

Spring 2014

# Killing Conscience: The Unintended Behavioral Consequences of "Pay for Performance"

Lynn A. Stout  
*Cornell Law School*, ls483@cornell.edu

Follow this and additional works at: <http://scholarship.law.cornell.edu/facpub>

 Part of the [Contracts Commons](#), and the [Corporation and Enterprise Law Commons](#)

---

## Recommended Citation

Lynn A. Stout, "Killing Conscience: The Unintended Behavioral Consequences of 'Pay for Performance'", 39 *Journal of Corporation Law* 525 (2014)

This Article is brought to you for free and open access by the Faculty Scholarship at Scholarship@Cornell Law: A Digital Repository. It has been accepted for inclusion in Cornell Law Faculty Publications by an authorized administrator of Scholarship@Cornell Law: A Digital Repository. For more information, please contact [jmp8@cornell.edu](mailto:jmp8@cornell.edu).

# Killing Conscience: The Unintended Behavioral Consequences of “Pay For Performance”

Lynn A. Stout\*

## Abstract

*Contemporary lawmakers and reformers often argue that ex ante incentive contracts providing for large material rewards are the best and possibly only way to motivate corporate executives and other employees to serve their firms’ interests. This Article offers a critique of the “pay for performance” approach. In particular, it explores why, for a variety of mutually reinforcing reasons, workplaces that rely on ex ante incentive contracts suppress unselfish prosocial behavior (conscience) and promote selfishness and opportunism. The end result may not be more efficient, but more uncooperative, unethical, and illegal employee behavior.*

|   |     |
|---|-----|
| I. INTRODUCTION .....   | 526 |
| II. OPTIMAL CONTRACTING AND THE IDEOLOGY OF INCENTIVES.....                 | 530 |
| A. <i>The Meaning of “Incentives”</i> .....                                 | 531 |
| B. <i>The Rise of Incentive Ideology</i> .....                              | 532 |
| C. <i>Does Incentive Pay Work? Evidence from the Corporate Sector</i> ..... | 534 |
| D. <i>Incentives and the Assumption of Selfishness</i> .....                | 535 |
| III. UNSELFISH PROSOCIAL BEHAVIOR: A PRIMER.....                            | 537 |
| A. <i>Lesson 1: Conscience Exists and Most People Know It</i> .....         | 539 |
| B. <i>Lesson 2: Social Context and the Jekyll–Hyde Syndrome</i> .....       | 541 |
| 1. <i>Instructions from Authority</i> .....                                 | 543 |
| 2. <i>Perceptions of In-Group Membership</i> .....                          | 543 |
| 3. <i>Beliefs About Others’ Prosociality</i> .....                          | 544 |

---

\* Distinguished Professor of Corporate and Business Law, Clarke Business Law Institute, Cornell School of Law. The author would like to thank participants in workshops held at Cornell Law School, the University of Colorado School of Law, Columbia Law School, Georgetown Law School, New York University’s Stern School of Business, the University of Pennsylvania Law School, Queen’s University Centre for Law in the Contemporary Workplace, Syracuse University Law School, the University of St. Thomas, and the University of Texas School of Law, for their helpful comments and suggestions. She is especially grateful for suggestions from Tamara Belinfanti, Bruce Buchanan, Donna Dabney, Michael Dorff, Ron Gilson, John Haidt, Robert Frank, Margaret Heffernan, Calvin Johnson, Miguel Padro, Barry Schwartz, Robert Scott, Jean Tirole, and Deborah Weiss. She apologizes to anyone whom she might have forgotten to include. She would also like to express her appreciation for the invaluable research assistance provided by Matthew Morrison.

4. *Magnitude of Benefits to Others*..... 545

5. *Conclusion: Understanding the Jekyll–Hyde Syndrome*..... 545

C. *Lesson 3: Prosociality and Personal Cost* ..... 546

D. *Lesson 4: The Role of Character* ..... 547

IV. THE UNINTENDED BEHAVIORAL CONSEQUENCES OF PAY FOR PERFORMANCE ..... 548

A. *Relational Contracts and Contractual Incompleteness* ..... 548

B. *Conscience As a Solution to Contractual Incompleteness* ..... 551

C. *Social Context and Crowding Out* ..... 552

D. *Conscience, Temptation, and Cognitive Dissonance*..... 554

E. *Selection Bias and the Question of Character*..... 555

V. CONCLUSION: ALTERNATIVES TO PAY FOR PERFORMANCE ..... 557

I. INTRODUCTION

In 2009, the Swiss bank UBS agreed to pay the U.S. government \$780,000,000 to settle charges that it had orchestrated a massive scheme to help wealthy Americans evade U.S. tax laws.<sup>1</sup> The UBS case—the largest tax fraud investigation in history<sup>2</sup>—began with the arrest of a single banker named Bradley Birkenfeld. Birkenfeld was one of several UBS employees who had repeatedly helped clients evade U.S. taxes, and he agreed to cooperate with the Justice Department in return for being allowed to plead guilty to a single count of conspiracy to defraud the U.S. government. When the judge who heard Birkenfeld’s plea asked him why he had participated in the scheme when he knew he was breaking the law, the 43-year-old banker replied: “I was incentivized to do this business.”<sup>3</sup>

The UBS tax scandal is only one of several recent high-profile cases in which incentive contracts supposedly tempted employees into opportunistic and even illegal behavior. In April of 2013, the nation was treated to the spectacle of dozens of Atlanta public school teachers and educators surrendering themselves to authorities after being indicted for allegedly conspiring to alter student test scores to earn cash bonuses.<sup>4</sup>

---

1. William P. Barrett & Janet Novack, *UBS Agrees to Pay \$780 Million*, FORBES (Feb. 18, 2009, 7:00 PM), [http://www.forbes.com/2009/02/18/ubs-fraud-offshore-personal-finance\\_ubs.html](http://www.forbes.com/2009/02/18/ubs-fraud-offshore-personal-finance_ubs.html).

2. Michael Rubinkam, *UBS Tax Evasion Whistle-Blower Reports to Federal Prison*, USA TODAY (Jan. 8, 2010, 5:23 PM), [http://usatoday30.usatoday.com/money/perfi/taxes/2010-01-08-ubs-tax-evasion-informant-prison\\_N.htm](http://usatoday30.usatoday.com/money/perfi/taxes/2010-01-08-ubs-tax-evasion-informant-prison_N.htm). Birkenfeld had approached the Department of Justice in 2007 with information about UBS’s activities after the IRS adopted a new tax-whistleblower reward program, but he was subsequently indicted when prosecutors concluded that he had not been fully forthcoming about his own involvement. *Id.* Birkenfeld received and served most of a 40-month prison sentence, but he also eventually was awarded \$104 million under the whistleblower reward program. M.V., *Whistleblowing: Birkenfeld’s Bonanza*, THE ECONOMIST (Sept. 11, 2012, 6:50 PM), <http://www.economist.com/node/21562860/print>.

3. Evan Perez, *Guilty Plea By Ex-Banker Likely to Aid Probe of UBS*, WALL ST. J. (June 20, 2008, 12:01 AM), [online.wsj.com/news/articles/SB121389134942588841](http://online.wsj.com/news/articles/SB121389134942588841).

4. David Beasley, *Atlanta Educators Surrendering in Cheating Scandal*, REUTERS (Apr. 2, 2013, 6:28

Incentive pay has been blamed for the savings and loan crisis of the late 1980s<sup>5</sup> and the Enron and Worldcom accounting frauds of the late 1990s.<sup>6</sup> Incentive contracts have also been identified as a root cause of the 2008 credit crisis, when they tempted mortgage brokers into approving loans to unqualified buyers<sup>7</sup> and enticed bank executives into embracing risky investing and lending practices.<sup>8</sup>

Despite such object lessons, public enthusiasm for pay for performance is still growing.<sup>9</sup> As a Wall Street Journal article put it, “Incentive programs are spreading like . . . kudzu.”<sup>10</sup> When things go wrong, incentive pay advocates typically argue that the problem lay not in using incentives but in using *poorly designed* incentives.<sup>11</sup> If we are sufficiently careful in measuring and rewarding individual performance, the “optimal contracting” argument goes, pay for performance schemes harness the forces of greed and self-interest to promote greater efficiency and better economic performance.

This Article challenges conventional wisdom by arguing that pay for performance strategies, by their very nature, often prove counterproductive and even disastrous “solutions” to complex social problems like corporate scandals and failing schools. Optimal contracting theory dominates the ongoing debate over executive compensation and is seeping into other policy discussions as well because reformers believe that even when we can’t do much else, we can at least “get the incentives right.” The assumption seems to be that *ex ante* incentives might help, and can’t possibly hurt. But pay-for-performance schemes *can* hurt.

Optimal contracting theory relies on a *homo economicus* model of purely self-interested behavior that predicts that legally enforceable, predetermined material

PM), <http://www.reuters.com/article/2013/04/02/us-usa-schools-atlanta-idUSBRE9310YP20130402>.

5. *Executive Compensation: How Much Is Too Much?: Hearing Before the H. Comm. on Oversight and Gov’t Reform*, 111th Cong. 238–39 (2009) (statement of William K. Black, Assoc. Professor of Economics and Law, Univ. of Missouri-Kansas City) [hereinafter *Executive Compensation Hearing*], available at <http://www.gpo.gov/fdsys/pkg/CHRG-111hhrg54553/pdf/CHRG-111hhrg54553.pdf>.

6. Margaret M. Blair, *Shareholder Value, Corporate Governance, and Corporate Performance: A Post-Enron Reassessment of the Conventional Wisdom*, in *CORPORATE GOVERNANCE AND CAPITAL FLOWS IN A GLOBAL ECONOMY* 53, 61–62 (Peter K. Cornelius & Bruce Kogut eds., 2003) (discussing role of stock options in Worldcom and Enron scandals); William W. Bratton, *Enron and the Dark Side of Shareholder Value*, 76 TUL. L. REV. 1275, 1327–28 (2003) (discussing role of incentives in Enron scandal).

7. See FIN. CRISIS INQUIRY COMM’N, *THE FINANCIAL CRISIS INQUIRY REPORT* 90 (2011) (discussing mortgage brokers’ incentive compensation).

8. *Id.* at 17 (discussing how executive compensation based on short-term gains increased riskiness of financial firms); see generally BD. OF GOVERNORS OF THE FED. RESERVE SYS., *INCENTIVE COMPENSATION PRACTICES: A REPORT ON THE HORIZONTAL REVIEW OF PRACTICES AT LARGE BANKING ORGANIZATIONS* 1 (2011) (“Risk-taking incentives provided by incentive compensation arrangements in the financial services industry were a contributing factor to the financial crisis that began in 2007.”).

9. RUTH W. GRANT, *STRINGS ATTACHED: UNTANGLING THE ETHICS OF INCENTIVES* 2 (2012) (“Increasingly in the modern world, incentives are becoming the tool we reach for when we wish to bring about change.”); see also Barry Schwartz, *Practical Wisdom and Organizations*, 31 RES. ORG. BEHAV. 3, 4 (2011), available at <http://www.swarthmore.edu/Documents/academics/psychology/Schwartz9o20ROB.2011-PracWis Organizations.pdf> (discussing modern tendency to rely on incentives to change behavior).

10. Eric Felten, *Age of Incentives: Paying Big Bucks for Puny Results*, WALL ST. J. (June 18, 2010, 12:01 AM), <http://online.wsj.com/news/articles/SB10001424052748704009804575308710787390320>.

11. See, e.g., Kevin J. Murphy, *Executive Compensation*, in 3B *HANDBOOK OF LABOR ECONOMICS* 2485 (Orley Ashenfelter & David Card eds., 1999) (describing problems as “consequences of poorly designed pay programs”).

incentives are the best, and possibly only, tool available to motivate an agent to do what the principal wants the agent to do. This behavioral model, while elegant and powerful, is also dangerously misleading. Extensive empirical evidence demonstrates that when employment contracts are incomplete (as all contracts must be to greater or lesser degrees), employers can get better results by emphasizing “internal” incentives, especially the internal force laymen call *conscience*. What’s more, conscience and self-interest often function as substitutes rather than complements. Ex ante incentive contracts—even well-designed ones—typically create “psychopathogenic” pressures that suppress or snuff out conscience. The result may not be more efficient agent behavior but more opportunistic, unethical, and illegal agent behavior.

By showing how incentive contracts suppress conscience, this Article does not suggest that pay itself (that is, some form of compensation) is unnecessary. Few employees are willing to work very long or very hard for free. Nor does this Article claim that incentive pay is always counterproductive. There may be agency tasks where ex ante incentive contracts perform quite well, despite their negative effects on conscience.

This Article does advance two counterintuitive claims. The first is that ex ante incentives are not always the only, or the best, means available for motivating employees. Extensive behavioral evidence demonstrates that with the right combinations of social cues and discretionary ex post rewards, many agents will work harder and more honestly than formal incentive contracts can induce them to. The second point is that for many complex tasks that principals might want agents to perform in the business world and elsewhere, employing ex ante incentives can be dangerous because this strategy suppresses conscience and promotes selfishness for a variety of reasons. The natural implication of the two points is that instead of relying on ex ante incentives, corporations and other employers often might do better to rely on ex post, trust-based compensation arrangements that recognize both the principal’s and the agent’s capacity to reciprocate prosocial behavior.

Part II of this Article begins by describing the optimal contracting approach and its history. It shows how when optimal contracting theorists speak of “incentives,” they are not using the word in a broad sense as a synonym for “motivations,” as in, “My love for my child gives me incentive to take her to the pediatrician.” Rather, they are speaking of predetermined financial or material rewards that are formally negotiated and specified ex ante. Part II shows how the notion that ex ante incentives provide the best, and possibly only, way to channel human behavior—an idea that implicitly assumes people are opportunistic and selfish—has exercised increasing influence in private employment markets and regulatory policy. As an important example, Part II describes the 1993 amendment of the tax code to encourage publicly held companies to use high-powered ex ante incentive schemes to compensate executives. Part II then explores how, despite the enthusiastic embrace of incentive pay by academics, policymakers, and reformers, there is remarkably little empirical evidence to support the claim that incentive contracts actually produce better results, in the business world or elsewhere. Yet rather than question the efficacy of incentive compensation schemes, many experts continue to insist the solution is simply to use more and better ones.

Part III explores some reasons why incentive pay often backfires. In particular, Part III surveys what behavioral science in general, and experimental gaming in particular, has revealed about the empirical phenomenon of unselfish prosocial behavior (*conscience*).

Contrary to the assumption of opportunistic selfishness that underlies optimal contracting theory, real people often act in an unselfish, prosocial fashion. In lay terms, they act as if they have a conscience that spurs them, at least sometimes, to sacrifice their own material payoffs to help or avoid harming others and to follow ethical rules. While different individuals show different proclivities toward conscientiousness, the data demonstrates that conscience is neither rare nor quirky. Almost anyone other than a clinical psychopath is likely to act unselfishly when certain social cues support unselfishness and the personal cost of acting unselfishly is not too high. Part III uses these findings to propose a simple model of conscience that offers four useful lessons for optimal contracting theory. First, conscience (unselfish prosocial behavior) exists and is a common behavioral phenomenon. Second, conscientious behavior seems triggered primarily by important social cues, especially instructions from authority, perceptions of common group membership, beliefs about others' prosocial behavior, and perceptions of benefits to others. Third, even when the social cues support conscience, it can disappear if the personal cost of acting conscientiously becomes too great. Fourth, individuals vary in their willingness and inclinations toward unselfish prosocial behavior.

