>163</sup>

This seems to be of particular concern with financial disclosures, given the increasing amount of evidence that people do not read them at all.<sup>164</sup> While every experiment may involve some tradeoff between the complexity of the experiment and the ability to extrapolate conclusions from it, the next section provides ways in which regulatory testing can become more accurate.

It is important to note that some concerns with regulatory testing may push in the opposite direction - people may underperform as they lack appropriate incentives. Both the CFPB and EU study provided compensation for participating in the study but no monetary incentives for answering questions

---

*Reading Comprehension: A Big Five Analysis*, 34 PERSONALITY & INDIVIDUAL DIFFERENCES 1069 (2003).

<sup>162</sup> Stark, Choplin & LaBoeuf, *supra* note 157. According to the authors, confirmation bias in this context means “cognitive biases wherein the individuals skim through documents seeking to confirm the truth of what they were told . . . and fail to skim for evidence that a statement is false”. *Id.* at 379.

<sup>163</sup> KIID DISCLOSURE RESEARCH REPORT, *supra* note 105, at 152.

<sup>164</sup> See *supra* section [the parts above that discuss failures of disclosure] of recent research on disclosure. See also MORE THAN YOU WANTED TO KNOW, *supra* note 7. In the online contract context see Florencia Marotta-Wurgler, *Some Realities of Online Contracting*, 19 SUPREME COURT ECON. REV. 11 (2011).

accurately. If comprehending disclosures requires cognitive effort we would expect more effort to be invested when provided an incentive to answer correctly. While in the context of a study, people may lack incentives. In real life, people may face potentially significant gains and losses when selecting a mutual fund to invest in or a low cost mortgage. Therefore, it is possible that real-life comprehension of disclosure documents may be higher than in regulatory studies.

While there might be some justification in designing studies so that they provide incentives to answer questions correctly, it is important to not overstate this concern. Firstly, it has been repeatedly shown that even when faced with significant financial stakes, people are sometimes reluctant to invest minimal cognitive effort, such as reading short disclosure documents.<sup>165</sup> Moreover, the extent to which providing performance based financial rewards, the standard practice for experimental economists, alters people's willingness to invest in responding to surveys has been debated extensively.<sup>166</sup>

#### D. Question Answering Format

The concerns raised in the previous section relate to the generalizability of testing disclosure materials in certain experimental settings. However, the specific format used by the CFPB and the EU in their testing raises additional concerns. As discussed in Part III, the CFPB and EU both test financial disclosures through questionnaires related to the content of the documents. The KIID study report never explicitly justifies its

---

<sup>165</sup> See *supra* Part II.A. and *supra* note 63.

<sup>166</sup> See, e.g., Vernon Smith & James Walker, *Monetary Rewards and Decision Costs in Experimental Economics*, 31 *ECON. INQUIRY* 245 (1993); Colin Camerer & Robin Hogarth, *The Effects of Financial Incentives in Experiments: A Review and Capital-Labor-Production Framework*, 19 *J. RISK & UNCERTAINTY* 7 (1999) (summarizing the results of 74 studies showing that financial incentives matter in some cases and not in others). See also Kathrin Pokorny, *Pay—But Do Not Pay Too Much: An Experimental Study on the Impact of Incentives*, 66 *J. ECON. BEHAV. & ORG.* 251 (2008).

More generally, researchers have raised myriad reasons that lab results are not generalizable. See, e.g., Steven Levitt & John List, *What Do Laboratory Experiments Measuring Social Preferences Reveal About the Real World?*, 21 *J. ECON. PERSP.* 153 (2007). On the possible importance of financial incentives, see Ondrej Rydval & Andreas Ortmann, *How Financial Incentives and Cognitive Abilities Affect Task Performance in Laboratory Settings: An Illustration*, 85 *ECON. LETTERS* 315 (2004).



decision to focus on comprehension despite the stated purpose of the disclosure to improve consumer decision-making and competition between funds. Similar to the European testing, the CFPB focuses its study on consumer comprehension when its stated purpose of financial disclosures is to help people make the best personal decisions that they believe are best and to insure “more knowledgeable choices”.<sup>167</sup> The CFPB does not explain why it chose not to test whether the financial disclosures allow consumers to make better decisions. The CFPB simply focused on comprehension. This is problematic for several reasons.

### 1. Saliency

First, an important aspect of consumers being able to use disclosures in a meaningful way relates to whether information is salient to them. In other words, people need to, first of all, notice important information in order to understand and then use the information. Research on financial disclosures has revealed that a primary concern is that people often fail to recognize important information, clearly a required stage for understanding the information.<sup>168</sup> If this is a major concern with financial disclosures, then the regulators should also test whether people notice important information. However, the format of asking comprehension questions somewhat undermines that inquiry, since the questions themselves focus people’s attention.<sup>169</sup> In a real-life context, people look over documents and make decisions, rather than being asked comprehension questions that shape their attention.

For example, in regards to mortgage disclosures, an important issue is whether consumers, when looking at the disclosures, notice if the loan has mortgage insurance. When subjects are explicitly asked whether the loan has mortgage

---

<sup>167</sup> CFPB MORTGAGE DISCLOSURE STUDY, *supra* note 133, at xvi.

<sup>168</sup> Knowing how to identify the important information in a disclosure may be a significant barrier to the effectiveness of a disclosure. This has led to a number of proposals to have disclosure focus on information that corrects misperceptions. See Oren Bar-Gill, *The Behavioral Economics of Consumer Contracts*, 92 MINN. L. REV. 749 (2008). See also Ayres & Alan Schwartz, *The No-Reading Problem in Consumer Contract Law*, 66 STAN. L. REV. 545 (2014) (discussing the proposal to have disclosure focus on unfavorable terms).

<sup>169</sup> In both studies, subjects answered questions when they had the disclosures in front of them. In other words, participants were not asked to recall information they had previously observed.

insurance (question 25),<sup>170</sup> it clearly increases the chance that consumers will look for the mortgage insurance rubric, as this information becomes important and therefore salient. However, we do not know whether the consumer would notice this information absent the question. As the CFPB study is designed primarily around asking subjects specific questions, the study by definition skews the saliency of information.

In the context of mutual funds, Choi et al.<sup>171</sup> indicates that people overlook fees. In many cases people cannot recall looking at the fees of a fund, let alone remember what the fees were. Looking to the way the EU tested the KIID, it is clear that by asking participants to compare fees among funds, it encourages participants to look at the fees. In other words, the European Commission's results may indicate that a certain percentage of people correctly answer questions about fees when asked, but tells us little about what people will know when they read the KIID and are not prompted by questions.

## 2. Comprehension Tests as "Under-Conclusive"

The regulation of consumer finance through disclosure relies crucially on people using disclosures to improve decisions. For example, the requirement that conflicts of interest be disclosed is based on the assumption that once these conflicts are known to people they will be able to accurately evaluate advice they receive given the advisor's incentives. This means that even if people understand the basic content of the disclosure, if they under-react or over-react to the conflict of interest, they may make distorted decisions.<sup>172</sup>

Therefore, people may pass comprehensibility tests regarding disclosures, however, the disclosure does not necessarily

---

<sup>170</sup> CFPB MORTGAGE DISCLOSURE STUDY, *supra* note 133, at 22.

<sup>171</sup> James Choi, David Laibson & Brigitte Madrian, *Why Does the Law of One Price Fail? An Experiment on Index Mutual Funds*, 23:4 REV. FIN. STUD. 1405, 1432 (2010).

