TROLLEY PROBLEMS

F.E. Guerra-Pujol*

"Do you kill one person to save five?"[1]

ABSTRACT

This Article proposes a novel and probabilistic solution to the famous "trolley problem" in moral philosophy. In short, the Author would solve the trolley problem by conducting an auction from behind a veil of ignorance.

TABLE OF CONTENTS

I. Introduction

Consider the following famous thought experiment: a trolley car with

---

1. *See* Cathy Lynn Grossman, *Do You Kill One Person to Save Five?*, USA TODAY (Oct. 10, 2010), http://content.usatoday.com/communities/Religion/post/2010/10/killing-people-morals-ethics-trolley-problem/1#.VEZ6laMo6bh.

defective brakes is barreling down a railway line at a high speed and will, with absolute certainty, kill five workers down the line unless the trolley is stopped or somehow diverted.[2] You have four choices:

(a) Release a switch in time to divert the trolley down a side track where there is only one worker present. If you choose option (a), only one person is killed.

(b) Push a fat man (who happens to be standing next to the rail line) directly in front of the trolley. If you choose option (b), only the fat man is killed.

(c) Do nothing. If you choose option (c), all five workers on the main track will be killed.

(d) Conduct a Coasean auction behind a Rawlsian "veil of ignorance."[3] That is, allow the parties in this thought experiment to bid for the right to remain alive. If you choose option (d), it is not clear who will live and who will die.

So, what would you do? What *should* you do? When is it morally or legally permissible to release the lever, to push the fat man in front of the trolley, to do nothing, or to conduct an auction behind a veil of ignorance? Furthermore, is there a right answer to this dilemma? That is, does the trolley problem have a solution?

The remainder of this Article consists of five Parts. Part II presents two well-known versions of the trolley problem—the standard version and the fat-man variation—and considers the differences and similarities between both versions. Part III considers the trolley problem from a legal perspective. Part IV presents a novel solution to the trolley problem and notes that the Author would solve the trolley problem by conducting an auction from behind a veil of ignorance. Part V explores some moral implications of this private-bargaining or auction approach without the veil of ignorance.

---

2. For an extremely melodramatic recap of the classical trolley problem in its usual iterations, see andrewjb2008, *Presentation Part 1 – Trolley Problem*, YOUTUBE (Apr. 4, 2008), http://www.youtube.com/watch?v=6WB3Q5EF4Sg.

3. The idea of using auctions to allocate legal rights can be credited to Ronald Coase. *See* R. H. Coase, *The Federal Communications Commission*, 2 J.L. & ECON. 1, 17–25 (1959) [hereinafter Coase, *FCC*]; R. H. Coase, *The Problem of Social Cost*, 3 J.L. & ECON. 1, 2–6 (1960) [hereinafter Coase, *Social Cost*]. The idea of negotiating behind a veil of ignorance is often credited to John Rawls. *See generally* JOHN RAWLS, A THEORY OF JUSTICE, 118–23 (rev. ed. 1999). Credit for this original idea, however, should go to John Harsanyi. *See* John C. Harsanyi, *Cardinal Utility in Welfare Economics and in the Theory of Risk-Taking*, 61 J. POL. ECON. 434, 434–35 (1953); *see also* John C. Harsanyi, *Cardinal Welfare, Individualistic Ethics, and Interpersonal Comparisons of Utility*, 63 J. POL. ECON. 309, 313–14 (1955) [hereinafter Harsanyi, *Cardinal Welfare*].

II. THE NATURE OF THE TROLLEY PROBLEM

The "trolley problem" is one the most famous thought experiments in the domains of law and morality, one that raises deep and difficult questions about the nature of right and wrong.[4] But what is the trolley problem? There are really two versions of the trolley problem, not one.[5]

A. *Trolley Problem No. 1: The Standard Version*

The philosopher Philippa Foot stated the original or standard version of the trolley problem as follows:

> Suppose that a judge or magistrate is faced with rioters demanding that a culprit be found for a certain crime and threatening otherwise to take their own bloody revenge on a particular section of the community. The real culprit being unknown, the judge sees himself as able to prevent the bloodshed only by framing some innocent person and having him executed. . . . *To make the parallel as close as possible it may rather be supposed that he is the driver of a runaway tram which he can only steer from one narrow track on to another; five men are working on one track and one man on the other; anyone on the track he enters is bound to be killed.* In the case of the riots the mob have five hostages, so that in both the exchange is supposed to be one man's life for the lives of five.[6]

In Professor Foot's original formulation of the trolley problem, the moral question is, "May one sacrifice the life of one person in order to save the lives of five persons?"[7] Thus, Professor Foot's version of the trolley problem poses a difficult moral dilemma because it forces one to make a "tragic choice"—save one

---

4. *See, e.g.*, Philippa Foot, *The Problem of Abortion and the Doctrine of the Double Effect*, *in* KILLING AND LETTING DIE 266 (Bonnie Seinbock & Alastar Nireress eds., 2d ed. 1994); Judith Jarvis Thomson, *The Trolley Problem*, 94 YALE L.J. 1395 (1985) [hereinafter Thomson, *Trolley Problem*]; Judith Jarvis Thomson, *Killing, Letting Die, and the Trolley Problem*, 59 THE MONIST 204 (1976), *reprinted in* ETHICAL THEORY 543 (Russ Shafer-Landau ed., 2013) [hereinafter Thomson, *Killing*]; *see generally* Michael Gorr, *Thomson and the Trolley Problem*, 59 PHIL. STUD. 91 (1990).