Part IV explores what the findings described in Part III imply about the effects of high-powered incentive schemes. In particular, through at least three different but mutually reinforcing mechanisms, incentive contracts tend to suppress conscience and encourage opportunistic and even illegal behavior that conscience otherwise would keep in check. First, incentive schemes frame social context in a fashion that encourages people to conclude purely selfish behavior is both appropriate and expected. As a result, pay-for-performance rules "crowd out" concern for others' welfare and for ethical rules, making the assumption of selfish opportunism a self-fulfilling prophecy. Second, the possibility of reaping large personal rewards from incentive schemes tempts people to cut ethical and legal corners, and once an individual succumbs to temptation, future lapses become more likely. The result can be a downward spiral into opportunistic and unlawful behavior. Third, industries and firms that emphasize incentive pay tend to attract individuals who, even if they are not clinical psychopaths, nevertheless are more inclined toward selfish and opportunistic behavior than the average. Once such relatively selfish actors come to dominate a workplace, less-selfish employees leave, and the employees who remain start acting in a more purely self-interested and opportunistic fashion.

Part V concludes by considering some implications for contemporary law and policy. The pay-for-performance approach dominates compensation practices in the executive suite today. It is also gaining popularity in our nation's schools, newsrooms, and medical centers. The scientific evidence suggests this may be a dangerous development. It may be counterproductive to compensate people primarily through large *ex ante* financial incentives. Sometimes, perhaps often, employers get better results by adopting exactly the opposite approach: emphasizing rewards that are modest, nonmonetary, and awarded *ex post*. This reality has important implications not only for the current debate over regulating executive compensation, but for other pressing issues of law and public policy as well.

## II. OPTIMAL CONTRACTING AND THE IDEOLOGY OF INCENTIVES

Economists and legal scholars have been studying the problem of how to best compensate corporate executives for decades.<sup>12</sup> From the beginning, the academic literature on executive compensation has typically analyzed the problem from an “optimal contracting” perspective.<sup>13</sup> Optimal contracting theory views the task of setting an executive’s compensation (or any employee’s or agent’s compensation) as a version of what economists call the agency cost problem.<sup>14</sup>

Economic theory predicts that “agency costs” arise whenever a rational and selfish principal hires a rational and selfish agent to accomplish something the principal wants done. Because the agent is selfish, if left to his own devices, he may not do what the principal wants done. To use the words of Michael Jensen and William Meckling (two of the earliest and most influential writers in the executive compensation debate): “If both parties to the relationship are utility maximizers there is good reason to believe that the agent will not always act in the best interests of the principal.”<sup>15</sup> By the same token, if the principal is purely selfish, she will do whatever she can to minimize any payment she makes to the agent. The solution for both parties is to draft an “efficient” or “optimal” contract that legally obligates the principal to pay the agent specific compensation that is tied *ex ante* to specific observable measures of the agent’s performance.

In the executive compensation debate, the agency cost problem is typically framed as a problem of getting self-interested corporate directors and executives to serve the interests of the firm’s shareholders. As Margaret Blair has put it, “[T]he conventional wisdom has been that directors and managers of companies will always make decisions in ways that serve their own personal interests unless . . . given very strong incentives . . . .”<sup>16</sup> Thus, the problem of executive compensation is framed as a problem of designing incentives that motivate executives to serve shareholders’ interests.<sup>17</sup>

---

12. See, e.g., Michael C. Jensen & Kevin J. Murphy, *CEO Incentives—It’s Not How Much You Pay, But How*, HARV. BUS. REV., May 1990, at 138; Bengt Holmstrom & Paul Milgrom, *The Firm As An Incentive System*, 84 AM. ECON. REV. 972 (1994); LUCIAN BEBCHUK & JESSE FRIED, *PAY WITHOUT PERFORMANCE: THE UNFULFILLED PROMISE OF EXECUTIVE COMPENSATION* (2004); see generally Iman Anabtawi, *Explaining Pay Without Performance: The Tournament Alternative*, 54 EMORY L.J. 1557 (2005).

13. Anabtawi, *supra* note 12, at 1561 (“The optimal contracting model underlies most scholarship in the area of executive compensation.”).

14. See generally Michael C. Jensen & William H. Meckling, *Theory of the Firm: Managerial Behavior, Agency Costs and Ownership Structure*, 3 J. FIN. ECON. 305 (1976) (describing the agency cost problem in firms).

15. *Id.* at 308.

16. Blair, *supra* note 6, at 60.

17. Elsewhere, Margaret Blair and I have pointed out at length that directors and executives owe duties not only to shareholders but also to the firm as a legal entity, suggesting there is reason to question whether directors are agents of shareholders or of the firm itself. See Margaret M. Blair & Lynn A. Stout, *A Team Production Theory of Corporate Law*, 85 VA. L. REV. 246, 292–309 (1999) (discussing the nature of directors’ duties). See generally LYNN STOUT, *THE SHAREHOLDER VALUE MYTH: HOW PUTTING SHAREHOLDERS FIRST HARMS INVESTORS, CORPORATIONS, AND THE PUBLIC* (2012) (arguing at length why directors should not be considered agents primarily of shareholders but rather of the firm itself).

### A. The Meaning of "Incentives"

It is important to understand exactly what executive compensation experts mean when they talk about incentives. Laypersons often use the word broadly as a synonym for motivation, as in, "My guilt gives me incentive to call my mother." But optimal contracting theorists typically do not concern themselves with internal, subjective motivations like guilt, love, or pride. In optimal contracting scholarship, the word "incentive" refers specifically to external punishments or rewards that share three important characteristics: (1) They are monetary or material in nature; (2) They are of a significant size; and (3) They are contractually predetermined, set in advance according to some *ex ante* algorithm or formula. Let us consider each characteristic in turn, as each is an important element of the optimal contracting approach that contributes to its limitations.

Addressing first the element of monetary value, the word "incentive" must be confined to monetary or material rewards, or optimal contracting theory—indeed economic analysis generally—loses its intellectual content. Using the word "incentive" to refer to anything that motivates behavior reduces economic theory to a tautology. (If economics is based on the principle that "people respond to incentives," and incentives then are defined as "anything people respond to," the logic becomes circular.)<sup>18</sup> Moreover, only monetary, or at least material, incentives lend themselves to formal incentive contracting. It is relatively easy to enforce a contract that says "you will get a million stock options at an exercise price of \$30 per share." It is much harder, and perhaps impossible, to enforce a contract that provides "you will be loved, honored, and esteemed." Pay for performance advocates are really advocating "pay money, or some other good with an ascertainable market value, for performance."

Second, as the phrase "high-powered incentives" implies, optimal contracting theory does not object to, and even embraces, very large incentive payments. After all, the larger the payment, the more it "incentivizes" the agent to perform.<sup>19</sup> Conversely, nominal or token rewards have little or no importance in the theory.

Third and perhaps most important, optimal contracting theory assumes that the rules for determining what exactly the agent must do to earn his or her pay, and for deciding the form and magnitude of the agent's pay, must be objective and must be agreed upon and specified in advance. *Ex ante* agreement to an objective performance goal is essential because optimal contract theory, like other theories that rely on the *homo economicus* model, leaves no room for trust. No rational and purely selfish agent would be so foolish as to rely on an employment contract that provides, "if you do a good job as we see it, we'll reward you with a bonus we think appropriate." Similarly, no rational and selfish principal would promise an executive "we'll give you a million dollar salary and trust you to do the best you can." Optimal contracting theory assumes that formal contracts can control the behavior of agents and principals only when the contract terms are

---

18. See STEVE E. LANDSBURG, *THE ARMCHAIR ECONOMIST: ECONOMICS AND EVERYDAY LIFE* 3 (1993) ("Most of economics can be summarized in four words: 'People respond to incentives.'").

19. See generally Holmstrom & Milgrom, *supra* note 12 (discussing high-powered and low-powered incentives); see, e.g., BEBCHUK & FRIED, *supra* note 12, at 8 (advocating for large payments when justified by results); Jensen & Murphy, *supra* note 12, at 3–4 (arguing for large pay packages that provide high-powered incentives).



objective, enforceable, and clearly specified ex ante.

### B. The Rise of Incentive Ideology

Judged by these standards, the methods that Corporate America used to compensate executives during the “managerialist” era of the 1920s through 1980s were hopelessly backward and inefficient.<sup>20</sup> Before optimal contracting theory gained influence, business executives were typically paid primarily with fixed salaries and the occasional modest bonus, both adjusted ex post on the basis of subjective criteria.<sup>21</sup> (“You did a great job this past year; we’re giving you a bonus and a raise.”) Nonmonetary rewards were coveted and common. (“You’ve earned a key to the executive washroom.”) Executive pay, though hardly stingy, was relatively modest and stable. In the early 1980s, the top three executives of the largest 50 U.S. corporations earned average inflation-adjusted compensation of approximately \$1 million annually—only slightly more than the top three executives of comparable companies earned in 1940. By 2000, this figure had quadrupled to \$4 million.<sup>22</sup>

Despite this, American executives did well for investors in the days before pay for performance. Public companies run by executives who were paid fixed salaries and modest bonuses provided investors with significant positive returns. For example, between 1950 (the year the Standard & Poor’s 500 Index was first published) and 1990, the Index produced inflation-adjusted total returns that averaged more than ten percent each decade.<sup>23</sup>

In the early 1990s, however, the idea of incentive pay captured the hearts and minds of reformers and business leaders alike. This enthusiasm for the optimal contracting approach was part of a broader social trend, the rise of law and economics.<sup>24</sup> As economic analysis became increasingly influential in legal and policy discussions, so did optimal contracting theory in executive compensation discussions.<sup>25</sup> The idea of pay for performance also benefited from the fact it seemed so obvious—people enjoy having

---

20. See INST. FOR GOVERNANCE, PAY FOR VALUE: CUTTING THE GORDIAN KNOT OF EXECUTIVE COMPENSATION 33 (2012) (describing managerialist era and its compensation practices); Tamara C. Belinfanti, *Beyond Economics in Pay for Performance*, 41 HOFSTRA L. REV. 91, 103 (2012) (citation omitted) (“Prior to 1993, corporations mostly compensated executives with fixed salaries and discretionary bonuses.”).

21. INST. FOR GOVERNANCE, *supra* note 20, at 19 fig.4 (showing that salary and bonus payments accounted for 74–99% of CEO pay in the 50 largest U.S. companies from the 1940s through the 1980s, falling to 40% of pay by the mid-2000s, while stock options, and long-term incentive plans rose to account for 60% of CEO pay).

22. Carola Frydman & Raven E. Saks, *Executive Compensation: A New View from a Long-Term Perspective, 1936–2005*, 23 REV. FIN. STUD. 2099, 2107 fig.1 (2007), available at [http://web.mit.edu/frydman/www/trends\\_rfs2010.pdf](http://web.mit.edu/frydman/www/trends_rfs2010.pdf).

23. This figure is calculated by taking the average of the annual returns for the four decades of the 1950s, 1960s, 1970s, and 1980s. See *S&P 500: Total and Inflation-Adjusted Historical Returns*, SIMPLE STOCK INVESTING, <http://www.simplestockinvesting.com/SP500-historical-real-total-returns.htm> (last visited Mar. 7, 2014).

24. In a 2008 study of conservative trends in legal thought, scholar Steven Teles described the law and economics movement as “the most successful intellectual movement in the law of the past thirty years.” STEVEN M. TELES, *THE RISE OF THE CONSERVATIVE LEGAL MOVEMENT: THE BATTLE FOR CONTROL OF THE LAW* 216 (2008).

25. Cf. Belinfanti, *supra* note 20, at 110–17 (tracing pay-for-performance ideology to economists).

money, so why not assume the prospect of having more money would necessarily motivate them to perform better? The end result is that today, the idea that executives can only be trusted to work hard and honestly if their pay is somehow tied to an objective performance metric has been accepted by a generation of corporate experts as a truth so obvious it does not need further examination.<sup>26</sup> As Michael Dorff puts it, “Questioning pay for performance is rather like questioning gravity.”<sup>27</sup>

The ideology of incentives has directly influenced the law. One of the clearest examples is in the U.S. tax code. In 1990, economists Michael Jensen and Kevin Murphy published an influential article in the *Harvard Business Review* calling for companies to tie their executives’ pay to objective metrics.<sup>28</sup> Only a few years later, the U.S. Congress passed a major revision of the Internal Revenue Code to encourage public companies to do just that.<sup>29</sup> Internal Revenue Code Section 162(m) provides that public corporations cannot deduct any annual compensation in excess of \$1 million that is paid to their top five executives unless that compensation is tied to an objective corporate performance metric.<sup>30</sup> Section 162(m) accordingly requires corporations that seek to minimize their tax burdens to adopt incentive pay schemes for their most highly paid executives.<sup>31</sup>

Of course, Section 162(m) has proven an utter failure when it comes to reining in the size of executive pay.<sup>32</sup> In the wake of Section 162(m)’s adoption, executive pay at public companies increased dramatically; even adjusting for inflation, total median annual compensation for CEOs of large public companies increased from about \$2.6 million annually in 1993 to more than \$14 million by 2000.<sup>33</sup> However, Section 162(m) has been quite successful in changing the way public companies compensate their CEOs and other executives. The years since 1993 have seen a seismic shift in the compensation practices of American business corporations, to the point where incentive pay now provides the bulk of compensation for top executives. In 1993, the percentage of CEO compensation attributable to incentive pay was only 35%.<sup>34</sup> Today this figure has risen to 85%.<sup>35</sup>

---

26. LYNN STOUT, CULTIVATING CONSCIENCE: HOW GOOD LAWS MAKE GOOD PEOPLE 42 (2011) (discussing acceptance of incentive ideology).

27. MICHAEL B. DORFF, INDISPENSABLE AND OTHER MYTHS: WHY THE CEO PAY EXPERIMENT FAILED AND HOW TO FIX IT (forthcoming 2014) (manuscript at 126) (on file with author).

28. See generally Jensen & Murphy, *supra* note 12, at 138; Michael C. Jensen & Kevin J. Murphy, *Remuneration: Where We’ve Been, How We Got to Here, What are the Problems, and How to Fix Them* (European Corporate Governance Inst., Working Paper No. 44/2004, 2004) [hereinafter *Remuneration*], available at <http://www2.gsu.edu/~wwwseh/Remuneration%20Where%20Weve%20Been.pdf> (providing numerous recommendations for “reforming” the system of executive compensation).