<sup>172</sup> For an experiment that tests this very question, see Daylian Cain, George Loewenstein & Don Moore, *The Dirt on Coming Clean: Perverse Effects of Disclosing Conflicts of Interest*, 34 J. LEGAL STUD. 1 (2005). For an additional perspective on the difference between conceptual financial knowledge and financial competence as related to welfare, see Sandro Ambuehl, Douglas Bernheim & Annamaria Lusardi, *Financial Education, Financial Competence and Consumer Welfare* (NBER, Working Paper No. 20618, 2014).

improve their financial decisions. Of course, one needs to clarify what is meant by “understanding” information, but having a thin notion of comprehension, one that focuses on the ability to answer questions on information that is clearly stated in the disclosure material, does not mean a person is better able to make a financial decision. This is because for comprehension to lead to better decisions more knowledge is required than understanding individual components of the disclosure. Making decisions requires processing information initially comprehended at face level, meaning that comprehending information does not automatically lead to better decisions, and therefore comprehension tests are “under-conclusive” relative to what is expected of disclosure.

I discuss three ways in which information is processed in order to reach a decision, all of which are overlooked by the current testing focused on comprehension.

Firstly, replicating information in the disclosure, the focus of many testing questions, does not necessarily reflect understanding. Second, even when disclosures are comprehended, it is the relative importance of different types of information contained in the disclosure that have significant impact on the quality of the final decision. As financial disclosures tend to contain myriad types of information, this issue is central to whether the disclosure helps people improve their decisions. Third, translating information into a decision requires understanding how the information applies to the individual, given their circumstances and preferences.

*Understanding Financial Concepts:* Current testing focuses primarily on replicating information in the disclosure without necessarily understanding this information. This is often the case when comprehension tests ask questions regarding financial concepts, when actually understanding these concepts is not required for a correct answer.

This is particularly stark in the case of the CFPB testing where the wording of many questions is identical to the disclosure itself. For example, question 19 asks “Which of the following settlements services, if any, can you shop for?” and then lists a number of services. The proposed disclosure contains a bolded heading “Services you can shop for” under which there is a list of

services. Therefore, a participant in the CFPB's study can easily answer this question correctly<sup>173</sup> without knowing what the various services are or even what shopping for services means.<sup>174</sup> A similar issue exists for questions about the escrow fund, APR and mortgage insurance. All these questions can be answered correctly even absent any basic understanding of what these concepts entail.<sup>175</sup> If testing does not even guarantee a narrow concept of comprehension it is unlikely to improve consumer decision-making.

*Understanding relative importance of information:* Financial disclosures contain many types of information, which carry different weight. In order to effectively use information for making decisions it is important to understand the relative importance of information for the actual decision. While the weight given to information may depend on consumers' preferences or on their particular situation, as discussed further in the next section, many optimal decisions depend on preference-independent information weighting, at least to some extent.

In the context of investment decisions, overlooking the issue of understanding the relative importance of information and focusing solely on information being clear, does not address the concerns raised by the extensive psychology and behavioral economics literature that has attempted to understand consumer behavior. This literature discusses two main problems with consumer decision-making. Firstly, consumers tend to overlook the fees charged by funds or else to discount their importance.<sup>176</sup> There are a number of possible explanations for this behavior including people's limited processing capacity and low financial literacy. Moreover, funds often have complex charging structures

---

<sup>173</sup> Indeed, 92.9% of those receiving the proposed disclosure answered this question correctly.

<sup>174</sup> Also see the examples of questions in which consumers did particularly well discussed earlier on in this section.

<sup>175</sup> See Question 6 "How do these two loans compare in terms of escrow account?", question 23 "What is the Annual Percentage Rate (APR) for this loan?" and question 25 "Does this loan have mortgage insurance?" CFPB MORTGAGE DISCLOSURE STUDY, *supra* note 133.

<sup>176</sup> See, e.g., John Beshears, James Choi, David Laibson and Brigitte Madrian, *How Does Simplified Disclosure Affect Individuals' Mutual Fund Choices?* page 14 (NAT'L BUREAU ECON. RES., Working Paper No. 14859, 2009), available at <http://www.nber.org/papers/w14859> [hereinafter: *Beshears et al. 2009*]; Choi, Laibson & Madrian, *supra* note 171.

making it difficult even for more financially educated investors to calculate fees<sup>177</sup> and may attempt to avoid charging more salient fees while increasing shrouded costs.<sup>178</sup> Second, consumers put disproportionate weight on factors that are far less likely to reflect the future success of their investment, such as the past performance of the fund.<sup>179</sup> To a large extent seeing past performance as predicting future performance is based on peoples general intuition about past behavior indicating future behavior, which may be a useful intuition in many other contexts.<sup>180</sup> Other explanations refer to the saliency and advertising of performance information to illuminate the importance past performance plays in investor decisions.<sup>181</sup>

Therefore, it is clear that for consumers to make informed investment decisions they must accurately understand the impact of fund fees and past performance on future overall returns. The current testing of financial disclosures would deem a financial disclosure to be successful if the consumer correctly answered questions regarding the fee structure and the past performance table of the fund, but would have not tested whether consumers knew the relative importance of the information on the final decision. Similarly, in the case of mortgage disclosures people may correctly answer questions on whether a loan has a floating interest rate but lack the tools to evaluate the meaning of the fact that a loan is a floating loan and how it should affect their decision. This breaks down the assumed connection between comprehension and decision-making and shows that the regulatory focus on testing

---

<sup>177</sup> See Choi, Laibson & Madrian, *supra* note 171, at 1422. In this study, MBA students who recognized the significance of fees were also unable to minimize fees based on the prospectus.

<sup>178</sup> Santosh Anagol & Hugh Hoikwang-Kim, *The Impact of Shrouded Fees: Evidence from a Natural Experiment in the Indian Mutual Funds Market*, 102 AM. ECON. REV., 576-593 (2012); Sendhil Mullainathan, Joshua Schwartzstein & Andrei Shleifer, *Coarse Thinking and Persuasion*, 123 Q.J. ECON. 577 (2008); Bruce Carlin, *Strategic Price Complexity in Retail Financial Markets*, 91 J. FIN. ECON. 278 (2009).

<sup>179</sup> See, e.g., John Kozup, Elizabeth Howlett & Michael Pagano, *The Effects of Summary Information on Consumer Perceptions of Mutual Fund Characteristics*, 42 J. CONSUMER AFF. 37 (2008); Ronald T. Wilcox, *Bargain Hunting or Star Gazing? Investors' Preferences for Stock Mutual Funds*, 76 J. BUS. 645 (2003); Erik Sirri & Peter Tufano, *Costly Search and Mutual Fund Flows*, 53 J. FIN. 1619 (1998).

<sup>180</sup> See Mullainathan, Schwartzstein & Shleifer, *supra* note 178, at 605.

<sup>181</sup> See Alan Palmiter & Ahmed Taha, *Mutual Fund Performance Advertising: Inherently and Materially Misleading?*, 46 GA. L. REV. 289, 301 (2012).

comprehension does not test the ultimate impact of the disclosures on decision-making and falls short of countering the true difficulties faced by consumers. Even if consumers were to read simplified disclosures and understand their content, problems with financial literacy and the behavioral economics literature suggests that consumers would continue to make suboptimal investment decisions.

*Knowledge of oneself:* Another important way in which information is used to make a decision involves applying the information to a particular circumstance. In the context of mortgages, loans vary along a number of characteristics, and the preference of one type over another depends on knowing ones circumstance and interpreting how the loan characteristics are compatible with the particular circumstance. For example, an Adjustable Rate Mortgage may be more appropriate for someone who does not plan to live long-term in the home they are purchasing. Making a correct decision requires knowing oneself (how long I plan to live in this home) and how this knowledge translates to a good decision (Adjustable Rate Mortgages are unlikely to be good if I plan to live long-term in the home). Similarly, many investment decisions require adjusting the investment's risk with the length of the investment and the proximity to retirement, knowledge that goes beyond understanding individual fund terms.