5. Readers who are already familiar with both versions of the trolley problem may feel free to skip this section of the paper.

6. Foot, *supra* note 4, at 269–70 (emphasis added). The trolley problem is a substantial part of Professor Foot's legacy. *See* William Grimes, *Philippa Foot, Renowned Philosopher, Dies at 90*, N.Y. TIMES (Oct. 9, 2010), http://www.nytimes.com/2010/10/10/us/10foot.html ("Philippa Foot, a philosopher who argued that moral judgments have a rational basis, and who introduced the renowned ethical thought experiment known as the Trolley Problem, died at her home in Oxford, England, on Oct. 3, her 90th birthday.").

7. The original trolley problem only offers choices (a) and (c) from Part I of this Article.

life or save five lives.[8]

Notice that this particular tragic choice poses a reciprocal problem: no matter how one decides the underlying moral question, someone will always be harmed. For instance, if the answer to Professor Foot's question is, "No, it is never moral to take a human life," and one does not push the switch on moral grounds, then the five workers will lose their lives. Conversely, one innocent person, who was safe before, will be harmed by those who answer that it is morally permissible or required on consequentialist or utilitarian grounds to throw the switch and divert the trolley.

B. *Trolley Problem No. 2: The Fat-Man Variation*

Another academic philosopher, Judith Jarvis Thomson, formulated a different version of the trolley problem as follows:

> [Y]ou are standing on a footbridge over the trolley track. You can see a trolley hurtling down the track, out of control. . . . [Y]ou know of one certain way to stop an out-of-control trolley: Drop a really heavy weight in its path. But where to find one? It just so happens that standing next to you on the footbridge is a fat man, a really fat man. He is leaning over the railing, watching the trolley; all you have to do is to give him a little shove, and over the railing he will go, onto the track in the path of the trolley. Would it be permissible for you to do this?[9]

Compared to the original version of the problem, the fat-man hypothetical raises an even deeper moral dilemma. Even if one were willing to sacrifice one life to save five lives in the standard version of the trolley problem, most would agree that it is morally and legally wrong to push the fat man in Professor Thomson's version of the problem, even though the moral arithmetic (one life versus five lives) is the same.[10]

Since these two versions of the trolley problem were introduced, many

---

8. The Author borrows this term from Guido Calabresi and Philip Bobbitt's classic book. *See generally* GUIDO CALABRESI & PHILIP BOBBITT, TRAGIC CHOICES 149–91 (1978).

9. Thomson, *Trolley Problem*, *supra* note 4, at 1409. This version of the trolley problem takes away choice (a) and forces the decision between choices (b) and (c) from Part I of this Article.

10. Researchers have found that subjects presented with the fat-man hypothetical find the prospect of pushing a person to his death "more emotionally salient than the thought of hitting a switch that will cause a trolley to produce similar consequences, and it is this emotional response that accounts for people's tendency to treat these cases differently." Joshua D. Greene et al., *An fMRI Investigation of Emotional Engagement in Moral Judgment*, 293 SCI. 2105, 2106 (2001).

scholars have attempted—without success, given the lack of consensus in the literature—to identify some relevant moral distinction between "killing" and "letting die."[11] For this Article, however, the Author shall not survey this literature or attempt to explain this metaphysical distinction. Like the luminiferous aether in Newtonian physics, the relevant moral distinction between both versions of the trolley problem—a fundamental moral distinction that so many scholars have been searching for in vain through the years—is a nonexistent entity.[12]

Before proceeding, it is important to note that although there are two versions of the trolley problem, the Author shall refer to both versions of the problem in the singular as the trolley problem because both problems present a tragic choice under conditions of time scarcity. In short, the person at the switch or the bystander on the bridge must not only make a life-and-death decision but must also make this decision quickly.

Next, this Article considers the trolley problem from a legal perspective, but as noted, the legal dimension of this problem is just as intractable and difficult as the moral dimension.

III. LEGAL TROLLEY PROBLEMS

Trolley problems not only pose deep and difficult moral questions, they also share a legal dimension testing the outer limits of the legal defense of necessity.[13] In particular: when, if ever, is it legally justified to take an innocent person's life?[14]

Necessity was a defense at common law and is still a valid legal defense today, but as one legal scholar noted, "There is considerable disagreement in both federal and state courts, as well as in modern legal scholarship, over the defense's elements, and over its contours and purpose."[15] For example, the two seminal criminal cases presenting this issue, *United States v. Holmes*[16] and *The Queen v. Dudley and Stephens*,[17] take different approaches to the defense of necessity.

---

11. *See* Gorr, *supra* note 4, at 98 (noting that most explanations suffer from "an insistence on finding a moral difference where there just doesn't seem to be one").

12. *See* Thomson, *Trolley Problem*, *supra* note 4, at 1408 ("But other things being equal, it is not morally required of us that we let a burden descend out of the blue onto five when we can make it instead descend onto one.").

13. *See* MODEL PENAL CODE §§ 3.02, 3.04, 3.05 (1985).

14. *See* Gorr, *supra* note 4, at 92–93.

15. Stephen S. Schwartz, *Is There a Common Law Necessity Defense in Federal Criminal Law?*, 75 U. CHI. L. REV. 1259, 1260 (2008); *see also* Edward B. Arnolds & Norman F. Garland, *The Defense of Necessity in Criminal Law: The Right to Choose the Lesser Evil*, 65 J. CRIM. L. & CRIMINOLOGY 289, 291 (1974).