29. See Belinfanti, *supra* note 20, at 104 (discussing adoption of I.R.C. § 162(m) (2012) and stating that employees of public corporations cannot deduct remuneration exceeding one million dollars unless it specifically fits into the defined category of “performance-based compensation”).

30. IRC § 162(m).

31. *Id.*

32. See Jeffrey D. Korzenik, *The Tax Code Encourages Big Wall Street Bonuses*, FORBES (Feb. 4, 2009, 3:00 PM), [http://www.forbes.com/2009/02/04/wall-street-bonuses-opinions-contributors\\_0204\\_jeffrey\\_korzenik.html](http://www.forbes.com/2009/02/04/wall-street-bonuses-opinions-contributors_0204_jeffrey_korzenik.html) (describing how I.R.C. § 162(m) has not changed the culture of large Wall Street firms and may in fact contribute to “higher volatility of share prices and stingier dividends.”).

33. *Remuneration*, *supra* note 28, at 31 fig.3; see also INST. FOR GOVERNANCE, *supra* note 20, at 17 fig.2 (showing an upward trend in CEO and other top employees’ median compensation).

34. Belinfanti, *supra* note 20, at 103.

35. *Id.*

*C. Does Incentive Pay Work? Evidence from the Corporate Sector*

Thanks in part to I.R.C. Section 162(m) and other regulatory changes that have encouraged U.S. public corporations to embrace incentive-based pay,<sup>36</sup> we now have two decades' extensive experience with pay for performance in the U.S. corporate sector. What have we learned from this massive field experiment in human motivation?

Researchers have published numerous empirical studies examining how adopting incentive pay plans at individual firms has influenced corporate performance. A few studies have found that certain types of incentive compensation schemes seem to be associated with slightly better stock performance measured over relatively short time periods.<sup>37</sup> Other studies, however, found little or no effect, or even negative effects.<sup>38</sup> Meanwhile, incentive pay has been statistically linked with opportunistic, unethical, and even illegal executive behavior, including earning manipulations, accounting frauds, and excessive risk-taking.<sup>39</sup>

These results have led experts who have surveyed the empirical literature to conclude that it provides little or no support for the claim that incentive plans reliably contribute to better corporate performance.<sup>40</sup> As legal scholar Michael Dorff wrote in his recent book on executive compensation, “[T]here is no empirically demonstrable relationship between firms’ use of performance pay and their success in the marketplace.”<sup>41</sup> Even economist Kevin Murphy—a long-time advocate for incentive pay—has conceded that “although there is a plethora of evidence on dysfunctional consequences of poorly designed pay programs, there is surprisingly little direct evidence

---

36. For example, in 2003, the Securities and Exchange Commission began requiring mutual funds to disclose how they were voting the shares held in their investment portfolios. Disclosure of Proxy Voting Policies, Exchange Act Release Nos. 33-8188, 34-47305, IC-25922 (2003). Funds responded largely by outsourcing their voting decisions to “investor advisory firms,” such as the Institutional Shareholders Services (ISS), which favors pay for performance schemes. DORFF, *supra* note 27 (manuscript at 126) (quoting ISS guidelines stating that “pay-for-performance should be a central tenet in executive compensation philosophy”).

37. See DORFF, *supra* note 27 (manuscript at 127) (describing the studies).

38. See *id.* (manuscript at 128) (describing the studies and noting that many “found that performance pay either has no effect or even hurts corporate results”); *id.* at 135 (“Many studies using different methodologies have failed to find consistent evidence that performance pay significantly improves outcomes for corporations.”). There is an interesting case study in HP’s experimental adoption of pay for performance at several divisions, which was abandoned when “some anti-social behavior began to emerge.” Belinfanti, *supra* note 20, at 134.

39. See DORFF, *supra* note 27 (manuscript at 145–46) (describing studies linking performance-based pay to poor risk controls, earnings management, and accounting restatements); Belinfanti, *supra* note 20, at 107–08 (discussing Morgan Stanley report that found pay for performance linked to “the manipulation of earnings, the externalization of risks, and the use of aggressive accounting practices to inflate a company’s stock price”); Bruno S. Frey & Margit Osterloh, *Yes, Managers Should Be Paid Like Bureaucrats*, 14 J. MGMT. INQUIRY 96, 97 (2005) (discussing studies linking incentive pay to financial fraud and accounting restatements); see generally Jared Harris & Philip Bromiley, *Incentives to Cheat: The Influence of Executive Compensation and Firm Performance on Financial Misrepresentation*, 18 ORG. SCI. 350 (2007) (finding link between incentive pay and misrepresentations); see also Calvin H. Johnson, *Corporate Meltdowns Caused by Compensatory Stock Options*, TAX NOTES 738–40 (May 16, 2011), available at [http://www.4texas.edu/law/faculty/calvinjohnson/meltdown\\_comp.pdf](http://www.4texas.edu/law/faculty/calvinjohnson/meltdown_comp.pdf) (concluding that Section 162(m), by favoring options as compensation, has increased the risk of corporate failures).

40. DORFF, *supra* note 27 (manuscript at 127–28).

41. *Id.* (manuscript at 128).

that higher pay-performance sensitivities [in executive compensation plans] lead to higher stock-price performance."<sup>42</sup>

But the story of pay for performance in modern Corporate America may be more disappointing and disturbing even than is suggested by academic studies of incentive plans at individual companies. To see why, let us look at the question from a higher altitude. Optimal contracting theory predicts that, other things being equal, public companies' dramatic shift toward incentive-based pay after the adoption of I.R.C. Section 162(m) should have significantly improved the performance and profitability of U.S. companies and produced a corresponding increase in investor wealth. If pay for performance were the panacea for poor corporate performance that optimal contracting theory predicts it should be, the adoption of Section 162(m) and the subsequent shift in compensation practices should have dramatically increased investor returns from holding stock in public companies.

Those increased investor returns have been noticeably absent. As noted earlier, the S&P 500 Index saw inflation-adjusted total annual returns averaging more than ten percent over the four decades of the 1950s, 1960s, 1970s, and 1980s.<sup>43</sup> During the 1990s, average returns rose to 14.7%, exactly what one would expect to see after executive compensation was first tied directly to share price.<sup>44</sup> But these gains proved unsustainable. The 2000s have been one of the worst decades for equity investors in history, with inflation-adjusted total returns averaging *negative* 3.4% annually.<sup>45</sup> Meanwhile, even as investor returns have plummeted, executive pay has increased. In 1991, two years before the adoption of Section 162, the average CEO of a large public company received pay approximately 140 times that of the average employee; today the ratio is approximately 300 times.<sup>46</sup> The shift to performance-based pay has also been accompanied by a disturbing outbreak of executive-driven corporate frauds, scandals, and failures at firms like Enron, Worldcom, Countrywide, AIG, Lehman Brothers, Goldman Sachs, and JP Morgan, all of which have or had pay-for-performance programs.<sup>47</sup>

#### *D. Incentives and the Assumption of Selfishness*

Of course, the question of what contributes to performance in individual companies and in the broader corporate sector is inevitably difficult and complex. Any discussion of how the shift to pay for performance has affected American public companies must inevitably remain speculative. Nevertheless, our collective experience with Section 162(m), combined with the unsettling results of academic studies, highlights just how little hard evidence supports the notion that pay-for-performance compensation is the

---

42. Murphy, *supra* note 11, at 2539.

43. SIMPLE STOCK INVESTING, *supra* note 23.

44. *Id.*

45. *Id.* More generally, using Dow Jones' S&P 500 Return Calculator (available at <http://dqydj.net/sp-500-return-calculator>), the author calculated that in the 20 years following the 1993 change to the tax code to encourage "pay for performance" in public companies (January 1994 to December 2013), the inflation-adjusted annual return to holding the S&P 500 and reinvesting dividends was 6.48 percent. This is significantly less than the 7.02 annual return investors enjoyed in the preceding 40 years (January of 1954 to December of 1993).

46. Belinfanti, *supra* note 20, at 107.

47. *Id.*

panacea incentive contracting theory predicts it should be.

Yet, rather than question the wisdom of relying on ex ante incentives, many policymakers and would-be reformers have responded to recent corporate crises and scandals by calling for even more use of them. For example, before the 2008 credit crisis, Harvard law professors Lucian Bebchuk and Jesse Fried were prominent advocates for tying executive pay to share price performance.<sup>48</sup> Post-2008, after incentive pay was identified as contributing to excessive risk-taking in financial firms, Bebchuk and Fried now have shifted to emphasizing tying pay to “long-term” stock performance.<sup>49</sup>

Meanwhile, the ideology of incentive pay has seeped into other important public debates. Experts urged the state of Georgia to adopt performance-based pay for teachers on the theory that “[t]o improve outcomes, the state must try to replicate market incentives.”<sup>50</sup> The result, we have since learned, may have been widespread cheating among Georgia educators seeking to improve their students’ test scores through the simple method of erasing and correcting students’ answers on tests.<sup>51</sup> The U.S. Department of Health and Human Services has launched a series of initiatives to explore using pay-for-performance systems for hospitals, physicians, and nursing homes.<sup>52</sup> When Bloomberg News bought *Business Week* magazine in 2009, Bloomberg’s chief editor announced that the company would start basing writers’ compensation on objective metrics like whether a story’s publication changed stock market prices.<sup>53</sup> Legal scholars have even advocated using performance pay to motivate regulators.<sup>54</sup>

As in the case of corporate executives, the ideology of incentives is being embraced in these areas despite the fact that there is little or no empirical evidence to demonstrate it actually works.<sup>55</sup> Perhaps we may eventually stumble upon a proven formula for using

---

48. See BEBCHUK & FRIED, *supra* note 12 (decrying lack of connection between executive pay and stock price performance).

49. Lucian A. Bebchuk & Jesse Fried, *Paying for Long-Term Performance*, 158 U. PA. L. REV. 1915, 1922–36 (2010). In prior work, Bebchuk and Fried had suggested corporate “myopia” was not a particularly significant problem. See BEBCHUK & FRIED, *supra* note 12, at 214–15 (discussing different methods of compensating corporate managers).

50. Noel D. Campbell & Edward J. Lopez, *Paying Teachers for Advanced Degrees: Evidence on Student Performance from Georgia*, 24 J. PRIVATE ENTERPRISE 33, 35 (2008); see also Victor Lavy, *Using Performance-Based Pay to Improve the Quality of Teachers*, 17 FUTURE OF CHILD. 87, 88–89 (2007) (discussing pay for performance in teachers’ compensation).

51. See *supra* note 4 and accompanying text; *infra* note 117 and accompanying text.

52. Press Release, U.S. Dep’t of Health and Human Servs., Centers for Medicare and Medicaid Services, Medicare “Pay for Performance” Initiatives (Jan. 31, 2005).

53. Stephanie Clifford, *An Uneasy Marriage of the Cultish and the Rumpel*, N.Y. TIMES, Apr. 26, 2010, at B1, available at [www.nytimes.com/2010/04/26/business/media/26bizweek.html?pagewanted=all&\\_r=0](http://www.nytimes.com/2010/04/26/business/media/26bizweek.html?pagewanted=all&_r=0).

54. See, e.g., M. Todd Henderson & Frederick Tung, *Pay for Regulator Performance*, 85 S. CAL. L. REV. 1003, 1032 (2012) (arguing that bank regulators should be compensated via a pay-for-performance approach); cf. Sharon Hannes, *Compensating for Executive Compensation: The Case for Gatekeeper Incentive Pay*, 98 CALIF. L. REV. 385, 368 (2010) (proposing incentive pay for auditors).

55. See David N. Figlio & Lawrence Kenney, *Individual Teacher Incentives and Student Performance*, 91 J. PUB. ECON. 901, 902 (2007) (“[T]here is no U.S. evidence of a positive correlation between individual incentive systems for teachers and student achievement.”); Dale B. Thompson, *The Next Stage of Health Care Reform: Controlling Costs by Paying Health Plans Based on Health Outcomes*, 44 AKRON L. REV. 727, 736 (2011) (“A number of pay-for-performance experiments were tried in the early 2000s. The results were not promising.”). See, e.g., Meredith B. Rosenthal et al., *Early Experience with Pay-for-Performance: From Concept to Practice*, 294 J. AM. MED. ASS’N. 1788, 1793 (2005) (analyzing the results of a pay for performance

financial incentives to motivate optimal performance from business executives—as well as doctors, teachers, journalists, and bureaucrats. There is reason to worry, however, that for many of our most important jobs and industries, the quest to tie pay to performance is quixotic at best, and destructive at worst. This is because optimal contracting theory rests on another, deeply flawed theory: the *homo economicus* theory of rational, selfish behavior.

Like most of economic theory, optimal contracting theory assumes that people are fundamentally selfish actors.<sup>56</sup> Even the most ardent enthusiasts of economics would likely admit there are times people seem to show concern for others and for following ethical rules. But optimal contracting theory assumes such departures from the *homo economicus* model are relatively rare and random. According to the theory, it is safe to presume employers and employees follow whatever course of action maximizes their own payoffs. This presumption is why incentives need to be monetary (or at least material), of significant magnitude, and predetermined *ex ante*. Only under these conditions can an employer rely on a contract to get the best out of a selfish employee, or the employee rely on a contract to make a selfish employer pay. Presumably, without incentive contracts, each side will opportunistically exploit the other.

The remainder of this Article argues that this assumption is the Achilles' heel of optimal contracting theory. In recent years, the *homo economicus* model has come under critique with the rise of “behavioral economics,” a school of economic thought that, rather than simply assuming people act rationally and selfishly, looks to empirical experiments to see how real people actually behave. Most contemporary work in behavioral economics tends to focus on departures from rationality more than on departures from selfishness.<sup>57</sup> However, behavioral science also demonstrates that, just as people often make choices that appear irrational, they also often make choices that seem unselfish and conscientious. We turn next to examine what behavioral science teaches about the phenomenon laymen call “conscience” and experts often call “unselfish prosocial behavior.” As we shall see, it teaches lessons that carry important implications for the wisdom of relying on pay for performance.

### III. UNSELFISH PROSOCIAL BEHAVIOR: A PRIMER

Before beginning, it is important to emphasize that the words “unselfish” and “prosocial” are used here to describe behavior and not emotions. We are talking about

---

study conducted in the health care industry); Meredith B. Rosenthal & Richard G. Frank, *What Is the Empirical Basis for Paying for Quality in Health Care?*, 63 *MED. CARE RES. & REV.* 135, 151–53 (2006) (arguing that pay-for-performance methods of compensation are not effective in the healthcare sector).

56. JOSEPH HENRICH ET AL., *FOUNDATIONS OF HUMAN SOCIALITY: ECONOMIC EXPERIMENTS AND ETHNOGRAPHIC EVIDENCE FROM FIFTEEN SMALL-SCALE SOCIETIES* 8 (2004) (discussing how the *homo economicus* model rests on a “selfishness axiom” that assumes “individuals seek to maximize their own material gains . . . and expect others to do the same”). Sometimes advocates for economic analysis try to soften *homo economicus*' sharp corners by arguing that people seek to maximize not their own material wealth but their “utility,” and may get utility from helping others, following ethical rules, and so forth. I have explained at length elsewhere how this stratagem robs economic analysis of usefulness and reduces it to a tautology with no predictive power. STOUT, *supra* note 26, at 26–27.