Therefore, even if information in the financial disclosure is understood but then not properly translated to the individual's circumstances, the disclosure does little to improve the financial decision. Current testing which focuses on abstract comprehension overlooks this required connection between comprehension and improved decision making.<sup>182</sup>

### 3. Comprehension Tests as "Over-Conclusive"

As discussed above, regulators often assume a connection between comprehension and improved decision-making; however,

---

<sup>182</sup> Interestingly, qualitative regulatory testing seems to be more concerned with actual decision-making. In a Federal Reserve publication from 2008 discussing lessons learnt from consumer testing, its summary of qualitative testing, unlike quantitative testing, considers how focus groups and interviews can be used "to identify why and how consumers make decisions and what information they use in the decision making process." FED. RES. BULL. 2011, *supra* note 12, at 8.

in many cases information can be understood but then incorrectly applied leading to suboptimal financial decisions. Comprehension tests may also be unhelpful for the opposite reason – they may require too much from the financial disclosure, since financial disclosures may be playing a role that is different from the role comprehension tests assume. Comprehension tests are meant to conclude more than is required from disclosure and are therefore “over-conclusive”.

It is hard to believe that regulators consider full comprehension as the benchmark for judging financial disclosures given their willingness to adopt disclosures even when comprehension results are low. The final report of the EU study suggests that although very few consumers who studied the KIID “understood every word or concept” they nonetheless “could benefit from using the KII document either as a quick first point of comparison before seeking more detailed information. . . and a means of arming themselves with questions to ask a financial advisor”.<sup>183</sup> If the purpose of the KIID for many consumers is an initial exposure to information laying down the foundation for a future inquiry, it is at odds with the methodology chosen by the EU to test the KIID. None of the questions in the study related to whether people intended to seek financial advice or what questions they would ask following the exposure to the KIID. However, despite the recognition in the conclusion of the report, the study was not set up in a way to test these possible benefits to disclosure.

The suggestion in the KIID study report, that financial disclosures may fulfill a function that comprehension tests fail to identify, may be correct. There are a number of possible mechanisms through which financial disclosures improve decision-making. By focusing on comprehension tests, regulatory testing today is limited to the mechanism of understanding financial disclosures and relying on them for financial decisions. If regulators tested financial decisions or financial attitudes, as discussed in the next section, rather than comprehension, it would be possible to widen the inquiry to other mechanisms through which financial disclosures could influence financial decision-making.

An interesting parallel can be made from financial literacy

---

<sup>183</sup> KIID DISCLOSURE RESEARCH REPORT, *supra* note 105, at 150.

research. Traditionally financial literacy tests, and the evaluation of financial education efforts, entailed asking people basic financial questions<sup>184</sup> like the famous Big Three test.<sup>185</sup> However, a recent study has taken a different approach in evaluating the impact of a financial education program.<sup>186</sup> Carpena *et al.* used a randomized experiment to measure the impact of a five-week education program in India. They suggest that while other measures of financial literacy rely on numeric and computational skills, financial education may influence decision-making in other ways:

“for instance by making individuals and households more aware of product choices available to them, equipping them to ask the right questions of financial providers, encouraging them to seek professional and personalized financial advice, and changing their attitudes towards purchasing and recommending formal financial products and services.”<sup>187</sup>

In their study, Carpena *et al.* find that while the financial education program did not improve financial decisions that required numeracy, it did improve participants’ financial product awareness and attitudes towards financial products.<sup>188</sup>

Similarly, in the context of financial disclosures, comprehension tests may be failing to identify the correct mechanism through which disclosure might improve financial decision-making. Current regulatory testing looks exclusively to full comprehension of financial disclosures as the standard of effectiveness. However, the notion of disclosure supporting decision-making through this mechanism has been recognized by previous literature on disclosure. For example, Kozup *et al.* argue that disclosures can educate consumers more broadly about

---

<sup>184</sup> For recent surveys on financial literacy see Lusardi & Mitchell, *supra* note 4; Hastings, Madrian and Skimmyhorn, *supra* note 35.

<sup>185</sup> Annamaria Lusardi & Olivia S. Mitchell, *Financial Literacy and Planning: Implications for Retirement Wellbeing* (PENSION RES. COUNCIL, Working Paper No. 17078, 2011), available at <http://www.nber.org/papers/w17078.pdf>.

<sup>186</sup> Fenella Carpena, *Unpacking the Causal Chain of Financial Literacy* (THE WORLD BANK, Working Paper No. 5798, 2011).

<sup>187</sup> *Id.* at 3.

<sup>188</sup> *Id.* at 13.



financial products or the industry and Perry and Blumenthal argue that disclosure can encourage consumers to engage in seeking additional information.<sup>189</sup> However, the current focus on comprehension omits the possibility of testing these hypotheses.

Another alternative mechanism through which disclosure can improve decisions is through influencing consumer decisions aiming at consumers' more intuitive processes without necessarily making them more informed.<sup>190</sup> This is what Ryan Bubb referred to as "System 1 disclosure", as it aims at our cognitive processes that are automatic and unconscious.<sup>191</sup> A classic example of such a disclosure is the mandated images on cigarette boxes that depict the negative health effects of smoking. These images are unlikely to be more informative than the warnings that have long existed on cigarette packages. Instead these images wish to influence intuitive rather than deliberative choices. In the context of financial disclosures, regulators could encourage the selection of certain loans or funds through designing disclosures that encourage an intuitive decision. These types of disclosures may not be more informative but can "nudge" people to make decisions that are more welfare increasing.

The use of such a mechanism relies on the identification of a welfare increasing decision<sup>192</sup>, which may not be uniform for all consumers for many financial decisions. In addition, such a policy would require regulators to justify why the use of influencing through disclosure is preferable to other regulatory options.<sup>193</sup>

---

<sup>189</sup> Vanessa Perry & Pamela Blumenthal, *Understanding the Fine Print: The Need For Effective Testing of Mandatory Mortgage Loan Disclosures*, 31 J. PUB. POL'Y & MARKETING 305 (2012); Kozup, Howlett & Pagano, *supra* note 179, at 49.

<sup>190</sup> Note that there is a sense in which System 1 disclosure may debias consumers and therefore may be considered informative. See Christine Jolls & Cass Sunstein, *Debiasing Through Law*, 35 J. LEGAL STUD. 199 (2006).

<sup>191</sup> See Bubb, *supra* note 7. For a general account of the two systems see DANIEL KAHNEMAN, *THINKING, FAST AND SLOW* (2011).

<sup>192</sup> For a discussion on the potential difference between revealed preferences and normative preferences, see John Beshears, *How are Preferences Revealed?*, 92 J. OF PUB. ECON. 1787 (2008) [hereinafter: HOW ARE PREFERENCES REVEALED?].

<sup>193</sup> Bubb, *supra* note 7, at 117.

#### 4. Creating Documents that Score Well

Another concern raised by the indirect comprehension tests is that regulators, either consciously or unconsciously, will create disclosure documents that score higher on comprehension tests or tailor the comprehension tests to the disclosure. This further increases the likelihood that the comprehension and decision-making will diverge. Testing actual impact of the disclosure directly mitigates the ability of regulators to design disclosures that score high on comprehension tests while providing little assistance to consumers.

The concern is similar to the problem of “teaching to the test” in teaching evaluations. In that context knowledge and skills test results are used to make broader inferences regarding the level of proficiency of students on a certain topic or quality of teaching. However, through focusing on preparing students to score highly on the tests, the ability of the test results to serve as a reflection of broader abilities than those tested breaks down.<sup>194</sup> In this context the problem arises because only a narrow part of the body of knowledge is being tested, whereas with financial disclosures testing focuses only on a proxy for improved decision-making. As discussed above, particularly in the case of the CFPB, the proximity between the wording of some questions and the headings used in the disclosure, raise serious doubts as to the general inference to be made from the testing.