16. United States v. Holmes, 26 F. Cas. 360 (No. 15,383) (C.C.E.D. Pa. 1842).

17. The Queen v. Dudley & Stephens, [1884] 14 Q.B.D. 273 (Eng.).

The *Holmes* court recognized that there are some limited situations in which taking an innocent person's life is legally justified under the necessity defense, when "all ordinary means of self preservation have been exhausted. The peril must be instant, overwhelming, leaving no alternative but to lose our own life, or to take the life of another person."[18] Thus, in this expansive view of the necessity defense, there are situations in which the act of taking an innocent life is legally justified: "[W]hen this great 'law of necessity' does apply, and is not improperly exercised, the taking of life is devested of unlawfulness."[19]

The *Dudley* case, by contrast, takes a narrow view of the necessity defense, refusing to recognize any principle of law that entitles persons to take the life of an innocent person to save their own.[20] In the immortal words of Lord John Coleridge, who delivered the court's judgment in this case:

> Who is to be the judge of this sort of necessity? By what measure is the comparative value of lives to be measured? Is it to be strength, or intellect, or what? It is plain that the principle leaves to him who is to profit by it to determine the necessity which will justify him in deliberately taking another's life to save his own.[21]

Under this narrow view, the necessity defense cannot apply to the taking of human life under any circumstances; the act of killing is never legally justified.

Why, then are trolley problems so problematic—not just for moral philosophers but for judges and legal scholars as well? One possibility is that trolley problems require fundamental value judgments about right and wrong and about the proper way to value human lives, a judgment over which deep disagreements are bound to flourish.[22] Or, in the words of Professors Edward

---

18. *Holmes*, 26 F. Cas. at 366, *quoted in* Arnolds & Garland, *supra* note 15, at 295.

19. *Id.* Notice, however, that this judicial quotation begs the all-important question: *when* does the defense of necessity apply, and *when* is this defense properly exercised? The *Holmes* court opined:

> An illustration of this principle occurs in the ordinary case of self-defense against lawless violence, aiming at the destruction of life, or designing to inflict grievous injury to the person; and within this range may fall the taking of life under other circumstances where the act is indispensably requisite to self-existence.

Id. But outside of the context of self-defense, *United States v. Holmes* did not trace the contours of the necessity defense to homicide. *See id.*

20. *Dudley*, 14 Q.B.D. at 287.

21. *Id.*

22. *Cf.* MODEL PENAL CODE, *supra* note 13, at §3.04, p. 46–47 (discussing the limitations of the necessity defense with regard to the use of deadly force); MODEL PENAL CODE, Comment (Tent. Draft No. 8, 1958) at 8 ("The life of every individual must be assumed in such a case to

Arnolds and Norman Garland, "That the decisions in *Holmes* and *Dudley* should come out on opposite sides of the question whether one innocent life can be sacrificed to save two, illustrates how difficult the 'value asserted' question can be."[23]

Given the moral dimension of both versions of the trolley problem, the legal controversy over the scope of the necessity defense should come as no surprise. Trolley problems are moral dilemmas by nature, and in the domain of ethics and morality, there is no way of testing or falsifying the "rightness" or "wrongness" of a person's deep moral intuitions, including one's intuitions about trolley problems.

Or is there? Both versions of the trolley problem present a tragic choice under conditions of time scarcity. Because scarcity is, at bottom, an economic problem, the Author proposes a market solution to the trolley problem. Specifically, the Author proposes a second-order thought experiment that imagines what hypothetical bargain or voluntary agreement the relevant parties to the trolley problem would reach if they were able to bargain or conduct an auction from behind a veil of ignorance.

IV. TROLLEY AUCTIONS

Here, the Author proposes a novel method of solving both versions of the trolley problem. In summary, the Author would solve these trolley problems by conducting an auction from behind a veil of ignorance. Although neither of these ideas is new—the idea of bargaining from behind a veil of ignorance goes back to John Harsanyi and John Rawls,[24] while the idea of allocating legal rights through auctions is attributed to Ronald Coase and is often referred to as the Coase Theorem[25]—both the combination of the two ideas into a single package and the application of this new framework to both versions of the trolley problem are new.

A. *Auctions Behind a Veil of Ignorance*

Both the Coase Theorem and the veil of ignorance—like the original pair of trolley problems themselves—are ingenious thought experiments or what-if

---

be of equal value and the numerical preponderance in the lives saved compared to those sacrificed surely establishes an ethical and legal justification for the act.").

23.   Arnolds & Garland, *supra* note 15, at 296.

24.   *See* RAWLS, *supra* note 3, at 102–05, 118–23; Harsanyi, *Cardinal Welfare*, *supra* note 3, at 314–16.