57. STOUT, *supra* note 26, at 77–78.

acts, not feelings.<sup>58</sup> Thus, this Article will describe an action as “unselfish,” “prosocial,”<sup>59</sup> or “conscientious” whenever the actor sacrifices time, money, or some other valuable resource to help or to avoid harming others or to follow ethical rules. This definition encompasses acts of active altruism, like running into a burning building to save a stranger. But it also applies to the more common phenomenon of “passive” altruism: declining to exploit others’ trust or vulnerability (e.g., refraining from shaking down schoolchildren for lunch money).

Unselfish prosocial behavior is so omnipresent in American society that it often goes unnoticed. For example, Americans watched with horror when their televisions showed scores of New Orleans residents looting in the wake of Hurricane Katrina. Few stopped to marvel at the tens of thousands of New Orleans residents who were *not* looting. I have explored at length elsewhere the many reasons why we tend not to see others’ unselfish behavior, including the nature of our language, various psychological quirks and biases, and the training offered in today’s colleges and universities.<sup>60</sup> However, another important reason the effects of conscience can be difficult to spot in daily life is that healthy societies tend to use extrinsic incentives to reinforce and promote prosocial behaviors.<sup>61</sup> This makes it hard to conclude with certainty that apparently unselfish behavior is driven at least in part by internal forces (conscience), and not only by fear of negative external consequences. I have never taken lunch money from a kindergartner, but it would be difficult for me to prove to a skeptic that it is conscience, and not fear of arrest and prosecution, that deters me from doing so.

Luckily, we can see the effects of conscience clearly in the experimental laboratory. In the lab, researchers control the environment in which behavior occurs. They can eliminate the external influences and incentives that muddy the waters of everyday life. Over the past half-century, behavioral scientists have taken advantage of this fact to design an ingenious variety of experiments to test what human subjects do in situations where self-interest, as measured by material gains and losses, conflicts with the interests of others. The result is an enormous body of empirical data that tells us a surprising amount about just how, when, and why conscience works. This Part surveys four basic lessons that behavioral science teaches about the nature of conscience.

---

58. When I pass up a convenient opportunity to relieve a kindergartner of her lunch money, it is easy to imagine any number of “selfish” subjective concerns that might motivate my restraint. I may want to avoid the internal pangs of guilt, seek the pleasant buzz of feeling virtuous, or simply avoid the fires of Hell. I might even suffer from an inchoate, irrational fear that, no matter what precautions I take, my misdeed inevitably will be detected. Whatever my subjective emotional state is, my objective behavior remains unselfish, in the sense I have declined an opportunity to make myself materially better off. *See generally* STOUT, *supra* note 26, at 54–55 (discussing the difference between unselfish behavior and unselfish emotions).

59. Acts do not always need to be unselfish to be prosocial. A selfish neurosurgeon who saves a dozen lives a week to pay for her third sports car is acting prosocially, albeit in a self-serving fashion.

60. STOUT, *supra* note 26, at 45–71.

61. Unselfish prosocial behavior is often consistent with legal incentives because a variety of legal rules are designed to promote prosocial behavior (e.g., criminal law, contract law, tort law). Similarly, many of the acts of altruism we observe in daily life occur between people who are acquainted with each other and who operate in the same community (the neighborhood, the workplace, the family). Thus, it is difficult to exclude the possibility that apparently unselfish prosocial behavior is actually motivated by concern for future consequences in the form of reciprocity or reputational loss. *Id.* at 65–66.

*A. Lesson 1: Conscience Exists and Most People Know It*

One of the most useful experiments for studying prosocial behavior is an experiment called a *social dilemma game*. A social dilemma resembles the familiar prisoner's dilemma of game theory. However, where the archetypal prisoner's dilemma involves two people, social dilemmas can be played by more (sometimes quite a few more) players. As in the prisoner's dilemma, each player is asked to choose between a "cooperative" strategy that helps the other players and a "defecting" strategy that maximizes the player's own personal returns. As in the prisoner's dilemma, an individual player always maximizes her personal payoffs by defecting, no matter what the other players do. However, the group gets the greatest aggregate payoff if all of its members cooperate.<sup>62</sup>

Over the past five decades, social scientists have published the results of hundreds of social dilemma experiments played by people of varying ages and backgrounds drawn from different cultures around the world.<sup>63</sup> Many experiments were cleverly designed to exclude any possibility that the players could rationally expect external rewards from choosing cooperation over defection.<sup>64</sup> There is a zero percent probability that a rational, selfish subject would cooperate in such games. Yet real people have a marked propensity to cooperate in social dilemmas.<sup>65</sup> As a rule of thumb, experimenters observe cooperation rates typically averaging about 50%.<sup>66</sup> This remarkable result has endured over nearly a half-century of testing.<sup>67</sup>

62. A common example is a "group contribution game." A group of  $n$  players—let us assume  $n$  is four—is assembled and each player is given an initial stake of, say, \$100. The players are told that they can choose between keeping all their newfound cash or contributing some or all of it to a common investment pool. Players are also told that any money contributed to the pool will be multiplied by some factor greater than one but less than  $n$  (assume the money will be tripled), then redistributed equally among all the players—including an equal share for players who did not contribute. The best individual strategy is to keep the \$100, while also hoping to receive an equal portion of the tripled funds that would result from any of the other players being foolish enough to donate to the common pool. For example, if you keep your \$100 and the other three players contribute theirs, you end up with \$325 (your original \$100 plus \$225 from the common pool). As a result, no rational, selfish player will cooperate, and selfish players walk away with only \$100 each. At the same time, the best group outcome (and the best average individual outcome) requires universal cooperation. If all unselfishly contributed, each would get \$300 back. Thus, the rational pursuit of self-interest in a social dilemma ultimately leaves both the group and its individual members worse off.

63. See, e.g., David Sally, *Conversation and Cooperation in Social Dilemmas: A Meta-Analysis of Experiments from 1958 to 1992*, 7 RATIONALITY & SOC'Y 58, 62–73 (1995) (summarizing over 100 studies done between 1958 and 1992); Robyn M. Dawes & Richard H. Thaler, *Anomalies: Cooperation*, 2 J. ECON. PERSP. 187, 188–92 (1988) (summarizing studies); Robyn M. Dawes et al., *Cooperation for the Benefit of Us—Not Me, or My Conscience*, in BEYOND SELF-INTEREST 97 (Jane J. Mansbridge ed., 1990) (summarizing studies). See also Simon Gächter, *Human Pro-Social Motivation and the Maintenance of Social Order* 8–9, in HANDBOOK ON BEHAVIORAL ECONOMICS AND THE LAW (Eyal Zamir & Doron Teichman eds., forthcoming 2014) (manuscript on file with editors) (discussing a variant of the social dilemma called the public good game).

64. For example, experiments often use subjects who are strangers, who are told they will play the game only once, and who play under anonymous double-blind conditions that ensure their choice of strategy (cooperate or defect) will not be revealed to either the other players or the experimenter.

65. HENRICH ET AL., *supra* note 56, at 5 ("[T]here is no society in which experimental behavior is even roughly consistent with the canonical model of purely self-interested actors . . .").

66. Sally, *supra* note 63, at 62–63.

67. In fact, the very first reported prisoner's dilemma experiment run at the RAND Corporation during the 1950s illustrated this phenomenon. The subjects were two RAND game theorists who had devoted their



Such results demonstrate that unselfish prosocial behavior is common, even endemic. Cooperating subjects in a social dilemma are choosing to serve others' interests rather than just maximizing their own. They are demonstrating a form of conscience, meaning that they are behaving as if they take account of more than just their own personal payoffs in making decisions. What's more, not only is conscience endemic, but most people *know* it is endemic. We can see this by comparing the results of two other common experimental games designed to test prosocial behavior: the *dictator game* and the *ultimatum game*.<sup>68</sup>

A dictator game is quite simple. Two subjects are asked to play and one subject (the dictator) is given a sum of money, say \$100. The dictator is then invited to divide the \$100 any way she chooses between herself and the second player. The second player gets what the dictator offers, no more and no less. (This is why the first player is called the dictator.) Interestingly, the majority of dictators in dictator games share at least some of their \$100 with the second player, despite the fact they receive no external reward for this sacrifice.<sup>69</sup> Dictators in dictator games thus demonstrate some degree of unselfish prosociality (conscience), just as subjects in social dilemmas do.

But the results of dictator games become still more interesting when compared with the results of a third experimental game that has been the subject of numerous studies: the ultimatum game. Like a dictator game, an ultimatum game involves two players. One player, called the "proposer," is given an initial stake of money: again, let us assume \$100. The proposer is then told that she can offer to share any portion she chooses—all, a lot, a little, or nothing—with the second player. The second player, called the "responder," then gets to make a choice of his own. He can accept the proposer's offer, in which case the \$100 will be divided as the proposer suggests. Or the responder can reject the offer, in which case both players get nothing.

It is clear what *homo economicus* would do in an ultimatum game. The proposer would offer the smallest possible amount of money (say, one dollar) and the responder would accept this minimal amount. After all, a dollar is better than nothing and should be accepted. Knowing this, no selfish proposer would offer more. Yet human subjects don't play ultimatum games this way. When real people play ultimatum games, the proposer usually offers the responder a substantial portion of the stake, often half.<sup>70</sup> And if the proposer does *not* do this, the responder frequently rejects the offer.<sup>71</sup>

Revenge is sweet. In an ultimatum game, however, it carries a cost. A responder who rejects any positive offer has made himself worse off in material terms than if he had

---

careers to studying rational selfishness. To the consternation of their colleagues, they showed a hearty willingness to unselfishly cooperate with each other. John Nash, a RAND game theorist who would go on to win a Nobel prize and become the subject of the Sylvia Nasar's biography, *A Beautiful Mind*, mused in a note to his colleagues, "[o]ne would have thought them more rational." SYLVIA NASAR, *THE LIFE OF MATHEMATICAL GENIUS AND NOBEL LAUREATE JOHN NASH* 119 (1998).

68. See generally Colin Camerer & Richard H. Thaler, *Anomalies: Ultimatums, Dictators and Manners*, 9 J. ECON. PERSP. 209 (1995) (describing experiments); Gächter, *supra* note 63, at 6 (same).

69. Camerer & Thaler, *supra* note 68, at 213.

70. See *id.* at 212 (summarizing studies that show a strong pattern of proposers offering 50% to responders); Martin A. Nowak et al., *Fairness Versus Reason in the Ultimatum Game*, 289 SCIENCE 1773 (2000) (concluding that most proposers offer 40–50% of the total sum, while about half of responders reject offers below 30%).

71. Camerer & Thaler, *supra* note 68, at 210 ("Offers of less than 20 percent are frequently rejected.").

accepted. Responders who reject positive offers in ultimatum games that they perceive as “too low” are sacrificing, not to benefit another, but to *harm* her.<sup>72</sup> This behavior is sometimes called spite.<sup>73</sup> Experimental proof of spite is interesting, but it becomes more interesting when the results of ultimatum games are compared with the results of dictator games. Offers in dictator games tend to be smaller than offers in ultimatum games.<sup>74</sup> Dictators share, but on average, they do not share *as much* as proposers in ultimatum games do.

This pattern suggests not only that people typically expect some degree of prosocial behavior from others, but also that they also believe *other* people typically expect some degree of prosocial behavior from others. Such expectations explain both why responders in ultimatum games incur a cost to punish a “too low” offer, and why proposers (apparently anticipating such punishment) offer more in ultimatum games where punishment is possible than in dictator games, where punishment is not possible.

The phenomenon of unselfish prosociality (conscience) thus affects human behavior on at least two levels. At the first level, people sometimes sacrifice to benefit (and in ultimatum games, to harm) others around them. At the second level, people expect that *other people* will sometimes sacrifice to benefit (or harm) others, and they alter their own behavior in reliance on this expectation.<sup>75</sup>

### *B. Lesson 2: Social Context and the Jekyll–Hyde Syndrome*

As we have just seen, experimental games demonstrate that the *homo economicus* model of purely selfish behavior is incomplete and often misleading. But to develop a better model that permits more accurate predictions for human behavior, we need to know a bit more. Most obviously, we need to have some idea about when and why people act unselfishly. To appreciate the nature of the problem, recall the 50% cooperation rate typically observed in social dilemmas.<sup>76</sup> This result supports the claim that people often behave unselfishly. But it also supports the claim that people often behave selfishly. After all, if people always showed concern for others, we would observe 100% cooperation rates. What explains why some people cooperate when others don’t, or why the same

72. There may be other forms of other-regarding behavior as well. People may not have only altruistic revealed preferences (willingness to sacrifice to help others) and spiteful revealed preferences (willingness to sacrifice to harm others), but also relative preferences (willingness to sacrifice to ensure that one enjoys a better position relative to others). See ROBERT H. FRANK, *LUXURY FEVER: WHY MONEY FAILS TO SATISFY IN AN ERA OF EXCESS* 107–21 (1999) (discussing relative preferences). Although relative preferences are important in explaining human behavior, they lie beyond the scope of this Article.

73. Spite involves harming another, but spiteful behavior may benefit third parties if it encourages cooperative behavior within a group. As a result, spite can be described at an evolutionary level as a form of altruism. See generally STOUT, *supra* note 26, at 122–47 (discussing the evolution of prosociality).

74. *Id.* at 89.

75. This second observation can hold true even for individuals who are themselves purely selfish and asocial. Consider the example of a purely selfish Jill who lacks a conscience and enters a contract with a conscientious Jack. Purely selfish Jill might rationally choose to make herself vulnerable to Jack by performing her part of the contract first, if she believes Jack will then unselfishly perform his part of the contract as well. We might call this rational trust in Jack’s conscience. Similarly, selfish Jill might refrain from taking opportunistic advantage of Jack if she believes that Jack would react by spitefully sacrificing to punish her, which we might call rational fear of Jack’s vengeance.

76. *Supra* Part III.A.

person may cooperate at one time and not at another?

Luckily, experimental gaming data again offers insight. In brief, most individuals' decisions whether to behave conscientiously in any particular situation (that is, to sacrifice to help or avoid harming others or to follow ethical rules), depends on something we might call *social context*.

From a purely economic perspective, social dilemma, ultimatum, and dictator games are highly standardized experiments that present subjects with fixed payoff functions determined by the nature of the game itself (social dilemma, ultimatum, or dictator game). But while the economic parameters of the games are fixed, researchers can run these experiments under a wide variety of *noneconomic* conditions. Thus, researchers in some experiments have requested that subjects either cooperate or defect;<sup>77</sup> have grouped subjects according to their tastes for abstract art;<sup>78</sup> have allowed them to exchange surnames;<sup>79</sup> and have raised or lowered the payoffs to other members of the group from choosing cooperation over defection.<sup>80</sup> One recent experiment even examined how subjects behaved when playing social dilemmas in the presence of a dog.<sup>81</sup>

None of these changes in social context change the economic structure of the games. For example, in a social dilemma, a subject always maximizes her personal payoffs by choosing defection over cooperation. Nevertheless, changes in social context produce dramatic changes in observed behavior. In a pioneering meta-survey of over 100 reported social dilemma experiments, David Sally found that researchers were able to elicit cooperation rates ranging from a low of 5% to nearly 97%.<sup>82</sup> To appreciate this astonishing behavioral flexibility, recall that payoffs in a social dilemma are such that rationally selfish players should *always* defect.