In the next section I discuss ways in which the testing of financial disclosures can be improved to examine a broader set of mechanisms through which financial disclosures can improve decision-making.

### V. CHANGING THE WAY WE TEST DISCLOSURE

Although some of the suggestions relate to the expansion of the scope of testing, such as retrospective analysis and dynamic testing, my main concern is not that insufficient resources are being invested in regulatory testing, but that these resources are misplaced. The concept of what needs to be tested is my focus, and not on the lack of regulatory resources allocated. In many cases, testing is already relatively extensive. The CFPB testing, for example, included several stages of qualitative and quantitative

---

<sup>194</sup> See James Popham, *Teaching to the Test?*, 58 EDUC. LEADERSHIP 16 (2001).

testing in addition to internal CFPB efforts to develop proposed disclosure documents. Moreover, previous consumer testing regarding the updating of mortgage disclosure took place over a few years, when the Federal Reserve was in charge of updating Regulation Z. Therefore, this paper relates to areas of regulation that government has already decided to invest significant efforts in testing while aiming at the wrong target.

### A. *Redesigning Testing Within Current Framework*

Even within the current framework, in which disclosures are tested in lab settings, testing methodologies should be better aligned with the desired outcome of disclosures. I discuss how to redesign testing if the desired outcome is improved decision-making. I also suggest ways to test other potentially desirable outcomes.

#### 1. Focusing on Decisions

An important preliminary question is whether regulators are (or should be) concerned with whether consumer decisions are good decisions.<sup>195</sup> The CFPB's mission is to ensure that consumers receive information that allows them to make decisions they believe are best for them.<sup>196</sup> Conceivably, the CFPB may hold that a person making a decision they believe to be good is the goal in and of itself.<sup>197</sup> One could argue that disclosure can provide people a sense of control or autonomy<sup>198</sup> regardless of its effect on their welfare.<sup>199</sup>

However, this too does not correspond to the way regulators

---

<sup>195</sup> This article does deal directly with the question of what a "good" decision is, and rather assumes that a good decision is one that increases a person's welfare. While there are many aspects of life in which there is significant uncertainty as to what increases one's welfare, in the context of consumer finance, in which one's objective is primarily pecuniary gain, these questions are less important.

<sup>196</sup> *About Us*, CONSUMER FIN. PROT. BUREAU, [www.consumerfinance.gov/the-bureau/](http://www.consumerfinance.gov/the-bureau/) (last visited October 1, 2015).

<sup>197</sup> See also Rubin, *supra* note 44 (who discuss the confusion TILA creates between the goal of legislation and the means of achieving it).

<sup>198</sup> Such a suggestion is made in Omri Ben-Shahar & Carl E. Schneider, *The Failure of Mandated Disclosure*, 159 U. PA. L. REV. 647, 650 (2011).

<sup>199</sup> I overlook the possibility that increasing one's sense of control in and of itself increases one's own welfare. For a more detailed discussion see WHY NUDGE?, *supra* note 19.

test proposed disclosures, or to what regulators ought to care about. Firstly, the EU or the CFPB did not test even this narrow sense of a decision being “good”. Moreover, it would also be incorrect to conclude that a mere decision reflects one’s belief that it is a good decision.<sup>200</sup> Second, it is doubtful whether this is a worthy purpose of regulation. The concern on which regulatory intervention is based focuses on problematic decisions consumers make because of cognitive limitations and their possible exploitations.<sup>201</sup> This is different than defining the concern as consumer’s subjective feeling that they are making suboptimal decisions. Therefore, it is unlikely that financial regulators themselves consider this the goal of regulation given that they have not articulated such a justification and often refer to goals of improved consumer outcomes.<sup>202</sup>

I therefore assume that regulators wish to improve consumer decision-making and would not hold that it is sufficient that disclosures are thinly comprehensible when tested even if consumers continue to make detrimental mistakes. My focus will be on measuring actual decision making rather than comprehension and trying to unravel the mechanism through which disclosures may improve financial decisions.

## 2. Directly Testing Decisions

If improving financial decisions is the aim of regulation, it should be the subject of regulatory testing. Beyond the indirectness of testing for improved decision-making through comprehension tests, the previous section also provided several reasons to doubt that comprehension may be a proxy for improved decisions.

Several studies have focused on decisions themselves as the measured outcome of experiments. For example, in the Beshears *et*

---

<sup>200</sup> See, e.g., HOW ARE PREFERENCES REVEALED?, *supra* note 192, which discusses the ways in which revealed preferences and normative preferences diverge. One of these cases might be when self reported preferences are different than choices, *id.* at 1792, so that one could make a decision being aware that it does not align with their preferences. This may be particularly true of complex decisions. *Id.* at 1788. This is clearly demonstrated in the study by Choi where participants invested in mutual funds with varying fees and were asked about their level of confidence in their decision. They showed that confidence that the decision was a right one fell with fees.

<sup>201</sup> See *supra* Part I for the discussion.

<sup>202</sup> See also *supra* Part I.

*al.*<sup>203</sup> and Choi *et al.* experiments, subjects were asked to make investment decisions and were paid in accordance to the performance of their portfolio. These experiments therefore provide two advantages over the CFPB and the EU's testing. Firstly, they ask subjects to actually make decisions based on the documents rather than simply ask them to answer questions about the documents. Second, they provide real monetary incentives for their decisions.<sup>204</sup>

There was one stage of the CFPB testing that involved making a decision. As mentioned above, Part 1 of the testing required subjects to examine two initial disclosures and select the loan they preferred. Although this may seem as the desired methodology in testing how disclosure impacts decisions, this was a minor part of the testing, both according to the CFPB's own account, and in terms of the evaluation of the task. The CFPB limited its analysis to the issue of whether subjects were able to articulate a justification for their choice and how many justifications they were able to provide. The CFPB designed the loan disclosures so that "neither loan was a "better" loan, so the CFPB's analysis made no judgment about which loan was selected".<sup>205</sup> This is a curious study design if one's purpose would be to see whether the disclosures improve consumer choices.

The conclusion is that regulatory testing should focus on asking people to make a financial decision, thereby testing how the exposure to different disclosures effects their decisions. In the previous section, I provided several reasons not to assume a direct connection between comprehension of a disclosure and disclosure leading to better decisions. Underlying many of my examples is that using disclosure to make an informed decision may require skills and knowledge external to the disclosure itself. Therefore, directly testing decisions may expose cases in which disclosure simply cannot be relied upon to protect consumers, as the disclosure alone cannot lead to improved decisions. These cases, in which disclosure may be an inappropriate regulatory tool, are currently obscured by the comprehension test methodology.

One challenge created by the testing of decisions is that one

---

<sup>203</sup> See *supra* note 176.

<sup>204</sup> See *supra* note 166 (discussing the influence of monetary incentives in experiments).

<sup>205</sup> CFPB MORTGAGE DISCLOSURE STUDY, *supra* note 133, at 20.

must decide how to evaluate decisions. In the case of comprehension tests, this concern does not arise, as a response to a comprehension question is either true or false. The next section deals with the question of how participants' choices can be evaluated to determine the impact of the financial disclosure. I wish to avoid overarching claims that regulators are able to always determine what a good financial decision is for every individual, and instead focus on the opportunity that experimental design offers for making inferences in easier cases. While I provide an outline and initial analysis of the various strategies regulators can use to evaluate decisions, further research is needed for a particular disclosure and financial decision being considered.

### 3. Benchmarks for Determining Improved Decisions

Any analysis of a disclosure's effectiveness must contain criterion for determining whether the decision is an improved decision. I suggest a number of ways in which decisions can be judged.