25.   *See* Coase, *Social Cost*, *supra* note 3, at 131–44; *see also* F. E. Guerra-Pujol, *Modelling the Coase Theorem*, 5 EUR. J. LEGAL STUD. 139, 140 (2013) ("Remove or disprove the Coase Theorem, and the economic approach to law is reduced to intellectual rubble or just another untestable or normative legal theory.").

propositions.[26] Let's start with the Coase Theorem. Although Coase's Theorem is not really a theorem or a mathematical or logical proof in the formal sense, it does make predictions or educated guesses about what the parties to a given conflict would do in a perfect world—a frictionless world of "zero transaction costs"[27]—if they could call a quick time-out and negotiate a private peace treaty or auction off the legal right to win the conflict.[28]

Next, consider the possibility of bargaining or conducting such an auction from behind a veil of ignorance.[29] In the context of the trolley problem, a person negotiating or bidding behind a veil of ignorance would not be aware of which version of the trolley problem she would find herself in, nor would she know what role or position she would occupy in either version of the problem.[30] In other words, a person negotiating or bidding behind a veil of ignorance would have an equal probability—or, better yet, an uncertain probability—of being a worker on the main track, the person at the switch, the worker on the side track, the bystander on the bridge, or the fat man.

So, what happens if the logic of auctions and the logic of the veil of ignorance are applied to either version of the trolley problem?

Consider the first version of the trolley problem. There are seven parties to this conflict. On the one side of this moral equation is the person at the switch, who must make the difficult decision whether to leave the trolley on the main track where five workers will be harmed or divert it to a side track where only one worker will be harmed. On the other side of this moral balance are a total of six workers who are all potentially at risk from the runaway trolley, depending on

---

26. *See generally* ROBERT NOZICK, ANARCHY, STATE, AND UTOPIA x (1974) (reluctantly adopting the style of "much contemporary philosophical work in epistemology or metaphysics" characterized by "puzzles, abstract structural conditions, challenges to find another theory which fits a specified range of cases, startling conclusions, and so on"). For some additional thoughts on the role of thought experiments in law, see F. E. Guerra-Pujol, *Confessions of a Latino Law Professor: My Scholarly Autobiography* 14, SOCIAL SCIENCE RESEARCH NETWORK (SSRN) (Jan. 1, 2012), http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1516485 ("[A] thought-experiment, especially an outrageous and far-fetched one, is often capable of producing the most rewarding and uncommon insights.").

27. *See* Elodie Bertrand, *What Do Cattle and Bees Tell Us about the Coase Theorem?*, 31 EUR. J.L. & ECON. 39, 44 (2011).

28. For Professor Coase's original statement of his conjecture, see Coase, *Social Cost*, *supra* note 3, at 15; see also Stephen G. Medema, *Legal Fiction: The Place of the Coase Theorem in Law and Economics*, 15 ECON. & PHIL. 209, 211–12 (1999).

29. *See* RAWLS, *supra* note 3, at 118–23.

30. The purpose of the veil of ignorance hypothetical is to "nullify the effects of specific contingencies which put men at odds" and enable parties to "evaluate principles solely on the basis of general considerations." *Id.* at 118.

whether the switch is pulled. But because of the veil of ignorance, none of the players knows their role ahead of time.

Given this counterfactual world, imagine what would occur if these seven unfortunate souls could call a Coasean time-out to take part in a Coasean auction behind a veil of ignorance. In truth, since this a second-order thought experiment,[31] the outcome of such a hypothetical auction is not obvious, unless a probabilistic approach to the trolley problem is taken.[32]

In summary, from a purely probabilistic perspective, each person has a five-in-seven chance of being a worker on the main track, a one-in-seven chance of being the person at the switch, and a one-in-seven chance of being the worker on the side track in the standard version of the trolley problem. Because each person has a greater chance of ending up on the main track than on the side track, each person would presumably agree to throw the switch in the event he were to find himself in this position in the moment of truth.

Moreover, this same probabilistic solution applies to the fat-man variation of the trolley problem. As before, there are seven parties in the fat-man hypothetical: the bystander on the bridge, the hapless fat man, and the five railroad workers at the end of the track. Again, assume the parties to this conflict could call a Coasean time-out and conduct a Coasean auction behind a veil of ignorance—and again, each person in the fat-man hypothetical has a five-in-seven chance of being a worker at the end of the track, a one-in-seven chance of being the bystander on the bridge, and a one-in-seven chance of being the fat man. As in the standard version of the trolley problem, because five-sevenths is greater than one-seventh, each person would presumably agree to push the fat man off the bridge if he were to find himself in the role of the bystander because such a decision triples each party's likelihood of survival—six players survive if the fat man is pushed, while only the fat man and the bystander will survive if the fat man is not pushed.[33]

---

31. That is, the original first-order thought experiment is the standard version of the trolley problem itself. The second-order thought experiment is the Coasean auction, which this Author imagines would be conducted behind a Rawlsian veil of ignorance.

32. For further work on probabilistic payoffs in Coasean auctions, see Guerra-Pujol, *supra* note 25, at 146–57.

33. For purposes of this second-order thought experiment, the Author ignores an important complication that occurs in the fat-man hypothetical: pushing the fat man off the bridge might expose one to criminal and civil sanctions. Notice, though, that the possibility of such criminal and civil sanctions is itself probabilistic in nature. For a discussion of the legal dimension of the trolley problem, see *infra* Part III.

B. *Auctions Without a Veil of Ignorance*

Will removing the veil of ignorance change the outcome of the auction? The answer to this question depends on whether the five workers are able to form a stable coalition and overcome free-rider problems. Assuming that the five workers are able to overcome free riding,[34] it is most likely that the five will pool their resources to bribe the person at the switch to divert the trolley onto the side track or bribe the bystander on the bridge to push the fat man.