Social context appears such a powerful determinant of conscientious behavior that it can trigger near-universal prosociality in some games, and near-universal selfishness in others. Although researchers have identified several different social cues that seem to trigger prosocial behavior, this Article will focus on four in particular that seem especially relevant to understanding the effects of pay-for-performance compensation plans: (1) instructions from authority; (2) perceptions of common "in-group" status; (3) expectations regarding others' selfishness or unselfishness; and (4) magnitude of the benefits to others from one's own unselfish action. Each deserves attention, for each has proven consistently important in triggering prosocial behavior in experimental games,

---

77. See generally Sally, *supra* note 63, at 64 (stating that instructions can affect the decision whether to defect or cooperate).

78. See Brent Simpson, *Social Identity and Cooperation in Social Dilemmas*, 18 *Rationality & Soc.* 443, 448 (2006) (discussing studies).

79. See, e.g., Gary Charness & Uri Gneezy, *What's in a Name? Anonymity and Social Distance in Dictator and Ultimatum Games* (Aug. 16, 2003), [http://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=292857](http://papers.ssrn.com/sol3/papers.cfm?abstract_id=292857) (finding that providing personal information increases cooperation); Gary Charness et al., *Social Distance and Reciprocity: An Internet Experiment* (Nov. 2003), [http://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=312141](http://papers.ssrn.com/sol3/papers.cfm?abstract_id=312141) (analyzing the interactions between people over the internet, and noting that social distance affects prosociality negatively).

80. Sally, *supra* note 63, at 66.

81. *Manager's Best Friend: Animal and Human Behaviour*, *ECONOMIST* (Aug. 12, 2010), <http://www.economist.com/node/116789216> (reporting results of study that found that cooperation rates rose when subjects were asked to play social dilemmas in the presence of a dog).

82. Sally, *supra* note 63, at 62.

and each also maps onto a well-studied and fundamental aspect of human psychology (obedience, in-group bias, imitation, and empathy). Moreover, as we will see in Part IV, each carries important implications for our understanding of the behavioral effects of ex ante incentives.

### *1. Instructions from Authority*

One of the most consistent findings in human psychology is that people tend to do as they are told. For example, in Stanley Milgram's infamous obedience experiments, subjects were told to administer a potentially lethal electric shock to another human being (in reality an actor pretending to be shocked). The vast majority did just that.<sup>83</sup> Of course, from a rational choice perspective, this was hardly surprising. After all, Milgram's subjects were being paid to follow instructions. More interesting, subjects in social dilemma games obey instructions to cooperate even though this means they get *less* money. In his meta-survey, for example, Sally found that giving formal instructions to cooperate raised cooperation rates by 34–40% compared to games where no instructions were given.<sup>84</sup> Conversely, formal instructions to defect increased defection by 20–33%.<sup>85</sup>

If Stanley Milgram's experiments showed us the dark side of obedience, social dilemmas show us a brighter side. People will often follow instructions to harm others. But they also often follow instructions to help or to avoid harming others, even when this requires some personal sacrifice.

### *2. Perceptions of In-Group Membership*

A second social variable that influences whether and to what extent subjects act prosocially in experimental games is whether researchers work to increase or decrease the "social distance" between players in the games.<sup>86</sup> For example, subjects in social dilemmas cooperate more when they are allowed to see or to speak to each other,<sup>87</sup> while players in dictator games are more generous when they know the surnames of their fellow players.<sup>88</sup>

83. See generally Stanley Milgram, *A Behavioral Study of Obedience*, 67 J. ABNORMAL & SOC. PSYCHOL. 371 (1963). Milgram's results are surprising and disturbing only if we in fact expect subjects to act prosocially, confirming that most people have consciences and that we know that most do.

84. Sally, *supra* note 63, at 78.

85. *Id.* People are so sensitive to directions from authority that they change their behavior in response to mere *hints* about what the experimenter desires. In one social dilemma experiment, experimenters observed a 65% cooperation rate when subjects were told they were playing the "Community Game." Lee Ross & Andrew Ward, *Naive Realism in Everyday Life: Implications for Social Conflict and Misunderstanding*, in VALUES AND KNOWLEDGE 103, 106–07 (Edward S. Reed et al. eds., 1996). Among similar subjects told they were playing the "Wall Street Game," cooperation dropped to 30%. *Id.* Similar results have been observed in dictator games, where dictators make larger offers when they are instructed to "divide" their stakes than when the experimenters use the "language of exchange." See Camerer & Thaler, *supra* note 68, at 211–13 (providing an example).

86. See STOUT, *supra* note 26, at 100–02, 146 (discussing the role of in-group perceptions). See also Elizabeth Hoffman et al., *Social Distance and Other-Regarding Behavior in Dictator Games*, 86 AM. ECON. REV. 653, 653–60 (1996).

87. Sally, *supra* note 63, at 76, 78, and 83.

88. Gary Charness & Matthew Rabin, *Understanding Social Preferences with Simple Tests*, 117 Q. J. ECON. 817 (2002).

Although this sort of in-group bias is sometimes associated with relatively immutable characteristics like racial or ethnic identity, perceptions of group membership are flexible and highly manipulable. In the famous “Robbers Cave” experiments organized at an Oklahoma summer camp for boys in the 1950s, sociologist Muzafer Sherif first created animosity and conflict between two otherwise similar groups of boys, and then eased the tension between the groups by forcing them to work together toward common goals.<sup>89</sup> Similarly, experimenters have been able to manipulate cooperation rates in social dilemmas by creating subgroup identities among players.<sup>90</sup>

### 3. Beliefs About Others’ Prosociality

A third social variable that seems to play a key role in eliciting unselfish behavior from subjects in experimental games is beliefs about whether *others* are acting, or would act, unselfishly. Not surprisingly, people tend to imitate what other people do. This includes acting nice when we think others would act nicely, and acting nasty when we think they would act nasty. Thus in dictator game experiments, dictators share more of their loot when they are given information indicating that other dictators in other games chose to share.<sup>91</sup> Similarly, numerous social dilemma studies indicate that subjects’ beliefs about how others are likely to behave strongly influence their own choices. Experimenters have found that subjects who believe that their fellow players in a social dilemma experiment are likely to defect become far more likely to defect themselves. Conversely, players who are led to believe their fellows will cooperate become more likely to choose cooperation.<sup>92</sup>

This last pattern is an especially striking example of how social considerations dominate economic concerns in experimental games, because in a social dilemma, believing one’s fellows are likely to cooperate actually increases the expected economic returns from defecting. Nevertheless, far from discouraging cooperation, a belief that other players are going to cooperate produces more cooperation—exactly the opposite of what the *homo economicus* model predicts.

There are a number of possible explanations for why people tend to imitate others.<sup>93</sup> For example, in the context of some experimental games, a belief that others would cooperate may reinforce in-group perceptions of the type discussed immediately above. Whatever the mechanism, perceptions about others’ prosociality seem to be important

89. See generally MUZAFER SHERIF ET AL., INTERGROUP CONFLICT AND COOPERATION: THE ROBBERS CAVE EXPERIMENT (1961).

90. Sally, *supra* note 63, at 78 (stating that “subgroup identity decreases the probability of cooperation”).

91. Erin Krupka & Roberto A. Weber, *The Focusing and Informational Effects of Norms on Pro-Social Behavior*, 30 J. ECON. PSYCHOL. 307, 313–14 (2009).

92. Scott T. Allison & Norbert L. Kerr, *Group Correspondence Biases and the Provision of Public Goods*, 66 J. PERSONALITY & SOC. PSYCHOL. 688, 688 (1994) (citation omitted) (“Numerous studies have reported that individuals are more likely to cooperate when they expect other group members to cooperate than when they expect others to defect.”); Gächter, *supra* note 63, at 20 (discussing role of conformist behavior in determining cooperation); Toshio Yamagishi, *The Structural Goal/Expectation Theory of Cooperation in Social Dilemmas*, in 3 ADVANCES IN GROUP PROCESSES 51, 64–65 (1986) (discussing experimental findings that “[e]xpectations about other members’ behavior is one of the most important individual factors affecting members’ decisions in social dilemmas”).

93. See generally STOUT, *supra* note 26, at 106–10 (discussing imitation).

triggers for one's own prosocial behavior.<sup>94</sup>

#### 4. *Magnitude of Benefits to Others*

Finally, a fourth social variable that influences behavior in experimental games is the magnitude of the payoffs to others from one's own unselfish behavior. Although this may on first inspection seem an economic variable, it is also a social variable; we are talking about economic returns to *others*, a subject of indifference to *homo economicus*. Real people, however, are more inclined to cooperate when they believe others benefit a lot, and not just a little, from their cooperation.

This has been seen in dictator games, where an experimenter's promise to double or triple the amount the dictator chooses to share has made some dictators so generous that their partners ended up with bigger payoffs than the dictators themselves.<sup>95</sup> Similarly, in explaining cooperation rates, Sally's meta-analysis of social dilemmas concluded that "the size of the loss to the group if strictly self-interested choices are made instead of altruistic ones . . . is important and positive."<sup>96</sup>

Concern about the size of benefits to others may be driven by empathy—the capacity to care about what happens to others, and not only to what happens to ourselves. Neoclassical economic theory does not quite know what to do with empathy. Nevertheless, it is a well-recognized and well-studied psychological phenomenon that may play an important role in determining when we do or do not behave in an unselfish, prosocial fashion.<sup>97</sup>

#### 5. *Conclusion: Understanding the Jekyll–Hyde Syndrome*

Taken as a whole, the experimental gaming data thus offers us a second, potentially very useful, lesson about prosocial behavior. In brief, *most people act as if they have at least two personalities* (or, as an economist might put it, two "revealed preference functions"). One personality is purely selfish. When this personality dominates, we maximize our personal payoffs without regard to how our choices affect others. Most people, however, have a second and more prosocial personality. When our prosocial personality dominates, we take account of others' interests, at least to some extent.

The result somewhat resembles the fictional protagonist of Robert Louis Stevenson's tale, *The Strange Case of Dr. Jekyll and Mr. Hyde*.<sup>98</sup> Sometimes we are caring, conscientious, and considerate of others' welfare (Dr. Jekyll). Sometimes we are selfish and asocial (Mr. Hyde). Which persona dominates in any particular situation seems determined largely by social context. And four of the most important aspects of social context are: instructions from authority; perceptions of in-group identity;

94. Some social scientists call this "generalized reciprocity." However, mutual cooperation in a one-shot, anonymous social dilemma cannot be true reciprocity because there is no rational hope that choosing cooperation could elicit benefits from others in future games. Nor can the recipient in a dictator game reciprocate the dictator's generosity. Thus, imitation seems a better word to describe such behavior.

95. See James Andreoni & John Miller, *Giving According to GARP: An Experimental Test of the Consistency of Preferences for Altruism*, 70 *ECONOMETRICA* 737, 745 (2002).

96. Sally, *supra* note 63, at 79.

97. STOUT, *supra* note 26, at 110–14 (discussing empathy).

98. ROBERT LOUIS STEVENSON, *THE STRANGE CASE OF DR. JEKYLL AND MR. HYDE* (1886).

expectations regarding others' prosociality; and perceived benefits to others.

This is not to say that when the social cues are lined up favorably, people always act prosocially. As will be discussed in greater detail, individuals differ in their proclivities toward conscientious behavior, and a blessedly small minority of psychopaths seem to lack any conscience at all.<sup>99</sup> But as we are about to see, even the most conscientious seem to take account of personal cost in choosing between selfish and unselfish behavior.

### C. Lesson 3: Prosociality and Personal Cost

Experiments have proven that social context plays a vital role in determining when people act in an unselfish, prosocial fashion. But saying that social context matters does not imply that economic context does not. A third fundamental lesson from experimental gaming is that prosocial behavior depends not only on social context, but personal payoffs as well.

Although people behave far more unselfishly than standard economic theory suggests, the supply of unselfish behavior seems (as an economist might put it) to be "downward-sloping." This means that as the cost of acting unselfishly increases, the quantity of unselfish behavior supplied declines. This phenomenon is perhaps most easily observed in social dilemma games. As the personal cost associated with cooperating in a social dilemma rises (that is, as the expected gains from defecting increase), the incidence of cooperation drops significantly. Sally's meta-survey found that doubling the reward from defecting decreased average cooperation rates in social dilemmas by as much as 16%.<sup>100</sup> Similarly, when proposers offer relatively larger shares in an ultimatum game, the likelihood that responders will spitefully reject it decreases.<sup>101</sup>

We seem more inclined to unselfishness when unselfishness is cheap. Conversely, when the cost of conscience is high, we are less inclined to "buy" it. It is important to emphasize this is not the same as saying people are basically selfish. *Any* cooperation in a social dilemma, and *any* sharing in an ultimatum or dictator game, is inconsistent with the *homo economicus* model. But when people indulge in conscience, they keep at least one eye on self-interest in doing so.<sup>102</sup>

This means that if we want to promote conscientious behavior, we need to give conscience breathing room to work. George Washington supposedly said few men have the honor to withstand the highest bidder. He may have been on to something: experimental gaming suggests that if we want people to be good, it is important not to

---

99. See *infra* Part III.D (discussing proclivity toward conscientious behavior and the role of character in altruism).

100. Sally, *supra* note 63, at 75.

101. See *supra* notes 70–71 and accompanying text. Although this pattern may be driven in part by responders' perceptions that proposers who offer larger shares are behaving more "fairly" and do not deserve punishment, it is also consistent with fact that as the size of the proposer's offer increases, so does the personal cost of spitefully rejecting the offer.

102. This does not mean unselfish behavior is economically unimportant. Acts of unselfishness that cost the unselfish actor relatively little can provide much larger benefits to others. (Anyone who has had a computer or wallet stolen can appreciate that the costs they would have avoided if these items had not been stolen outweigh the benefit to the thief.) Summed up by many different individuals and for many different social interactions, the total gains from such small acts of altruism can be enormous. Thus, even a limited human capacity for unselfish action generates enormous benefits over long periods and for large populations.

tempt them too much to be bad. As we shall see in Part IV, this carries important implications for the modern ideology of incentives.

#### *D. Lesson 4: The Role of Character*

Finally, we turn to a fourth important lesson to be learned from behavioral experiments: although almost everyone is capable of unselfish action, a very small percentage of the population seems not to be, and the rest of us vary in our inclinations toward unselfishness. In other words, while unselfish behavior is determined in large part by social context and personal cost, it is also related to what laymen might call *character*.