*Cases in which there is clearly an inferior decision.* In many contexts a financial decision is clearly worse than another decision, such as selecting a mortgage that is dominated by another mortgage offered. In the context of regulatory testing it is also possible to design the experiment to include dominated options. For example, two mutual funds may follow the same index and differ only on fees, so that selecting the fund with lower fees is clearly preferable.<sup>206</sup>

Similarly, mortgage disclosures can be designed in a way that one loan is more expensive than another. In an experiment sponsored by the European Commission on retail investment consumer decision-making, over 6,000 participants from eight Member States were asked to make a number of investment decisions.<sup>207</sup> In each case one investment option had the highest expected value, and therefore the experiment was designed to see

---

<sup>206</sup> The experiment by Choi had such a design, where subjects were asked to allocate funds to essentially identical index funds, differing only on the fees they charge. See Choi, et. al, *supra* note 171.

<sup>207</sup> The purpose of the project was to study consumer decision-making and was unrelated to the testing of mandated disclosures, Nick Chater, Steffen Huck, & Roman Inderst, *Consumer Decision-Making in Retail Investment Services: A Behavioural Economics Perspective*, EUR. COMM'N (2010) [hereinafter EC DECISION-MAKING REPORT 2010].

whether consumers selected the optimal investment choice.<sup>208</sup> Such experiment designs allow for easy evaluation of consumer decisions.

Consumers often face dominated options in real-life situations, which strengthens the case for using such an experimental design, and also creating opportunities for testing real-life decisions. One recent example appears in a study of Bhargava *et al.*, which discusses a US firm that permitted employees to select their own health insurance plan from a standardized menu that included many dominated options. They found that the majority of employees chose dominated health insurance options.<sup>209</sup>

Perry and Blumenthal suggest a way of evaluating decisions that is slightly more complex than when one option is clearly better than another, which can be used when the study is a field experiment and not in a controlled environment. They suggest that performance of loans selected by consumers be tracked over time to see the incidence and severity of late payments and defaults to determine whether disclosure led to consumers selecting loans with lower rates, late payments and defaults.<sup>210</sup>

*Articulation of justification.* There may be instances in which evaluating a decision may depend on a consumer's preferences. Therefore another way to evaluate decisions is to require participants to explain their decision. In the case of the CFPB study, justifications were simply counted to see the "robustness" of participants' justifications for their decisions. In a quantitative study co-sponsored by several federal agencies relating to financial privacy notice required under the Gramm-Leach-Bliley Act, a different approach was taken.<sup>211</sup> In that study, participants were

---

<sup>208</sup> The study justifies considering the investment with the highest expected value as being optimal based on expected utility maximization over one's wealth. *See id.*, at 264.

<sup>209</sup> Saurabh Bhargava, George Loewenstein & Justin Sydnor, *Do Individuals Make Sensible Health Insurance Decisions? Evidence from a Menu with Dominated Options* (AM. ECON. REV., Working Paper No. 21160, 2015).

<sup>210</sup> Perry & Blumenthal, *supra* note 189, at 309.

<sup>211</sup> Alan Levy & Manoj Hastak, *Consumer Comprehension of Financial Privacy Notices: A Report on the Results of the Quantitative Testing*, FED. TRADE COMMISSION (2008), available at [http://www.ftc.gov/system/files/documents/reports/reports/quantitative-research-levy-hastak-report/quantitative\\_research\\_-\\_levy-hastak\\_report.pdf](http://www.ftc.gov/system/files/documents/reports/reports/quantitative-research-levy-hastak-report/quantitative_research_-_levy-hastak_report.pdf).

asked to compare two bank privacy notices, assuming they provide identical products, and select which bank they prefer and explain why.<sup>212</sup> They then coded the justifications for the decision to identify when participants provided well-reasoned responses.

Although this study falls short of a full examination of consumer decisions, it evaluates the justifications given for consumer choices in order to determine whether a decision was informed or whether the reasons given for the decision were irrelevant or incorrect.

*Consistency with stated preferences.* The FTC financial privacy notice testing mentioned above dealt with disclosures that mostly differed on two dimensions – the extent of information sharing and number of opt-out possibilities. More sophisticated testing of financial disclosures may involve comparing multiple loans with several characteristics in which no loan is strictly preferable to another.

One way to evaluate decisions is on their consistency with a person's stated preferences. So for example, participants' risk preferences can be elicited before or after being asked to select a mutual fund to invest in. Researchers can then evaluate decisions in terms of their consistency with preferences.<sup>213</sup>

However, such a methodology relies on the people's stated preferences reflecting their true preferences and on the belief in the stability of people's preferences. There are many reasons to doubt these two assumptions, as people do not always accurately state their preferences<sup>214</sup> and preferences depend heavily on the context in which they are extracted.<sup>215</sup> Therefore, attempts to evaluate

---

<sup>212</sup> See also FED. RESERVE BD., CONSUMER RESEARCH AND TESTING FOR PRIVATE EDUCATION LOANS: REPORT OF FINDINGS (2009), available at <http://www.federalreserve.gov/newsevents/press/bcreg/bcreg20090311a8.pdf>.

<sup>213</sup> See, e.g., Glenn Harrison, Ronald Harstad & Elisabet Rutstrom, *Experimental Methods and Elicitation of Values*, 7 EXPERIMENTAL ECON. 123 (2004).

<sup>214</sup> On the difficulty of eliciting true risk preferences see for example Ronald Cummings & Laura Taylor, *Unbiased Value Estimates for Environmental Goods: A Cheap Talk Design for the Contingent Valuation Method*, 89 AM. ECON. REV. 649, (1999). See also Steven Levitt & John List, *What Do Laboratory Experiments Measuring Social Preferences Reveal About the Real World*, 21 J. ECON. PERSP. 153 (2007) (in the context of social preferences).

<sup>215</sup> See e.g., Steffen Anderson, et al., *Eliciting Risk and Time Preferences*, 76 ECONOMETRICA 583 (2008) (in the experiment the joint elicitation of risk and



decisions based on their consistency with elicited preferences should be done with caution.<sup>216</sup>

*Do the decisions correlate with expectations of group preferences?* In cases in which personal preferences are difficult to extract, one could use group preferences and examine whether they correlate with people's decisions. Rather than comparing an individual's preferences to a decision made after exposure to disclosure materials, inference is statistical based on group characteristics.

Peter Ubel discusses this strategy in the context of medical decision aids in which one would expect, for example, that the stage of cancer would correlate with preferences for particular treatments.<sup>217</sup> While on an individual level people may have personal preferences on other treatment dimensions, we would expect certain patterns to be revealed when comparing groups in aggregate. If aggregate decisions using medical decision aids show no correlation between levels of risk of cancer and the chosen treatment, one might be skeptical as to the effectiveness of the decision aid. In the context of consumer finance, whether to take a 30-year loan versus 15-year loan will depend partially on whether this is the first home being purchased or whether the consumer is above a certain age. While consumers within these two groups may have many additional idiosyncratic preferences beyond the length of the loan, if consumers who are purchasing their first home take out shorter loans than older consumers, concerns are raised regarding consumer decision-making.

While I do not wish to undervalue the importance of

---

time preferences yielding different results than when they were elicited separately); Ben McQuillin & Robert Sugden, *Reconciling Normative and Behavioral Economics: The Problems to be Solved*, 18 SOC. CHOICE & WELFARE 560 (2012).

<sup>216</sup> For a broader discussion on how to determine preferences see *How are Preferences Revealed?*, *supra* note 192. A similar concern is raised by Peter Ubel who discusses deriving preferences in the context of quantifying utility values for the purpose of medical treatment and the evaluation of medical decision aids, as an individual's risk appetite is relevant to their preferred medical treatment. Ubel references a number of common methods such as deriving utility from the standard gamble method. However accurately estimating risk preferences, even when examined through revealed preferences and not stated preferences, can be misleading as people are often deeply confused when dealing with probabilities and risk. See Peter Ubel, *Beyond Comprehension: Figuring Out Whether Decision Aids Improve People's Decision*, in *THE BEHAVIORAL FOUNDATIONS OF PUBLIC POLICY* (Eldar Shafir ed. 2012).