But where does this leave the seventh man—the sole worker stranded on the side track or the fat man standing on the bridge—in both versions of the trolley problem? In short, unless either of them is able to outbid or outbribe the five workers in peril and offer a larger payment to the person at the switch or the bystander on the bridge, the outcome is the same as when there is a veil of ignorance. In either case, the worker on the side track and the fat man on the bridge will most likely be outbid by the five workers on the main track[35]

Of course, this analysis assumes that the person at the switch or the bystander on the bridge will actually accept the Coasean bribe from the five workers.[36] But the Coasean analysis of the trolley problem is nevertheless attractive where moral intuitions about right and wrong diverge—that is, when no "reflective equilibrium" about a given moral dilemma is possible.[37] Rather than getting bogged down in an intractable moral discussion about right and wrong or pretending that some reflective equilibrium is possible, which requires asserting that everyone who can

---

34. This is a very big and perhaps unrealistic assumption. *But see* Gerald Marwell & Ruth G. Ames, *Economists Free Ride, Does Anyone Else? Experiments on the Provision of Public Goods*, 15 J. PUB. ECON. 295, 307 (1985) (summarizing the results of their experimental studies: "over and over again, in replication after replication, regardless of changes in a score of situational variables or subject characteristics, the strong version of the free rider hypothesis is contradicted by the evidence").

35. However, the result is different if the single player happens to be fabulously wealthy, which would undermine the usefulness of that particular auction as a tool for allocating legal rights. *See* Coase, *FCC, supra* note 3, at 19. When no meaningful bargaining can take place, "the final allocation of resources will indeed depend on to whom the rights are initially assigned." Medema, *supra* note 28, at 215. For further discussion, see *infra* Part V.A.

36. The worker at the switch or the bystander at the bridge may simply refuse to accept responsibility for causing the death of another person, even to save five even when offered a bribe. Coasean theories are poor at incorporating these types of moral objections; "the nasty little implication of the Coasean perspective is that traditional notions of causation and harm go out the window." Medema, *supra* note 28, at 217.

37. "From the standpoint of moral theory, the best account of a person's sense of justice is not the one which fits his judgments prior to his examining any conception of justice, but rather the one which matches his judgments in reflective equilibrium." RAWLS, *supra* note 3, at 42. For an extended discussion of the idea of reflective equilibrium, see *id.* at 42–45.

think rationally enough will reach the same conclusion (usually one's own), Coase's conjecture instead poses a question that can, in principle, be answered with testable hypotheses.

In addition, this alternative approach to both trolley problems reveals a new insight that both the philosophical and legal approaches miss: the importance of free riding in deciding the outcome of trolley problems. That is, the fate of the worker on the side track and the fate of the hapless fat man depend not so much on hair-splitting moral distinctions but rather on whether the five workers can overcome mundane free-rider problems when forming their coalition. If not, the worker on the side track and the fat man stand a good chance of offering the decisionmaker a better bribe and staying alive.

This Coasean analysis thus makes trolley problems more tractable. It demonstrates that both versions of the trolley problem are really free-rider problems.[38] Broadly speaking, a free rider is a "freeloader"—someone who enjoys the benefits of an activity without paying for it. Thus, free riding is a form of "group theft," but is more subtle than outright theft—it occurs through shirking when free riders withhold effort or resources or through cost shifting when free riders externalize the costs of their activities on others.[39]

Originally, the term "free rider" was based on the classic example of someone using public transportation without paying the fare.[40] If too many people ride for free, the system will collapse because it will not have enough money to operate. Another example of free riding is tax evasion. When people do not pay their share of taxes, they are free riding off the contributions of others by enjoying the public benefits that come with taxes—such as roads, water treatment plants, and fire services—without the burden of paying the taxes.[41]

In general, free riding is most likely to occur in large, anonymous groups and

---

38. It is worth noting that there is some debate about the empirical significance of free riding in the real world. *See, e.g.*, Marwell & Ames, *supra* note 34 (concluding that experimental studies show much less evidence of observable free riding than one would expect based on the predictions of economic theory).

39. *See id.* at 296–99 (discussing free rider hypotheses and methods for operationalization of the theory in experimental and observational studies).

40. *See* Timothy Taylor, *More on the Origins of the Free Rider Idea*, CONVERSABLE ECONOMIST (Oct. 11, 2014), http://conversableeconomist.blogspot.com/2014/10/more-on-origins-of-free-rider-idea.html.

41. *See* Timothy Taylor, *How the Free Rider Idea Evolved*, CONVERSABLE ECONOMIST (Sept. 16, 2014), http://conversableeconomist.blogspot.com/2014/09/how-free-rider-idea-evolved.html (concluding that "free riding offers a broad explanation for why government needs power of compulsory taxation and regulation").

in one-off interactions.[42] Thus, the free-rider problem could still pose an insurmountable obstacle to Coasean bargaining in situations involving a multitude of group members or in one-shot prisoners' dilemmas.[43] In these situations, the larger and more anonymous the relevant group is, the more likely free-rider problems will arise.[44] In either version of the trolley problem—where the size of the groups are small and the stakes are high (life and death)—one could assume that the five workers would be able to overcome any free-rider problems and that each worker would chip in a fair share to the common bribe fund.