The most obvious example can be found in psychopaths. Psychopaths (sometimes called sociopaths) are relatively rare individuals who, for reasons of nature or nurture, seem incapable of acting unselfishly or showing empathy for others. (Not to put too fine a point on it, it can be argued that *homo economicus* is a psychopath.) Luckily, only about one to three percent of the population is estimated to suffer “antisocial personality disorder” (the formal psychiatric label for psychopathy), and many of those individuals are safely confined in prison.<sup>103</sup>

The rest of us are capable of acting unselfishly, at least in the right circumstances. Recall that experimenters have observed cooperation rates of over 97% in some social dilemmas, and sharing rates of 100% have been observed in some dictator games (presumably dictator games without any psychopathic subjects).<sup>104</sup> When the stars are aligned—when social context supports unselfishness and the personal cost of acting unselfishly is not too high—conscience seems a near-universal behavioral phenomenon.

But in real life, the stars are not always aligned. Sometimes social context is ambiguous, and sometimes large temptations raise their heads. In ambiguous or tempting circumstances, different individuals show different propensities to act conscientiously. Gender seems to play a role in some experimental games, as does religion, although both variables have only modest and quirky effects.<sup>105</sup> Another significant demographic variable may be age. Prosocial behavior in games increases throughout childhood and young adulthood, and some evidence exists that the process of becoming more prosocial continues with age<sup>106</sup> (stereotypes of grumpy old men to the contrary).

But in addition to such relatively weak demographic variables, intriguing evidence suggests that one’s proclivity toward prosociality—one’s “character”—may be in large part a product of one’s experience.

In 2004, a consortium of behavioral scientists published the results of a large study of social dilemma, ultimatum, and dictator games played by subjects from 15 small, non-Western hunting, herding, fishing, and farming cultures around the globe.<sup>107</sup> The consortium found that people of all ages, genders, and backgrounds—Machiguenga subsistence farmers from the rainforests of South America, Torguud nomads in Mongolia, Lamalera whale-hunters in Indonesia—routinely behaved in an unselfish,

---

103. STOUT, *supra* note 26, at 47–48.

104. *Id.* at 98.

105. *Id.* at 100.

106. HENRICH ET AL., *supra* note 56, at 5.

107. *Id.* at 10.



prosocial fashion in playing the games.<sup>108</sup> As the researchers put it, “[T]here is no society in which experimental behavior is even roughly consistent with the canonical model of purely self-interested actors.”<sup>109</sup>

Nevertheless, there were clear differences between cultures. For example, Machiguenga on average contributed 22% in social dilemmas, while the more generous Orma cattle-herders of Kenya contributed 58%.<sup>110</sup> The researchers also found that individual demographic variables—such as gender and wealth—did a poor job of predicting behavior.<sup>111</sup> Rather, behavior seemed driven by social experiences, and especially by whether the culture was one in which people frequently engaged in market transactions with strangers (like hiring themselves out for wages) and whether economic production required people to cooperate with non-kin (whale-hunters necessarily cooperate a lot, while slash-and-burn subsistence farmers need to cooperate very little).<sup>112</sup> The researchers concluded, “Our data suggest that these between-group behavioral differences . . . are the product of the patterns of social and economic interaction that frame the everyday lives of our subjects.”<sup>113</sup> In layman’s terms, character may be largely a product of experience.

But whatever the underlying cause of differences in individuals’ inclinations toward prosociality, it seems clear that individual variations exist. This last lesson will prove important as we investigate what behavioral science teaches about the likely consequences of the ideology of incentives.

#### IV. THE UNINTENDED BEHAVIORAL CONSEQUENCES OF PAY FOR PERFORMANCE

As we have seen, the optimal contracting literature presumes people are self-seeking, opportunistic actors who will shirk (if they are agents) or renege on promises to pay (if they are principals) unless constrained by enforceable contracts that provide the correct ex ante incentives. As we have also seen, real people often depart from this behavioral model. Business firms, school systems, and medical centers must deal with real people. Thus, this Part explores the question: what does behavioral science tell us about the real behavioral consequences of using ex ante incentives?

##### *A. Relational Contracts and Contractual Incompleteness*

The fundamental nature of this question becomes apparent once we recognize, as contracts experts do, that it is impossible to design a truly “optimal” agency contract that always creates perfect incentives.<sup>114</sup> Like other contracts, employment contracts are

108. *Id.* at 5, 14–15 tbl.2.1.

109. *Id.* at 5.

110. *Id.* at 23 tbl.2.3.

111. HENRICH ET AL., *supra* note 56, at 28.

112. *Id.* at 33, 49.

113. *Id.* at 45.

114. See generally Robert E. Scott, *A Theory of Self-Enforcing Indefinite Agreements*, 103 COLUM. L. REV. 1641 (2003) (observing that all contracts must be to some degree incomplete); Oliver Hart & John Moore, *Incomplete Contracts and Renegotiation*, 56 ECONOMETRICA 755, 776 (1988) (discussing how contracting parties are “forced to write an incomplete contract because of their inability to specify the state of the world in sufficient detail that an outsider can verify whether it has occurred”).

always incomplete, meaning they do not address all the potential issues or disputes that might arise in the future between the parties. For example, an employment contract between parents and a babysitter might address the sitter's hourly wage and the number of hours of work to be provided. However, the contract is unlikely to address what should happen if the child wanders off and becomes lost, or if the child scrawls on the sitter's shoes with indelible markers, or if the parents return home hours late due to some emergency.

Contracts are incomplete for good reasons. One is that humans aren't omniscient. As Melvin Aron Eisenberg has put it, "Contracts concern the future, and are therefore always made under conditions of uncertainty."<sup>115</sup> Problems can arise during performance that neither party thought of, much less discussed in the contract. For example, a bank might hire a derivatives trader only to have the position become obsolete as a result of unexpected financial reform legislation.

Complexity also leads to incompleteness because complexity makes negotiating and drafting contracts expensive. When a corporation hires a CEO, even if the parties could anticipate every issue that might arise in the course of managing the business—from a sudden advance in production technology to a nationwide quarantine due to a flu pandemic—they might find that attempting to draft a formal contract that addressed each and every possible contingency is prohibitively expensive and time-consuming. Instead, they might prefer a short, incomplete contract that addresses only the most important and obvious aspects of the employment relationship (e.g., responsibilities and salary) and leaves other matters to be dealt with in the future, should they arise.<sup>116</sup>

Finally, contracts are often intentionally incomplete about matters that, while important to the parties, are difficult to observe or to prove in court. For example, suppose a New York City school teacher's contract provides for a performance bonus if the teacher's students achieve certain test scores. Suppose further that the contract specifically provides no bonus will be paid if students' scores rise because the teacher tampered with the students' answer sheets. Even if (as allegedly happened in Georgia)<sup>117</sup> a statistical analysis of student answer sheets showed improved test scores, but also a suspiciously high number of changed and corrected test answers, it would be difficult and expensive, and perhaps impossible, for the school district to determine whether the teacher or the students changed the answers—much less prove the matter in court.

Because uncertainty, complexity, and unobservability are endemic, incomplete contracts are everywhere. Even a relatively simple agency contract—say, a contract with a real estate broker to sell a house—contains gaps. What if the homeowner thinks the agent is not marketing the home as enthusiastically as he should? As Steven Shavell puts it, "Contracts typically omit all manner of variables and contingencies that are of potential relevance to contracting parties."<sup>118</sup> Contracts scholar Robert Scott goes further:

---

115. Melvin Aron Eisenberg, *The Limits of Cognition and the Limits of Contract*, 47 STAN. L. REV. 211, 213 (1995).

116. Similarly, uncertainty and complexity can defeat a court's attempt to provide optimal "implied" contractual terms. STOUT, *supra* note 26, at 179–82.

117. Steve Osunsami & Ben Forer, *Atlanta Cheating: 178 Teachers and Administrators Changed Answers to Increase Test Scores*, ABC NEWS (July 6, 2011), <http://abcnews.go.com/us/atlanta-cheating-178-teachers-administrators-changed-answers-increase/story?id=14013113>.

118. STEVEN SHAVELL, *ECONOMIC ANALYSIS OF LAW* 63 (2004).

“All contracts are incomplete.”<sup>119</sup>

But some contracts are more incomplete than others. Contracts fall along a spectrum of completeness. At one end lie “discrete” contracts—simple contracts for exchanges between parties who never expect to deal with each other again. A contract to purchase a laptop computer from an online catalog is an example of a relatively discrete contract. At the other end of the spectrum lie “relational” contracts that involve complex, long-term, uncertain exchanges—for example, a contract to employ a teacher, surgeon, or business executive.<sup>120</sup> Drastic incompleteness is a hallmark of most employment contracts. An empirical study of Fortune 500 CEOs, for example, found that nearly a third had no written employment contract at all, and another third had only bare-bones contracts that spelled out their pay and incentives but few of their duties.<sup>121</sup>

This observation raises the question of how relational contracts, including many and possibly most employment contracts, can work. Purely selfish actors can be counted upon to opportunistically exploit the gaps in relational contracts and perform poorly or not at all. Anticipating this, purely selfish actors would avoid relational exchanges with other purely selfish actors.<sup>122</sup> Yet, real people do enter incomplete relational contracts. In fact, many of our most economically significant exchanges—joint business ventures, apartment leases, building contracts, and of course employment agreements—are relational. Somehow, despite the problems of uncertainty, complexity, and unobservability, relational exchanges take place. How?

Sometimes, opportunistic behavior in relational contracts can be discouraged by fear of loss of reputation. As organizational economist Oliver Williamson has put it (with typical academic style), “reputation effects attenuate incentives to behave opportunistically in interfirm trade—since the immediate gains from opportunism in a regime where reputation counts must be traded off against future costs.”<sup>123</sup> But as Williamson has also noted, “[T]he efficacy of reputation effects is easily overstated . . . .”<sup>124</sup> There are good reasons to question whether reputation can always, or even often, motivate purely selfish actors not to act opportunistically in relational exchanges. For example, reputation becomes an unreliable guarantee as one nears retirement. (This does not seem to deter corporations from hiring executives and directors in their fifties, sixties, and seventies.) It can also be hard for outside observers to determine which party was at fault when a relational deal breaks down; for example, there was widespread public disagreement over the wisdom—or folly—of Hewlett-Packard’s 2010 decision to fire CEO Mark Hurd.<sup>125</sup>

---

119. Scott, *supra* note 114, at 1641.

120. See generally Ian R. Macneil, *Relational Contract Theory: Challenges and Queries*, 94 NW. U. L. REV. 877 (2000) (discussing relational contracts); Stewart Macaulay, *Relational Contracts Floating on a Sea of Custom? Thoughts About the Ideas of Ian Macneil and Lisa Bernstein*, 94 NW. U. L. REV. 775 (2000) (same).

121. Stewart J. Schwab & Randall S. Thomas, *An Empirical Analysis of CEO Employment Contracts: What Do Top Executives Bargain For?* 63 WASH. & LEE L. REV. 231, 240–41 (2006).

122. STOUT, *supra* note 26, at 183–84 (discussing the difficulties of relational contracting between purely self-interested actors).

123. OLIVER E. WILLIAMSON, *THE MECHANISMS OF GOVERNANCE* 116 (1996).

124. *Id.*

125. See Ashlee Vance, *Oracle Chief Faults H.P. Board for Forcing Hurd Out*, N.Y. TIMES, Aug. 9, 2010, [www.nytimes.com/2010/08/10/technology/10hewlett.html?dbk](http://www.nytimes.com/2010/08/10/technology/10hewlett.html?dbk) (noting Oracle CEO Larry Ellison’s critique of the H.P. board’s decision to force Hurd out).

*B. Conscience As a Solution to Contractual Incompleteness*

Given the limits of formal contracts and reputation, how can purely selfish actors participate successfully in relational exchange? Maybe purely selfish actors cannot—at least, not with other purely selfish actors. However, the empirical evidence on prosociality suggests another possibility. Although conventional economic analysis treats contract law as a vehicle for allowing self-interested actors to bind themselves to perform their promises, the story of relational contract may be just the opposite—not a tale of self-interest, but a story of prosocial partners who trust each other and, to at least some extent, look out for each other.

The key to understanding this idea is to understand that when two people contemplate entering a relational contract, each wants protection against the possibility that the other might opportunistically exploit the many gaps that necessarily exist in the contract. Neither the parties nor the courts can reliably fill the gaps because of uncertainty, complexity, and unobservability. Reputational concerns sometimes can check opportunistic behavior, but reputation alone often is not enough. So, parties entering a relational contract may seek to employ a third possible check on opportunism—their contracting partner's conscience.<sup>126</sup>

Suppose, for example, some unanticipated problem or opportunity arises while two parties are performing a contract. Where two purely selfish actors would instantly find themselves locked in conflict over who should bear the loss or claim the gain, prosocial partners could resolve the question far more easily—say, by splitting the unanticipated gain or loss—because they share, to at least some extent, the common goal of promoting their mutual (not only individual) welfare. Nor do prosocial partners need to reduce every detail of their bargain to writing. They trust each other to focus on performing, not on selfishly searching for loopholes. Finally, even when some element of performance is unobservable or unverifiable, prosocial partners will try to hold up their end of the deal.

In brief, an implicit “term” of relational contracts seems to be that each party agrees that, in performing, she will suppress her Mr. Hyde personality and adopt a Jekyll-like attitude toward her counterparty. As Ian Macneil has put it, a relational contract is just that—a relationship—characterized by (among other attributes) “role integrity,” “flexibility,” and “reciprocity.”<sup>127</sup> Using the language of behavioral science, a relational contract creates a social context that promotes unselfish prosocial behavior. The spectrum from simple discrete contracts to complex, incomplete relational contracts, accordingly, can be viewed as a spectrum from Hydish behavior toward one's counterparty, to Jekyllish behavior.

This approach offers a number of insights into the questions of how relational exchanges really work, and how contract law and contract lawyers can make them work

---

126. See generally Margaret M. Blair & Lynn A. Stout, *Trust, Trustworthiness, and the Behavioral Foundations of Corporate Law*, 149 U. PA. L. REV. 1735 (2001) (discussing how trust can fill gaps in incomplete contracts); STOUT, *supra* note 26, at 185–88 (same). For empirical evidence, see generally Martin Brown et al., *Contractual Incompleteness and the Nature of Market Interactions* (Inst. for Empirical Research in Econ., Working Paper No. 38, 2002), available at <http://www.iew.uzh.ch/wp/iewwp038.pdf>; Peter Kollock, *The Emergence of Exchange Structures: An Experimental Study of Uncertainty, Commitment, and Trust*, 100 AM. J. SOC. 313 (1994) (investigating the roles of reputation and social exchanges in the relationships).

127. Macneil, *supra* note 120, at 897.

better.<sup>128</sup> But it carries especially important implications for ex ante incentive contracts. This is because the behavioral evidence indicates that employment contracts that rely on material incentives to motivate performance suppress the vital force of conscience—essential for relational contracting—and encourage undesirably selfish, opportunistic, and even illegal behavior. Incentive pay does this through at least three separate but mutually reinforcing mechanisms: (1) by changing perceptions of social context in ways that encourage selfishness; (2) by creating material temptations that can extinguish conscience; and (3) by introducing selection bias against individuals with relatively prosocial characters.