<sup>217</sup> See Ubel, *supra* note 216, at 357.

personal preferences in financial decisions, the difficulty in extracting these preferences and comparing people's decisions to these preferences should not deter regulators from evaluating decisions when they are not preference sensitive, or at least not strongly preference sensitive. Regulatory testing can make significant progress before attempting harder questions regarding idiosyncratic preferences, as financial decisions are often easier to evaluate objectively.

### *B. Testing Alternative Mechanisms*

When discussing the shortcomings of the comprehension tests currently used by regulators, one concern was that regulators assumed that the way in which disclosures helped consumers was by making consumers fully informed after comprehending the disclosure. This, however, overlooks other mechanisms by which disclosures may lead to better financial outcomes.<sup>218</sup>

One way in which disclosure may lead to better financial outcomes, not directly through full comprehension, is by flagging to the consumer the various aspects of the decision and the different components to consider. For example, a consumer who is considering taking a mortgage loan for the first time may be unaware of prepayment penalties, balloon payments or refinancing options. After reading the disclosure, a consumer still may have difficulty understanding what these payments are and accurately estimating their significance, but the consumer may now know these are important aspects of the loan and may choose to further seek information on these payments or seek advice.<sup>219</sup>

This is not to argue that disclosures necessarily work in this manner or improve financial outcomes this way, but to illustrate that there are other probable mechanisms through which disclosure may impact people that are overlooked by current testing methodologies. Comprehension questions like "How much principle will you pay in 5 years?"<sup>220</sup> or "What is the highest

---

<sup>218</sup> See *supra* Part IV.D.3, talking about comprehension as over-conclusive.

<sup>219</sup> See Ian Ayres & Alan Schwartz, *The No-Reading Problem in Consumer Contract Law*, 66 STAN. L. REV. 545 (2014). They suggest that disclosures emphasize information that is surprising to consumers. Although quite different to my proposal, they too suggest that disclosure play an alternative role to making consumers fully informed.

<sup>220</sup> CFPB MORTGAGE DISCLOSURE STUDY, *supra* note 133, at Question 24.

possible monthly principal and interest payment?<sup>221</sup> may be incorrectly answered by consumers who nonetheless have increased awareness after being exposed to disclosures and may be better equipped when seeking professional advice.

Regulatory testing needs to include a way to test whether these alternative mechanisms exist. The best way to test the impact of disclosures over time is through randomized control trials, discussed in the next section. There are also ways to test such mechanisms in more experimental settings, such as asking questions about future intentions to seek advice. People can be asked about awareness of certain payments and elements of financial decisions, without being asked comprehension questions.

### C. Expanding Testing Efforts

In many cases the reliance on lab testing to determine the impact of disclosure is insufficient. Particularly in areas in which regulators rely heavily on disclosure to protect consumers, testing should be done through randomized control trials and retrospective analysis. In cases in which there are grave concerns regarding misalignment of incentives of disclosers and consumers, dynamic testing should be used to predict the effect of disclosure requirements on regulated entities.

#### 1. Real-life Testing

Experiments have become a common way of testing decision-making in the social sciences and should be used in the regulatory context too. However, experiments raise several concerns with regards to their ability to test real-life behavior and long-term effects. While observational studies can provide some information on real-life effects of financial disclosures, randomized controlled trials (RCT) allow for experimentation in which causality can be inferred.<sup>222</sup> In a RCT, a group is allocated at random to receive an intervention or treatment to examine the effect of the treatment on a certain outcome. This allows a study to avoid various problems with uncontrolled studies such as the impact of external factors and selection bias.

---

<sup>221</sup> CFPB MORTGAGE DISCLOSURE STUDY, *supra* note 133, at Question 38.

<sup>222</sup> Cass Sunstein, *Empirically Informed Regulation*, 78 U. CHI. L. REV. 1349, 1371 (2011).

RCTs have been used to test the effect of medication for over 60 years and have become increasingly central in shaping development policy in low and middle-income countries.<sup>223</sup> However, RCTs continue to play a far more limited role in testing effectiveness of public policy interventions in the developed world.<sup>224</sup> Policy makers for several reasons have shown resistance towards RCTs such as the possible costs RCTs entail and various ethical objections.<sup>225</sup> For example, some claim that if a new intervention is beneficial it would be unethical to deny it to a certain group for the sake of the study.

It is questionable what weight these concerns should carry. Regarding the costs of RCTs, considering the resources that the CFPB and EU study required it is doubtful whether an RCT would be significantly costlier given that a real-life context requires less participant recruitment.<sup>226</sup> Moreover, the cost of RCTs should be considered in the context of their estimated benefit to consumers. With regards to the inequitable treatment necessitated by RCTs, the gradual introduction of new policies is quite common, and is rarely objected to on moral grounds.<sup>227</sup> In

---

<sup>223</sup> Christopher Deeming, *Trials and Tribulations: The 'Use' (and 'Misuse') of Evidence in Public Policy*, 47 SOC. POL'Y & ADMIN. 359, 360 (2013).

<sup>224</sup> See LAURA HAYNES, *ET AL.*, TEST, LEARN, ADAPT: DEVELOPING PUBLIC POLICY WITH RANDOMISED CONTROLLED TRIALS, BEHAV. INSIGHTS TEAM 6 (2012), available at [https://www.gov.uk/government/uploads/system/uploads/attachment\\_data/file/62529/TLA-1906126.pdf](https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/62529/TLA-1906126.pdf). Various legal scholars have advocated the use of field tests. See, e.g., ALAN SCHWARTZ, *REGULATING FOR RATIONALITY* (2014) (suggesting that field tests are a way to enable to regulator to overcome the uncertainty of how biases from experimental research will interact in reality); Thomas Ulen, *A Behavioral View of Investor Protection*, 44 LOYOLA UNIV. CHI. L. J. 1357 (2013).

<sup>225</sup> For a discussion of the use of RCTs for implementing law and the possible objections, see Michael Abramowicz, Ian Ayres & Yair Listokin, *Randomizing Law*, 159 U. PENN. L. REV. 929, 961 (2010); See also Rubin, *supra* note 44, at 302 (Rubin also suggests that political considerations avoid field experiments, as the opportunity window to implement new rules may be narrow and not to be wasted on experimentation).

<sup>226</sup> Haynes, *et al.*, *supra* note 224 at 15. Where they argue that the costs of RCTs are often exaggerated.

<sup>227</sup> This has led to the articulation of the conditions under which RCTs are ethically permissible. See R. Boruch, *et al.*, *Randomized Controlled Trials for Evaluation and Planning*, in THE SAGE HANDBOOK OF APPLIED SOCIAL RESEARCH 147 (Leonard Bickman & Debra J. Rog eds., 2009); Fives et al., *The Ethics of Randomized Control Trials in Social Settings: Can Social Trials be Scientifically Promising and Must There be Equipose?*, INT'L J. RES. & METHOD IN EDUC. (2014); See also Abramowicz, Ayres & Listokin, *supra* note

general, a shift away from determining the benefits of regulation in abstract means that there is uncertainty regarding the effect of regulation and so RCTs cannot be characterized as a denial of a known benefit to one group.<sup>228</sup> Matthew Spitzer and Eric Talley have recently discussed how the ability of regulators to engage in experimentation of new rules presents significant option value, since the regulator can simply revert back to the existing regulation if the new regulation is demonstrated as not beneficial.<sup>229</sup>

To date a number of RCTs have been conducted by legal academics,<sup>230</sup> as well as some regulatory attempts, such as the RCTs run by the Behavioural Insights Team in the UK.<sup>231</sup> One challenge to the implementation of RCTs is accurately defining the control group. A possible set up would be to expose the randomized treatment group to the proposed disclosure document while the control group would receive the existing disclosures.<sup>232</sup> Since people may compare loans from different originators randomizing at the level of geographical areas may be preferable to the individual originator, although this may create a heavy burden for originators that cover multiple areas.