Of course, these Coasean conjectures do not inform the underlying morality of these outcomes and bribes. Coase's Theorem is a non-normative test. It is not designed to answer intractable moral questions. Instead, its purpose is to imagine how the parties in a given conflict, situation, or moral dilemma would actually behave if they could call a Coasean time-out; specifically, Coase's Theorem asks what kind of auction the parties would design or what kind of peace treaty the parties would draw up. But the resulting outcome or bribe itself poses deep moral problems—problems that Coasean analyses are unable to answer. In short, the possibility of an auction without a veil of ignorance raises several moral red flags.

V. MORAL RED FLAGS

Although auctions offer a novel solution to the trolley problem, auctions without a veil of ignorance are morally problematic for several reasons. In fact, ever since Ronald Coase introduced his so-called theorem in 1960,[45] scholars have mounted a moral counterattack, offering a wide variety of criticisms of Coase's Theorem and of economic analysis of law generally.[46] The following identifies and

---

42. For instance, Elinor Ostrom and other researchers have found that social norms and institutions can limit the extent of free riding through the imposition of informal sanctions on those who do not contribute or those who take more than their share from the common pool. *See, e.g.*, Elinor Ostrom, *Collective Action and the Evolution of Social Norms*, 14 J. ECON. PERSP. 137, 137–58 (2000).

43. *See id.* at 142 (noting that a large number of one-shot prisoners' dilemma experiments show "initial cooperation rates ranging from 40 to 60 percent").

44. *See* Marwell & Ames, *supra* note 34, at 302–04.

45. In his autobiography, economist George Stigler presents a vivid reconstruction of the first time Coase presented his argument at the University of Chicago. *See* GEORGE J. STIGLER, MEMOIRS OF AN UNREGULATED ECONOMIST 75–80 (1988).

46. This literature is vast. For a small sample of this literature, see, e.g., CALABRESI & BOBBITT, *supra* note 8; JULES L. COLEMAN, MARKETS, MORALS AND THE LAW 311–42 (1988); Ronald Dworkin, *Why Efficiency? A Response to Calabresi and Posner*, 8 HOFSTRA L. REV. 563, 566–73 (1980); Russell Hardin, *The Morality of Law and Economics*, 11 L. & PHIL. 331, 358–60 (1992); Mark Kelman, *Consumption Theory, Production Theory, and Ideology in the Coase Theorem*, 52 S. CAL. L. REV. 669, 678–95 (1979); Margaret Jane Radin, *Market-*

considers the smaller subset of moral concerns that might apply to the Coasean approach to the trolley dilemma. In summary, an auction without a veil of ignorance raises at least three moral red flags:

(a) The well-known problem of "wealth effects"—will the person or group with the most money always win the Coasean auction and thus distort the allocation of legal rights?[47]

(b) The subsidiary problem of "commensurability"—are all human lives worth the same, and if not, how should economic value be assigned to a person's life?[48]

(c) The problem of "repugnant markets"—the Coasean solution presupposes an explicit market in human life, but is such a life-or-death auction repugnant and unfathomable?[49]

A. *Wealth Effects*

The moral critique of wealth effects is this: why should the person with the largest bribe win the Coasean auction?

Recall that in either version of the trolley problem, the five workers on the main track—assuming they are able to band together and overcome free-rider problems—will most likely save their skins by outbidding the sixth worker on the side track or the fat man on the bridge, unless the sixth player is somehow able to outbid the five workers on the main track. In other words, in a perfect or frictionless world of zero transaction costs, the outcome of either version of the trolley problem depends on who is able to offer the largest bribe. But, here is the rub: from a moral standpoint, why should the person with the largest bribe win?[50]

In non-life-threatening situations, the answer is easy enough: property rights—and all scarce resources generally—should gravitate toward those who value such rights the most.[51] In Coase's original cattle-trespass problem, for

---

Inalienability, 100 HARV. L. REV. 1849, 1864–70 (1987).

47. *See infra* Part V.A.

48. *See infra* Part V.B.

49. *See infra* Part V.C.

50. This question echoes the call for equal opportunities for survival that was set forth in *Holmes*: when choices must be made regarding who survives and who does not, "there should be consultation, and some mode of selection fixed, by which those in equal relations may have equal chance for their life." United States v. Holmes, 26 F. Cas. 360, 367 (No. 15,383) (C.C.E.D. Pa. 1842).

51. *See, e.g.*, Richard A. Posner, *Utilitarianism, Economics, and Legal Theory*, 8 J. LEGAL STUD. 103, 122, 127 (1979); *see also* Coase, *FCC*, *supra* note 3, at 19 ("First of all, it must be observed that resources do not go, in the American economic system, to those with the most

example, the party willing to pay more for the legal right being auctioned (i.e. the right to trespass or the right to be free from trespass) will outbid the other party for the right in question.[52] The problem with this analysis in the context of the trolley problem, however, is that the relevant right in question is the right to life. How does one go about assigning a value to human life? This question leads to our second moral critique of Coase's Theorem as applied to the trolley problem: the problem of commensurability.

B. *Commensurability*

One method for assigning a value to human life is by finding out, either directly or indirectly, how much people are willing pay to avoid a small risk of death.[53] This method involves a two-step process. First, find out how much extra pay workers in a certain industry demand in compensation for assuming a specific on-the-job risk, or in the alternative, find out how much extra money people are willing to pay for enhanced safety features on consumer products.[54] Next, divide this dollar amount by the probability of harm—the probability of being injured on the job or the probability of being injured using a product. The resulting number, expressed in dollars, gives the value of one "statistical life." For example, if people are willing to pay a maximum of $1 to avoid a risk of death of one in a million, then the value of a statistical life is $1 million because, in the aggregate, that corresponds to a willingness to pay $1 million to avoid the death of one unidentified person.[55]

There are two problems with this method, however. One is psychological; the other is inescapably moral. The psychological problem is that the value of avoiding or accepting a given risk might vary in some nonlinear and nonrational fashion with the probability or magnitude of that risk.[56] The risk range used in most value-of-life studies, for example, is somewhere between a one-in-10,000 chance

---

money but to those who are willing to pay the most for them.").