### C. Social Context and Crowding Out

Let us first consider how pay-for-performance contracts frame social context. As discussed in Part III, when choosing between asocial and prosocial behavior, people pay close attention to social context. Contract negotiations provide a social context that can be ambiguous. Is the contract in question a discrete contract, in which case mutually selfish behavior may be appropriate and expected? Or is it a relational contract calling for trust, cooperation, and mutual regard for each other's interests? In extreme cases (buying a car versus negotiating a prenuptial agreement) the distinction is clear. But in many other cases, including employment contracts, the contract may have both discrete and relational elements.

In an ambiguous situation, an actor who wants to rely on her contract partner's conscience wants to signal as clearly as possible that performance calls for mutually considerate, rather than arm's length, behavior. Yet what signal does an employer send when it uses ex ante incentive contracts to motivate its employees? The pay-for-performance approach inevitably signals that the employer in question views the employment relationship as an arm's length exchange in which self-interested behavior is appropriate, expected, and even encouraged. This is likely to induce the behavioral phenomenon social scientists call "crowding out."<sup>129</sup>

In one classic study of crowding out, researchers studied ten day-care centers where parents occasionally arrived late to pick up their children, forcing the teachers to stay after closing time.<sup>130</sup> The researchers convinced six of the centers to introduce a new policy of fining the parents who arrived late.<sup>131</sup> The result? Late arrivals increased significantly.<sup>132</sup> In another famous study, people were asked to donate blood either for free, or for a modest payment. More people agreed to donate blood for free than for cash.<sup>133</sup>

From an economic perspective, these are bizarre results. How can raising the price of an activity prompt people to purchase more of it, or paying people to do something

---

128. See generally STOUT, *supra* note 26, at 175–99 (discussing acceptance of incentive ideology).

129. See generally Frey & Osterloh, *supra* note 39, at 102–06 (discussing crowding out theory and evidence).

130. See generally Uri Gneezy & Aldo Rustichini, *A Fine is a Price*, 29 J. LEGAL STUD. 11 (2000).

131. *Id.* at 3.

132. *Id.*

133. Carl Mellström & Magnus Johannesson, *Crowding Out in Blood Donation: Was Titmuss Right?*, 6 J. EUR. ECON. ASS'N 845, 852–56 (2008).

cause them to do it less? The answer, according to crowding out theory, is that associating a particular kind of behavior or interaction with monetary payments changes the social context, making the interaction look like a market transaction in which purely selfish behavior is deemed appropriate. Thus, fining parents who arrived late to the day-care center signaled that lateness was not a social faux pas, but a market decision parents were free to make without worrying about teachers' welfare. Similarly, paying people to donate blood signals that donation is a voluntary decision to sell personal property, rather than a social obligation to contribute to group welfare. By emphasizing external material incentives, day-care center fines and blood donation payments crowd out internal "incentives" like guilt and empathy.<sup>134</sup>

For similar reasons, incentive pay can be expected to crowd out less-selfish employee motives like trust, loyalty, and commitment. This is because emphasizing material incentives manipulates each of the four social cues we have examined—instructions from authority, perceptions of common group membership, beliefs about others' prosocial or asocial behavior, and perceptions of benefits to others—in a fashion that promotes selfishness.<sup>135</sup> First, offering a material incentive to induce an employee to perform a particular act inevitably sends the unspoken signal that the employer believes the employee would not otherwise perform; in other words, the employer expects selfish behavior and indeed views it as appropriate to the task at hand. Second, traditional incentive-pay schemes undermine a sense of group identity and common fate because they encourage individuals to believe their compensation and success is tied only to their own efforts, rather than the group. Third, when pay-for-performance schemes are used widely in an organization, they support the perception that *other* employees are likely to behave selfishly (not to mention signaling the selfishness of the employer, who proposes to withhold compensation regardless of circumstances unless its performance metrics are met). Fourth, pay-for-performance schemes imply that the employee's selfishness actually benefits the employer. Otherwise, why would it be rewarded?

Understanding how incentive pay reframes social context and crowds out ethics and conscience offers insight into what Bradley Birkenfeld might have been trying to say when he told the judge that UBS had "incentivized" him to help its clients evade paying taxes.<sup>136</sup> Birkenfeld probably was not suggesting he was excused from breaking the law simply because he received a material benefit from doing so; surely, he realized no judge would be sympathetic to the notion that self-interest justifies illegality. Rather, Birkenfeld was saying that, by incentivizing its employees to help its clients evade taxes, UBS had created a social context that gave him *permission* to do so.

Incentive contracts, it turns out, do more than change behavior. At a very deep level they change *motivations*. Emphasizing self-interest turns out to be a self-fulfilling prophecy. By treating people as if they care only about their own material rewards, we ensure that they do.<sup>137</sup>

---

134. The framing effects of monetary exchange are so powerful that merely earning money playing a game of Monopoly makes experimental subjects less likely to help a researcher pick up "accidentally" dropped pencils immediately afterwards. MARGARET HEFFERNAN, *WILLFUL BLINDNESS: WHY WE IGNORE THE OBVIOUS AT OUR PERIL* 186–87 (2011).

135. STOUT, *supra* note 26, at 249–52.

136. *Supra* note 3 and accompanying text.

137. See generally STOUT, *supra* note 26, at 247–52 (discussing how emphasizing selfishness as a motive

*D. Conscience, Temptation, and Cognitive Dissonance*

Even if only at an intuitive level, many employers recognize the value of creating a workplace that promotes employee trustworthiness and loyalty, and (incentive schemes notwithstanding) they attempt to manipulate social context to promote unselfish employee behavior. The strategy can be as simple as posting a sign that reads, "Customer Service Is Our Priority," or as elaborate as sponsoring week-long corporate retreats where executives listen to motivational speakers and go white-water rafting together.

What is probably less well recognized, however, is that even the most careful effort to create a workplace that supports conscientiousness and prosociality can run aground on the rock of employee self-interest. This is because, as we have seen, conscience works best when it does not conflict too directly with self-interest. Unlike Oscar Wilde, most of us can resist small temptations. It is the big ones that do us in.

Pay-for-performance schemes can create very big temptations indeed. This is especially true in corporate environments, because a hallmark of the American public corporation is that it permits the accumulation of enormous wealth.<sup>138</sup> Incentive contracts based on metrics subject to executives' influence, especially metrics that executives can manipulate or falsify, create tempting opportunities for executives to try to extract this wealth for themselves through behavior that imposes costs on the corporation or on third parties.<sup>139</sup> In effect, once corporate directors agree to compensate executives with ex ante incentive contracts, the board has effectively ceded a great deal of control over the firm's assets to the executives themselves. To the extent the incentive contracts are incomplete—as all incentive contracts must be—they also inevitably present executives with opportunities to try to expropriate corporate assets through opportunistic, illegal, or otherwise undesirable behavior.

Thus, corporate employers that rely primarily on ex ante material incentives to motivate their employees are playing a dangerous game. It is almost always possible, and sometimes far easier, for an executive to meet a performance metric through unethical or illegal behavior rather than hard work. For example, in the case of Enron, executive stock option grants intended to motivate employees to "maximize shareholder wealth" in fact motivated them to commit a massive accounting fraud. As Franklin Raines, then-CEO of Fannie Mae, described the causes of the Worldcom and Enron scandals in an interview in *Business Week*: "[Y]ou put enough money in front of people, good people will do bad things."<sup>140</sup> Employees who would never think of shoplifting or other small acts of larceny will ignore the voice of conscience if the opportunity for a hugely profitable fraud comes

---

increases the incidence of selfishness).

138. See generally Margaret M. Blair, *Locking in Capital: What Corporate Law Achieved for Business Organizers in the Nineteenth Century*, 51 UCLA L. REV. 387 (2003) (discussing a corporation's capacity to own and accumulate assets).

139. For example, incentive contracts that allow executives to personally profit from increased profits create incentives for those executives to cause the firm to load up on risk. In a survey of 562 risk managers, compensation practices were identified as a chief cause of banking failures. HEFFERNAN, *supra* note 134, at 189.

140. *Executive Compensation Hearing*, *supra* note 5, at 221. One of the great ironies of Raines' testimony was that he was subsequently sued for manipulating Fannie Mae's financial statements in order to maximize his own incentive pay. The parties settled the suit. James R. Hagerty, *Fannie Mae Ex-Officials Settle*, WALL ST. J., Apr. 19, 2008, at A3.

along. Thus, a workplace that relies on large material incentives to motivate employees is also a workplace that suppresses the force of conscience.

Moreover, once otherwise honest individuals succumb to temptation and indulge in unethical or illegal behavior, they become more likely to cross ethical lines again in the future, and more easily. It is a truism among those who study business frauds that white-collar offenders usually start with small violations, before escalating into full-blown criminality. Many psychologists believe the reason has to do with the phenomenon known as cognitive dissonance.<sup>141</sup> Cognitive dissonance theory posits that people desire consistency between their beliefs and their actual behavior. When their actions become inconsistent with their attitudes, rather than change their behavior, they tend to change their attitudes by rationalizing their actions to restore apparent consistency. The result is that when incentives tempt people to do things they are otherwise reluctant to do, they respond to the inconsistency between their beliefs (“I should not do this”) and their behavior (“I did this”) by changing their beliefs (“Since I did this, it must be something okay to do”).

Thus, “induced compliance” shifts people’s views about the appropriateness of their own conduct because “in the battle between changing one’s attitude and changing one’s behaviour, attitudes are the easiest to change.”<sup>142</sup> Once incentive pay tempts employees into opportunistic or illegal behavior, they change their beliefs about what is “opportunistic” or “illegal” so they can continue thinking of themselves as fundamentally ethical and law-abiding. This change makes it much easier for them to justify similar unethical or illegal behavior to themselves in the future.

Pay-for-performance schemes accordingly can create criminogenic environments that first tempt honest individuals into unethical or illegal behavior, then invite them to adopt looser views about what is unethical or illegal in the first place. It is sometimes said in the business world that pressure makes diamonds. We should bear in mind it also makes felons.

#### *E. Selection Bias and the Question of Character*

Thus far we have focused on how incentive pay discourages prosocial behavior in individuals who are fully and equally capable of acting prosocially. As noted earlier, however, experimental gaming demonstrates that individuals differ substantially in their inclinations toward prosocial action.<sup>143</sup> Few of us are psychopaths, utterly without conscience. Nevertheless, some people are more inclined toward conscientious behavior than others are. This too has implications for the wisdom of using incentive contracts.

In particular, because contracts are incomplete, most incentive contracts create opportunities for employees to try to reap rewards through behavior that technically satisfies the contracts but is illegal, unethical, or not truly in the employers’ interests. To the extent this is true, we can expect employers who rely on incentives to attract more than their share of opportunistic employees. Incentive schemes naturally attract the relatively opportunistic, because relatively opportunistic individuals see potential for

---

141. See generally JOEL COOPER, *COGNITIVE DISSONANCE: FIFTY YEARS OF A CLASSIC THEORY* (2007).

142. *Id.* at 15.

143. See *supra* notes 103–113 and accompanying text (discussing studies that find individual differences).



personal gain that individuals who are more constrained by personal ethics would discount as out-of-bounds and unavailable. Thus, it is perhaps no coincidence that Wall Street executives are widely perceived to lack both empathy and ethics.<sup>144</sup> Investment banks and other financial firms are notorious for offering their employees incentive compensation packages that create opportunities to reap millions of dollars.<sup>145</sup>

Moreover, once a workplace begins to attract more than its share of relatively opportunistic or unethical employees, through a variety of different but mutually reinforcing effects, it will also begin to repel the relatively prosocial and to subvert the prosocial employees who remain at the firm into committing their own ethical lapses (which, given cognitive dissonance, diminishes their prosociality). This phenomenon has been described by William Black, an expert on white-collar crime, as a “Gresham’s dynamic in which bad ethics drives good ethics out of the marketplace.”<sup>146</sup> Black, who served as deputy staff director for the federal commission that investigated widespread fraud in the savings and loan industry in the late 1980s, concludes that incentive-based pay schemes created a Gresham’s dynamic in the recent subprime mortgage crisis. The practice of compensating loan brokers with incentive pay based largely on the number of loans they originated led to a rapid deterioration in broker ethics and a subsequent loosening of mortgage lending standards, with disastrous results.<sup>147</sup>

There are several reasons why workplaces that attract more than their share of opportunists drive out prosocial behavior. First, relatively ethical employees conclude they suffer a competitive disadvantage, and decamp for greener pastures where their prosocial proclivities are less of a handicap. Alternatively, the relatively ethical may conclude they can no longer afford to be so squeamish, and decide to dispense with their ethics. Second, as the population of a workplace becomes dominated by opportunists, with fewer and fewer conscientious employees, the risk that an opportunist’s misconduct will be revealed by a whistleblower declines. Third, as discussed in Part III, when a workplace becomes crowded with opportunists, this changes social context. When “everybody does it” (whether “it” is approving low-quality mortgage loans, committing accounting fraud, or cheating on income taxes), it is easy to conclude that you can do it, too.

Adverse selection pressures accordingly lead workplaces that rely on pay for performance to attract a disproportionate share of relatively unethical and opportunistic employees. For example, between 2000 and 2007, more than 10,000 individuals with criminal records became mortgage brokers in Florida, leaving one to wonder how many individuals attracted to that business were criminals who simply hadn’t been caught yet.<sup>148</sup> Once this occurs, the result can be a self-reinforcing dynamic in which prosocial individuals and prosocial behaviors are driven out. It may even be possible to reach a

---

144. Consider *Rolling Stone*’s description of investment bank Goldman Sachs as “a great vampire squid wrapped around the face of humanity, relentlessly jamming its blood funnel into anything that smells like money.” Matt Taibbi, *The Great American Bubble Machine*, ROLLING STONE, July 9, 2009, <http://www.rollingstone.com/politics/news/the-great-american-bubble-machine-20100405>.

145. Margaret Heffernan provides an interesting case study in how incentive pay led to criminal behavior at Lehman brothers. HEFFERNAN, *supra* note 134, at 128–31.

146. *Executive Compensation Hearing*, *supra* note 5, at 201.

147. *Id.* at 201.

148. FIN. CRISIS INQUIRY COMM’N, *supra* note 7, at 11.

tipping point in which opportunistic behavior becomes so prevalent that prosocial behavior within the company virtually disappears. Think of Enron, Countryside Financial, or (in *Rolling Stone's* opinion) Goldman Sachs. The firm becomes, in effect, a criminal enterprise populated primarily by employees who act like psychopaths—at least until they get on the elevator and go home.