RCTs would allow for several benefits in comparison to current experimentation practices. As discussed above, lab conditions may make people sufficiently focused on the disclosure to be able to understand the document, even when in reality people do not read them, or that they do not devote the needed cognitive

---

225.

<sup>228</sup> On this topic see JIM MANZI, UNCONTROLLED: THE SURPRISING PAYOFF OF TRIAL AND ERROR FOR BUSINESS, POLITICS AND SOCIETY (2012).

<sup>229</sup> See Matthew Spitzer & Eric Talley, *On Experimentation and Real Options in Financial Regulation*, 43 J. LEGAL STUD. 121 (2014). The argument is more sophisticated than presented here. They develop a game theoretic model in which benefits and costs of experimentation are borne differently by regulators and courts leading to tension between the two.

<sup>230</sup> See, e.g. Dalie Jimenez, et al., *Improving the Lives of Individuals in Financial Distress Using a Randomized Control Trial: A Research and Clinical Approach*, 20 GEORGETOWN J. ON POVERTY L. & POL'Y 449 (2013); James Greiner, Cassandra Wolos Pattanayak & Jonathan Hennessy, *The Limits of Unbundled Legal Assistance: A Randomized Study in a Massachusetts District Court and Prospects for the Future*, 126 HARV. L. REV. 903 (2013).

<sup>231</sup> See Haynes, et al., *supra* note 224.

<sup>232</sup> The cleanest control group may be a group that receives no disclosure at all, however, given how old TILA disclosure requirements are this may be unrealistic to consider a control group as one that receives no disclosure at all. See Perry & Blumenthal, *supra* note 189, at 310.

resources to understand them. RCTs would allow regulators to overcome this problem with current testing practices as well as allow them to test the impact of disclosures with human interaction of brokers and other service suppliers.

Ben Shahar and Schneider argue that the one type of cost of disclosure overlooked by regulators is the “accumulation problem”.<sup>233</sup> They claim that disclosures extract from one another by competing for consumers’ limited attention and ability to understand disclosure. Because different regulatory agencies develop disclosure materials in isolation from other disclosures, they are unaware of the interaction between disclosures. Field-testing of disclosures is likely to expose the extent of Ben Shahar and Schneider’s disclosure externalities concern, bringing what they claim to be a covert phenomenon to the surface.

## 2. Retrospective Analysis

Another important element of effective testing is to include retrospective testing after the financial disclosure has been adopted. After the adoption of a financial disclosure there is more information available on how people use the disclosure and the extent to which it improves decisions. Mechanisms that work over time, such as financial disclosure, creating awareness of the different aspects of a mortgage loan or mutual fund investment, are hard to test in a short experiment, so that post adoption analysis can reveal long-term outcomes.<sup>234</sup> Despite the fact that more is known post adoption, most testing efforts focus on the stage prior to adoption.<sup>235</sup>

The importance of regulatory look-back has been somewhat recognized since the Carter Administration in 1978; however, retrospective review has not consistently been applied.<sup>236</sup>

---

<sup>233</sup> See *Futility of CBA*, *supra* note 42.

<sup>234</sup> See also Executive Order No. 13563 (2011) (supporting retrospective analysis). For an overview of retrospective analysis techniques see Joseph Aldy, *Learning From Experience: An Assessment of the Retrospective Review of Agency Rules and the Evidence for Improving the Design and Implementation of Regulatory Policy*, NAT’L BUREAU OF ECON. RES., at 17 (2014).

<sup>235</sup> CARY COGLIANESE, OECD, MEASURING REGULATORY PERFORMANCE: EVALUATING THE IMPACT OF REGULATION AND REGULATORY POLICY (2012).

<sup>236</sup> For a review of retrospective analysis see Aldy, *supra* note 234, at 27.

The Obama Administration issued several Executive Orders that recognize the importance of retrospective testing, with three executive orders emphasizing the need for analysis and review of existing federal rules.<sup>237</sup> A recent analysis of retrospective review based on the 2011 OMB report to Congress on federal regulation by Randall Lutter<sup>238</sup> revealed that retrospective analysis was incomplete in the cases discussed in the report. Another report by Joseph Aldy discusses similar results.<sup>239</sup>

Regarding independent agencies, like the SEC and the CFPB, guidelines should be developed requiring outlines of expected retrospective analysis at the time of adoption, as well as designing implementation of regulation to facilitate such analysis.<sup>240</sup> For example, implementation of regulation is a way that resembles a RCT allowing for meaningful analysis post adoption.<sup>241</sup>

### 3. Dynamic Testing

A central issue to the success of consumer financial regulation is the response of the regulated entity. In the past, regulated entities have significantly undermined consumer financial regulation. For example, financial institutions weakened

---

The requirement for retrospective analysis appeared in earlier Executive Orders, such as Reagan's Executive Order No. 12291 (1981) and Clinton's Executive Order No. 12866 (1993). However, the review of regulation was not consistent prior to Obama's Executive Orders.

<sup>237</sup> Executive Order No. 13563 (2011); Executive Order No. 13579 (2011), applying to independent agencies; Executive Order No. 13610 (2012). For a detailed discussion of the different provisions requiring retrospective analysis see Lutter.

<sup>238</sup> Randall Lutter, *Regulatory Policy: What Role for Retrospective Analysis and Review?*, 4 J. COST BENEFIT ANALYSIS 17 (2013).

<sup>239</sup> Aldy, *supra* note 234.

<sup>240</sup> Executive Order No. 13579 relates to independent agencies, encouraging them to expand retrospective analysis. Independent agencies have dedicated some attention to retrospective analysis. *See, e.g.*, SEC Request for Information on Retrospective Review of Existing Regulation, 76 Fed. Reg. 176 (Sept. 12, 2011), available at <http://www.gpo.gov/fdsys/pkg/FR-2011-09-12/pdf/2011-23179.pdf>. Regarding the CFPB, under Section 1022 of Dodd-Frank, defining the CFPB's rule-making authority, the CFPB is instructed to provide retrospective analysis of its regulation: "The Bureau shall conduct an assessment of each significant rule or order adopted by the Bureau under Federal consumer law." 12 U.S.C. § 5511 (Dodd-Frank § 1022.d).

<sup>241</sup> For the proposal the prospective analysis include an outline of future retrospective analysis, see Aldy, *supra* note 234, at 17.

the default for consumer checking account overdraft coverage<sup>242</sup> and have undermined regulation through shifting credit card expenses to non-salient fees.<sup>243</sup>

In many cases disclosure regulation lays down the basic requirements of the disclosure document but allows the industry considerable discretion. This discretion is not always exercised in a way that is optimal for consumers. For example, the SEC summary prospectus Final Rule did not contain a page limitation, but rather stated that its intent was “that funds prepare a concise summary (on the order of three or four pages) that will provide key information”.<sup>244</sup> However, the SEC Guidance Regarding Mutual Fund Enhanced Disclosure from June 2014<sup>245</sup> stated that: “it is not unusual for the staff to review filings with Summary Sections that are longer than ten pages for a single mutual fund and sometimes almost twenty pages in length.”<sup>246</sup>

The SEC summary prospectus rules include additional levels of discretion that may be of more concern, such as the drafting of the principal risks and investment strategies of the summary prospectus, and the general formatting of the document, which is left to the regulated entity to determine. Oehler *et al.* found that in the case of KIID in Europe, disclosures created by issuers and suppliers were evaluated as inadequate in assisting consumers relative to the neutral document developed by the researchers.<sup>247</sup>

Even when the regulators closely dictate the content of the disclosure, such as in the case of TILA and RESPA disclosures, the industry often has control over the context in which consumers receive the disclosure. As discussed above in beginning of this section, the effect of disclosure can be diluted by providing many

---

<sup>242</sup> See Lauren Willis, *When Nudges Fail: Slippery Defaults*, 80 U. CHI. L. REV. 1155 (2013).