    52.   Coase, *Social Cost*, *supra* note 3, at 2–8.

    53.   *See, e.g.*, RICHARD A. POSNER, CATASTROPHE: RISK AND RESPONSE 165–71 (2004).

    54.   Judge Richard Posner notes that this data is readily available.

> From data on wage premia in dangerous occupations, the response of housing prices to proximity to hazardous sites, seat-belt use, cigarette smoking, and other behavior toward risk, actual or perceived, of death, economists have calculated how much money the average person would demand to incur a given such risk.

Id. at 165.

    55.   *Id.* at 165–66.

    56.   *See id.* at 166 ("[T]here is no reason to think that the relation between the risk of death and the perceived cost of the risk is linear. Quite the contrary.").

of death and a one-in-100,000 chance.[57] But what if, as in either version of the trolley problem, the relevant risk is much lower or much higher than this range? In the words of Judge Richard Posner, a person might "demand an infinite amount of money" to take a one-in-two risk of death, while at the same time completely ignoring very small risks of one in a billion or even one in several million—and thus valuing his or her own life at zero at that end of the risk continuum.[58]

The trolley problem presents a textbook illustration of this psychological problem. Simply put, the relevant risks in the trolley problem are too high. If you were one of the five workers on the main track, for example, how much money would you be willing to pay to avoid the risk of a runaway trolley? An infinite amount? In either version of the problem, the death of the five workers will occur with absolute certainty unless the lever is released or the fat man is pushed. If you were the fat man or the solitary worker on the side track, how much money would you be willing to accept from the five workers to avoid death?

Contrary to the logic of the Coase Theorem, it is understood that no person will accept any amount of money in exchange for a 100 percent probability of death. This problem may be avoidable by setting up a Rawlsian auction, which would require the parties involved in either trolley problem to bid for the right to life from behind a veil of ignorance, not knowing what role they will play in the trolley problem.[59] But in reality, no amount of sophisticated "mechanism design" avoids the underlying psychological problem because a hypothetical Rawlsian auction conducted from behind a veil of ignorance does not alter the relevant risks in the trolley problem.[60]

Another possible solution to this psychological problem is to reframe the risk involved. Specifically, instead of asking the five workers to consider the certain risk of death to them in either trolley scenario, they could be asked to consider the much lower risk (in the abstract) of there being a runaway trolley in the first place. But there is an inescapable moral problem with this technical reframing-of-the-risk solution. But the inescapable moral problem is a deeper one. Whatever value is assigned to the lives of the five workers, the fundamental moral problem posed by the two trolley problems cannot be avoided: when, if ever, is it morally permissible

---

57. *See id.* at 167 (observing that "almost all of the studies involve risks [of death] that are between [.000008] and .00014, with the average being .0000915").

58. *Id.* at 166–67.

59. *See supra* Part III.

60. That is, even behind a veil of ignorance, parties would pay any amount of money to avoid death and would not accept any amount of money as a bribe "for the excellent reason that if they were dead they wouldn't get to enjoy the money they'd been given to assume the risk." POSNER, *supra* note 53, at 166.

to kill or let die?

Put another way, the moral problem is that the process of assigning a value to human life suggests that, all things being equal, the lives of the five workers on the main track will be worth at least five times more than either the life of the worker on the side track or the life of the fat man on the bridge. But this simple arithmetic ignores the complex moral dimensions of the trolley problem. That is, even if the lives of the five workers are statistically worth five times more than the life of the hapless worker on the side track or the fat man, this fact alone does not provide what the morally correct action should be in either version of the trolley problem.[61]

In short, the Coasean solution to the trolley problem, although imaginative and intriguing, ends up neglecting the central moral dilemma posed by the runaway trolley altogether.[62] But this conclusion raises new questions. Why is the moral dilemma in either version of the trolley problem inescapable? Why does the Coasean approach appear to fall short from a moral perspective? The Author conjectures that the problem is not so much with the Coase Theorem or economic analysis per se. The problem is with us.

C. *Market Aversion*

The most salient moral objection to the Coasean approach to trolley problems is market aversion, or the problem of repugnant markets.[63] Broadly speaking, a repugnant market can be defined as an area of commerce that is considered by most to be outside the range of morally permissible market transactions.[64] In summary, some real-world examples of repugnant markets include the sale of horsemeat for human consumption, price gouging after natural disasters, prediction markets for terrorist events, and the purchase and sale of human organs.[65]

In other words, people often find certain commercial transactions, even purely voluntary and consensual ones, to be repugnant or revolting and thus

---

61. *See* Thomson, *Trolley Problem, supra* note 4, at 1408 ("Not just any distributive intervention is permissible: It is not in general morally open to us to make one die to save five.").

62. Indeed, this same critique could be made of economic analysis at large, at least when human lives are at stake, which is almost always.