#### V. CONCLUSION: ALTERNATIVES TO PAY FOR PERFORMANCE

As we have seen, optimal contracting theory embraces the quest for the complete employment contract that perfectly aligns the interests of agent and principal so that all “agency costs” disappear. This quest is a bit like the quest for the Holy Grail. No perfect contract is possible, and gaps inevitably remain. What fills the gaps? According to standard optimal contracting theory, only reputational concerns can, and any contractual gap that cannot be filled by reputation will be exploited opportunistically and become a source of agency costs. But (again according to the theory) this should not discourage us from the quest to design *ex ante* incentive contracts that are as complete as possible, for without such contracts, opportunism is inevitable and uncontrollable.

Behavioral science offers a different perspective. The human capacity to act prosocially can also fill gaps in incomplete relational contracts, and motivate contract partners to perform even when there is no realistic threat, or insufficient threat, of legal or reputational sanction if they don't. This possibility deserves our attention, for behavioral science also teaches that a workplace that emphasizes *ex ante* financial incentives will tend to suppress the force of conscience in at least three ways: by shifting social context, by creating temptations, and by introducing selection biases that favor less conscientious individuals.

But if we don't use incentive contracts to motivate employee performance, what can we use to motivate them? One can only get so much commitment, loyalty, and hard work for free. In a capitalist society—perhaps in any society—few people are willing to work long and hard for nothing. (When you take from each according to his ability and give to each according to his needs, you are likely to end up with lots of needy, incompetent people.) Eventually the siren call of self-interest invites even the most dedicated agent to ask, “What's in it for me?”<sup>149</sup>

Something must be. Many Americans volunteer their time for various worthy causes. But when it comes to full-time employment, most insist on being paid. It is important to emphasize that critiquing incentive pay is not the same thing as critiquing the general idea of paying. There are lots of ways to compensate and reward executives and other agents for their efforts, beyond using large, material, *ex ante* incentives.

Thus, this Article concludes by addressing the question: if we don't use high-powered *ex ante* financial incentives to motivate people, what can we use? Behavioral science suggests an intriguing and counterintuitive answer. When employment contracts are highly incomplete, instead of relying on material, large, *ex ante* contractual incentives, employers might do well to adopt the *opposite* approach: emphasize rewards

---

149. Put differently, prosocial behavior tends to disappear when the personal cost of behaving prosocially gets too high. See *supra* notes 100–102 and accompanying text (discussing effects of personal cost on prosociality).

that are nonmaterial, relatively modest, and determined ex post on the basis of subjective evaluations.

Let us first consider the advantages of using nonfinancial rewards. Behavioral science supports human resource experts' belief that employing nonmaterial rewards—greater job responsibilities, a better parking space, an “Employee of the Month” plaque—can work as well or better than emphasizing material rewards like cash bonuses or stock options.<sup>150</sup> Most obviously, job titles and award plaques cost firms much less to provide. But there are psychological as well as economic advantages. Unlike monetary rewards, which have intrinsic value unrelated to social context, nonmonetary rewards appeal to employees' desires for status, esteem, and feelings of in-group membership. Such social motivations naturally focus employee attention on social context rather than personal financial circumstances—exactly where we want to focus employee attention to encourage prosocial behavior.<sup>151</sup> Nonmonetary rewards also seem to do a better job of preserving intrinsic employee motivations, such as interest, creativity, and desire for mastery. In his bestseller, *Drive*, Daniel Pink emphasizes this advantage, surveying the extensive experimental evidence that demonstrates how the prospect of monetary rewards often reduces individuals' performance on tasks requiring creativity and persistence.<sup>152</sup>

Of course, man or woman cannot live on “Employee of the Month” awards alone. Even the most prosocial employee has to pay the rent and buy groceries. Thus, most employers must pay their employees reasonably competitive financial compensation. But there are advantages, apart from the obvious cost savings, in trying to keep financial compensation relatively modest. As we have seen, firms that avoid offering very large financial incentives may benefit from employee selection bias because they are less likely to attract selfish opportunists.<sup>153</sup> Of course, some businesses—used car dealerships, hedge funds—may want to attract selfish opportunists, because employees perform tasks that are relatively simple, the desired outcome is certain, and employee performance is easy to observe, making it feasible to design relatively complete employment contracts with few gaps for employees to exploit. But many businesses—schools, hospitals, public corporations—must necessarily rely on employment contracts that are far more incomplete and leave greater room for opportunistic behavior. For example, James Sinegal, the CEO of Costco, works under an employment contract, whose terms

---

150. Susanne Neckermann et al., *What Is an Award Worth? An Econometric Assessment of the Impact of Awards on Employee Performance* (Ctr. for Econ. Studies, Working Paper No. 2657, 2009), available at [http://www.cesifo-group.de/portal/page/portal/DocBase\\_Content/WP/WP-CESifo\\_Working\\_Papers/wp-cesifo-2009/wp-cesifo-2009-05/cesifo1\\_wp2657.pdf](http://www.cesifo-group.de/portal/page/portal/DocBase_Content/WP/WP-CESifo_Working_Papers/wp-cesifo-2009/wp-cesifo-2009-05/cesifo1_wp2657.pdf).

151. A meta-analysis of 12 major research studies on employee engagement found that such social factors as the degree to which the employer is perceived to care about employees, pride in the company, attachment to coworkers, and personal relationships with managers were all important drivers of employee performance. CONFERENCE BD., *EMPLOYEE ENGAGEMENT: A REVIEW OF CURRENT RESEARCH AND ITS IMPLICATIONS 6* (2006), available at <http://montrealoffice.wikispaces.com/file/view/Employee+Engagement++Conference+Board.pdf>.

152. See generally DANIEL H. PINK, *DRIVE: THE SURPRISING TRUTH ABOUT WHAT MOTIVATES US* (2009) (discussing evidence on best ways to promote creativity and persistence).

153. See generally ROBERT H. FRANK, *WHAT PRICE THE MORAL HIGH GROUND? HOW TO SUCCEED WITHOUT SELLING YOUR SOUL* (1990) (discussing how businesses can thrive by appealing to ethics and conscience rather than selfishness and greed).

supposedly “fit on a cocktail napkin.”<sup>154</sup> In such cases, firms that can attract conscientious employees—teachers who want students to learn, doctors who want to help patients, CEOs who want to leave a legacy rather than simply take as much money as possible—have an advantage. In addition to limiting adverse selection bias, avoiding incentive contracts that provide for very large rewards also reduces the risk that otherwise prosocial employees might be tempted to ignore their consciences, and avoids creating a social context that suggests the employer believes employees work hard only for self-serving reasons.

Finally, and perhaps most importantly, behavioral science cautions against the common yet often unspoken belief that compensation must be tied to predetermined, objective metrics in order to be effective. One of the most dangerous consequences of incentive ideology is that it blinds us to the possibility of using subjectively determined *ex post* rewards rather than *ex ante* objective incentives to motivate performance in relational contracts.<sup>155</sup> The notions that employees might work conscientiously even without predetermined incentives, or that employers might voluntarily reward extracontractual efforts, conflict directly with optimal contracting theory’s basic assumption that both principals and agents are opportunistic and purely self-interested actors. According to optimal contracting theory, no rational agent would ever work hard simply because a principal said, “Do a good job, and I’ll reward you appropriately.” Nor would any rational principal ever voluntarily reward an agent foolish enough to put in extra effort. These sorts of prosocial behaviors require trust and trustworthiness, which play no part in optimal contracting theory.

Yet trust and trustworthiness do play an important part in explaining real human behavior.<sup>156</sup> This can be seen quite clearly in an interesting experimental variation on the social dilemma game called, appropriately enough, the “trust game.” A trust game is simply a social dilemma in which two players act sequentially rather than simultaneously. One of the two subjects (the “trustor”) is first given a sum of money, say \$100. Both subjects are told the trustor can choose to contribute some or all of the \$100 to an investment fund, which the researchers will triple and then give to the second subject (the “trustee”). The trustee then gets to choose whether she wants to keep the tripled funds entirely for herself, or return all or some portion back to the trustor.

In a well-designed trust game where subjects play only once and anonymously, no rational and selfish trustee would ever donate any of the tripled funds back to the trustor. Anticipating this, no rational and selfish trustor would ever donate any of his initial stake to the investment fund. In real trust games, however, the trustor typically shares more than half his funds, and the trustee typically repays the trustor with a slightly larger amount.<sup>157</sup>

Employment relationships that rely on *ex post* rewards rather than *ex ante* incentives

---

154. Gretchen Morgenson, *Executive Pay: A Special Report; Two Pay Packages, Two Different Galaxies*, N.Y. TIMES, Apr. 4, 2004, [www.nytimes.com/2004/04/04/business/executive-pay-a-special-report-two-pay-packages-two-different-galaxies.html](http://www.nytimes.com/2004/04/04/business/executive-pay-a-special-report-two-pay-packages-two-different-galaxies.html).

155. Cf. HEFFERNAN, *supra* note 134, at 56 (“Economic models work in ways very similar to . . . ideologies: pulling in and integrating the information that fits the model, leaving out what can’t be accommodated.”).

156. See generally Blair & Stout, *supra* note 126 (discussing role of trust and trustworthiness in business).

157. STOUT, *supra* note 26, at 9; Gätcher, *supra* note 63, at 8.

are directly analogous to the trust game. The employer first trusts the employee by committing to pay him a salary that is not contingent on meeting objective metrics. Next, the employee trusts the employer by working harder and more honestly than the employer could force him to work under the terms of the formal contract. Then, the employer reciprocates the employees' trust by giving the employee a raise and more job responsibilities. This process of reciprocal trust and trustworthiness continues until either the employee retires, or one of the parties fails to reciprocate, and the employment relationship is severed because the employee either quits or is fired.<sup>158</sup> Experimental tests of compensation arrangements that rely on employee trust and employer trustworthiness this way show that they can be more effective than ex ante incentive contracts at inducing employee effort in repeated interactions.<sup>159</sup>

At this point, any seasoned businessperson over the age of 50 should be experiencing *déjà vu*. Optimal contracting theory recommends that employers seek to negotiate employment contracts that are as complete as possible and that emphasize large material rewards that are tied to objective performance metrics determined ex ante. Behavioral science, however, counsels that when contracts are seriously incomplete, we might do better to adopt the opposite approach: offer relatively modest pay, emphasize nonmaterial rewards, and adjust financial compensation ex post on the basis of the employer's subjective satisfaction with the employee's performance. This second approach is *exactly what the business world mostly relied on before Congress passed tax legislation requiring corporations to tie executive pay to performance*.

As discussed in Part II, before the 1993 adoption of I.R.C. Section 162(m), stock options and other forms of incentive pay tied to objective metrics played a far less important role in executive compensation practices than they do today.<sup>160</sup> CEOs and other executives typically received relatively modest salaries along with a variety of noncash perquisites such as nicer offices, better parking spaces, and promotions to larger divisions. Cash bonuses were common but relatively modest and set after-the-fact, on the basis of the employee's performance as viewed subjectively by the company's board of directors or senior managers. In other words, the business world followed exactly the sort of compensation practices behavioral science recommends when contracts are seriously incomplete. Judging from pre-1993 corporate performance and investor returns, the system worked reasonably well.<sup>161</sup>

---

158. Cf. Ronald J. Gilson et al., *Braiding: The Interaction of Formal and Informal Contracting in Theory, Practice, and Doctrine*, 110 COLUM. L. REV. 1377 (2010) (describing a similar reciprocal trust dynamic in commercial contracting).

159. See generally Martin Brown et al., *Relational Contracts and the Nature of Market Interactions*, 72 ECONOMETRICA 747, 749 (2004) (investigating compensation and employee efforts); Ernst Fehr & Klaus M. Schmidt, *Fairness and Incentives in a Multi-Task Principal-Agent Model*, 106 SCANDINAVIAN J. ECON. 453 (2004) (same); Ernst Fehr et al., *Fairness and Contract Design*, 75 ECONOMETRICA 121 (2007) (same). A real-life example can be found in the Mayo Clinic, perhaps the most highly regarded health care practice in the nation, which insists on keeping all staff on fixed salaries in order to maintain its culture of collaboration and patient focus. Leonard L. Barry & Kent D. Seltman, *Building a Strong Services Brand: Lesson from Mayo Clinic*, 50 BUS. HORIZONS 199, 202 (2007).

160. See *supra* notes 20–31 and accompanying text. A 1988 paper by George Baker, Michael Jensen, and Kevin Murphy critiqued prevailing compensation practices as “largely independent of performance.” George P. Baker et al., *Compensation and Incentives: Practice v. Theory*, 43 J. FIN. 593, 593 (1988).

161. See *supra* notes 43–47 and accompanying text.

History accordingly suggests that when they are left to their own devices, businesspeople tend to be pretty good intuitive behavioral scientists. Indeed, they seem superior in this regard to the academics and regulators who continue to argue that we must tie pay to ex ante performance metrics. (These academics and regulators seem not to have noticed the rather obvious fact that their own relatively modest pay isn't tied to much of anything.) As evidence, some companies have already found an end run around Section 162(m) that allows them to use ex post subjective rewards in a fashion quite reminiscent of standard executive compensation practices before the rise of pay-for-performance ideology. This can be seen in the increasing use of "plan within a plan" pay schemes that give executives ex ante objective performance targets, but also allow directors to retain subjective ex post discretion to *reduce* the maximum compensation paid to the executive, even if the objective goal is met.<sup>162</sup>

What do such developing business practices imply about the ideology of incentives? Most important, that it is just that: merely ideology, and counterproductive ideology to boot. America's great achievements in the twentieth century—sending humans to the moon, winning World War II, beating polio, building great global public corporations like IBM, Ford, Xerox, and General Electric—were all accomplished without the aid of "optimal contracting." Yet, despite the absence of reliable empirical evidence to support it, a belief that incentive contracts are essential to good performance not only captured our corporate boardrooms, but it is now spreading to our schools, hospitals, and newsrooms as well. Behavioral science and history both caution against this development.

---

162. See SHEARMAN & STERLING, LLP, DELAWARE FEDERAL COURT ISSUES SIGNIFICANT RULING CONCERNING IMPACT OF EXECUTIVE COMPENSATION PROVISIONS OF THE INTERNAL REVENUE CODE ON DELAWARE CORPORATE LAW GOVERNING SHAREHOLDER VOTING (2013), available at [http://www.Shearman.com/~media/Files/NewsInsights/Publications/2013/07/Delaware-Federal-Court-Issues-Significant-Ruling\\_/Files/View-full-memo-Delaware-Federal-Court-Issues-Sig\\_/FileAttachment/Delaware-FederalCourtIssuesRulinonImpactofExecut\\_.pdf](http://www.Shearman.com/~media/Files/NewsInsights/Publications/2013/07/Delaware-Federal-Court-Issues-Significant-Ruling_/Files/View-full-memo-Delaware-Federal-Court-Issues-Sig_/FileAttachment/Delaware-FederalCourtIssuesRulinonImpactofExecut_.pdf) (discussing "plan within a plan" under Section 162(m) in relation to Judge Sue L. Robinson's ruling in *Freeman v. Redstone*, Civ. No. 12-1052-SLR, 2013 WL 3753426 (D. Del. July 16, 2013)).