<sup>243</sup> See Oren Bar-Gill & Ryan Bubb, *Credit Card Pricing: The Card Act and Beyond*, 97 CORNELL L. REV. 967 (2011). See also Sumit Agarwal, *et al.*, *Regulating Consumer Financial Products: Evidence from Credit Cards*, Q.J. OF ECON. 111 (2015).

<sup>244</sup> SEC Summary Prospectus, *supra* note 67, at 4551.

<sup>245</sup> SECURITIES EXCH. COMM'N, GUIDANCE REGARDING MUTUAL FUND ENHANCED DISCLOSURE (Jun. 2014), available at <http://www.sec.gov/investment/im-guidance-2014-08.pdf>.

<sup>246</sup> *Id.* at 2.

<sup>247</sup> Oehler, Hofer & Wendt, *supra* note 126



other documents or by crowding out their effect with human interaction.<sup>248</sup>

Another concern is that financial disclosures are sometimes designed to make certain costs and features salient that are currently worrisome, however regulated entities may respond by restructuring their fees and shifting charges towards more shrouded costs.<sup>249</sup> A paper by Anagol and Hoikwang Kim<sup>250</sup> shows that a regulatory change allowing closed-end funds to shroud front load fees led to investments being diverted from open-end funds to closed-end funds. Concerns along these lines were raised in the context of the SEC's summary prospectus that does not require funds to include the turnover rate,<sup>251</sup> relating to the percentage of the fund's assets that have changed over the past year resulting in potentially significant transaction costs,<sup>252</sup> in the fees section. This may result in funds increasing the fees charged through this shrouded channel.

In the context of disclosure of conflicts of interest there is evidence that the disclosure may backfire and lead to changes in behavior of advisors.<sup>253</sup> When the disclosure of conflicts of interest itself alters the behavior of the discloser, regulatory testing should not focus solely on consumer responses to the disclosure.

Despite the centrality of the regulated entities' responses to the success of regulation, current testing does not consider these reactions prior to the adoption of disclosures. Therefore, another important element of effective testing of proposed financial

---

<sup>248</sup> See also *supra* Part IV.B., discussing the problems with the development of the "neutral" KIID.

<sup>249</sup> In the credit card context, see Bar-Gill & Bubb, *supra* note 243.

<sup>250</sup> Anagol & Hoikwang-Kim, *supra* note 178.

<sup>251</sup> SEC SUMMARY PROSPECTUS, *supra* note 67, at 30.

<sup>252</sup> Roger Edelen, Richard Evens & Gregory Kadlec, *Scale Effects in Mutual Fund Performance: The Role of Trading Costs*, at 3 (Mar. 17, 2007), available at: [http://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=951367](http://papers.ssrn.com/sol3/papers.cfm?abstract_id=951367).

<sup>253</sup> In an experiment by Cain, Loewenstein, and Moore participants who were "estimators" were rewarded by the proximity of their estimate to the actual amount of coins in a jar. Some participants who were "advisors", who were provided more information than estimators, were paid more when estimators responded with high rather than accurate values. The results showed that advisors that were required to disclose this conflict of interest exaggerated their advice more than those who were not required to disclose. See Cain, Loewenstein & Moore, *supra* note 172.

disclosures or retrospective analysis should be the impact on the regulated entities required to create or provide the disclosures. Since industry responses to regulation may be hard to predict, there is a need to develop a methodology for anticipating industry responses, including dynamic testing, prior to adoption.<sup>254</sup>

Even without experimentation a significant amount can be learned from experience with previously adopted financial disclosures and the industry's response to various requirements. Through retrospective analysis of regulation, regulators can infer when discretion was appropriate and when not. They can also infer the types of regulation that are particularly sensitive to regulatory arbitrage.<sup>255</sup>

Regarding prospective regulation, one way in which this can be tested is by simulating the dilemma that regulated entities might encounter. On the one hand regulated entities presumably wish to abide by the legal requirements in terms of the content and delivery of disclosure. On the other hand regulated entities wish to remain profitable, and if profitability was partially the result of consumers taking loans with higher interest rates or investing in funds with higher fees, this interest will persist. Experiments could involve participants who are likely to face similar dilemmas in the future, such as business school students. Participants could be asked to develop disclosure materials and loan terms that could then be used as the basis for an experiment comparing consumer decisions based on those disclosures relative to disclosures developed by a group with incentives more aligned with consumer welfare.

More extensive experiments could involve participants being given the role of the loan originator or mortgage broker, required to maximize their profits through offering loans to another group of people. Beyond developing the disclosures, participants could be asked to act as brokers whereby determining

---

<sup>254</sup> Willis suggests another way to respond to this problem by requiring firms to demonstrate that their consumers understand products and services they purchase. See Lauren Willis, *The Consumer Financial Protection Bureau and the Quest for Consumer Comprehension* (May 14, 2015), available at [http://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=2606373](http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2606373).

<sup>255</sup> For an initial attempt to analyze when regulation may be undermined by regulated entities see Michael Barr, Sendhil Mullainathan & Eldar Shafir, *Behaviorally Informed Financial Services Regulation*, NEW AM. FOUND. (2008).

the environment in which other participants receive the disclosure. Experiment compensation could be tied to profitability of the loans consumers eventually select.

Mandated disclosures shape financial products and the behavior of financial institutions and intermediaries. Regulators have dedicated very little analysis to industry reactions to mandated disclosure. Consideration in the abstract of these reactions can be insufficient as industry responses may be hard to predict. Instead regulators should expand testing of proposed regulation to test industry responses.

## VI. CONCLUSION

Fundamental changes need to take place in the way financial regulators validate disclosures. The prevalent use of disclosure requires that regulators adequately demonstrate that their mandated disclosures actually assist people in making better financial decisions.

This article has discussed two recent examples of quantitative testing in depth, which demonstrate the significant resources dedicated to consumer testing. Considering the future of testing, many of the issues currently under consideration by the Consumer Financial Protection Bureau and other consumer financial regulators, such as payday lending, credit cards and insurance, also rely heavily on disclosure. Accordingly, we can expect that consumer testing will increasingly be used to adopt and justify these disclosures.

It is therefore worrying that current testing methodologies are on the wrong track and do not fulfill the purpose of testing whether disclosures will improve decision-making. Firstly, regulators have inadequately articulated the purpose of disclosures, and therefore also the purpose of testing. While regulators seem to be concerned with improving consumer decision-making, they confuse improved decisions with improved comprehension of disclosures. Since comprehension and improved decisions often diverge, the current testing methodology is misguided. Second, regulators limit their testing to one type of benefit of disclosure, namely making consumers fully informed, when in fact disclosures are possibly beneficial in other more subtle ways, such as creating general awareness of the need to seek

further advice and information. Even if the purpose of disclosure is to make consumers fully informed, the results from regulatory testing cast doubt on whether this goal is achieved by newly adopted disclosures.

Given these shortcomings I suggest that regulators rethink the design of their tests for validating disclosure. Financial regulators need to reconsider their theory regarding the potential benefit of disclosure and the purpose of testing. They must focus on consumer decision-making by evaluating decisions made based on disclosure, and not only on comprehension, and so provide a more complete picture of the impact of disclosure. The shortcomings of current testing practices have direct implications for other types of disclosure, such as medical decision aids and nutritional value labels, which are often also tested through indirect comprehension tests rather than decision-making tests.