63. *See generally* Alvin E. Roth, *Repugnance as a Constraint on Markets*, 21 J. ECON. PERSPS. 37, 38–40 (2007); *see also* MARTHA NUSSBAUM, HIDING FROM HUMANITY: DISGUST, SHAME AND THE LAW 134–47 (2004).

64. See Roth, *supra* note 63, at 38.

65. *See id.* at 39 tbl.1 (listing "markets in which some transactions are, or were once, repugnant").

morally wrong.[66] In the words of Professor Alvin Roth,

> When my colleagues and I have helped design markets and allocation procedures, we have often found that distaste for certain kinds of transactions can be a real constraint on markets and how they are designed, every bit as real as the constraints imposed by technology or by the requirements of incentives and efficiency.[67]

There is a certain class of transactions that is considered off-limits or morally out-of-bounds, and feelings of market aversion, if widely held, can operate as an ever-present endogenous constraint on a certain class of voluntary transactions.[68]

Before proceeding, it is worth noting that the market-aversion problem poses an important subsidiary question: when is the use of law (coercion) justified in order to prevent repugnant trades from taking place? Although one could argue that people should abstain from transacting in repugnant markets on moral grounds when such transactions are morally wrong, one could also argue that the use of law to deter or punish purely voluntary and consensual transactions is also morally wrong.[69]

Putting aside this question about the proper role of law in regulating repugnant markets, why is market aversion relevant to trolley problems? In short, market aversion operates as a real constraint on the operation of the Coase Theorem, thus limiting its scope and domain. To the extent most people would find any kind of explicit bargaining or bidding in trolley problems to be morally repugnant or morally wrong, the Coasean solution breaks down. Why? Because even when transactions costs are zero and property rights are well-defined (as Coase's Theorem requires)—and thus even when a Coasean bargain or auction is theoretically possible—market aversion will prevent most trades from taking place.[70]

---

66. For an example of a repugnant market from the world of fiction, see Enrique Guerra-Pujol, *Buy or Bite?*, *in* THE ECONOMICS OF THE UNDEAD: ZOMBIES, VAMPIRES, AND THE DISMAL SCIENCE 123, 125–29 (Glen Whitman & James Dow eds., 2014) (proposing the legalization of blood markets for vampires).

67. Roth, *supra* note 63, at 38; *see also* Radin, *supra* note 46, at 1850 (citing Elizabeth Landes & Richard Posner, *The Economics of the Baby Shortage*, 7 J. LEGAL STUD. 323 (1978)).

68. It is worth noting, however, that not all forms of market aversion rise to the level of repugnance or revulsion or raise a moral concern. In these weaker cases of aversion, "a milder word might be more apt—some transactions may be called distasteful, inappropriate, unfair, undignified, or unprofessional." Roth, *supra* note 63, at 40.

69. *See generally* JOHN STUART MILL, ON LIBERTY (1859) (defending liberty, a priori, against encroachment for politically or socially desirable ends).

70. Here, again, economists may be the sole exception. "Although economists see very few tradeoffs as completely taboo, noneconomists often decline to discuss tradeoffs at all,

VI. CONCLUSION

Although auctions offer a novel solution to the deep moral dilemma posed in the trolley problem, the market approach assumes away the very real problem of market aversion. The Coase Theorem, for example, has nothing at all to say about repugnant markets.[71] In addition, the wealth-effects and commensurability problems might also call into question the morality of a market solution to the trolley problem.

Nevertheless, we might be able to solve the problems of market aversion, commensurability, and wealth effects by introducing a veil of ignorance to the market solution. With a veil of ignorance, the parties to the trolley problem do not know their role or position ahead of time and thus must weigh the interests of all the parties when deciding upon the correct course of action. Ironically, however, the outcome of a trolley-problem discussion behind a veil of ignorance is probably the same as that of an auction without a veil of ignorance.

---

preferring to focus on the repugnance of transactions like organ sales." Roth, *supra* note 63, at 54.

71.   *Cf.* LUDWIG WITTGENSTEIN, TRACTATUS LOGICO-PHILOSOPHICUS 189 prop. 7 (C. K. Ogden trans., 2010) ("Whereof one cannot speak, thereof one must be silent.").

APPENDIX[72]

THE TROLLEY PROBLEM

A trolley is running out of control down a track. In its path are five
people who have been tied to the track by a mad philosopher. Fortu-
nately, you could flip a switch, which will lead the trolley down a differ-
ent track to safety. Unfortunately, there is a single person tied to that
track. Should you flip the switch or do nothing?
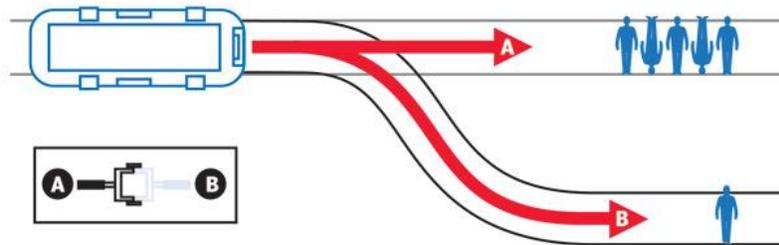
THE FAT MAN VARIATION

Another trolley is running out of control down another track. In its path
are five people who have been tied to the track by a different mad
philosopher. You are on a bridge under which the trolley will pass; you
can stop it by dropping a heavy weight in front of it. As it happens, there
is a very fat man next to you— your only way to stop the trolley is to push
him over the bridge and onto the track, killing him to save five. Should
you proceed?
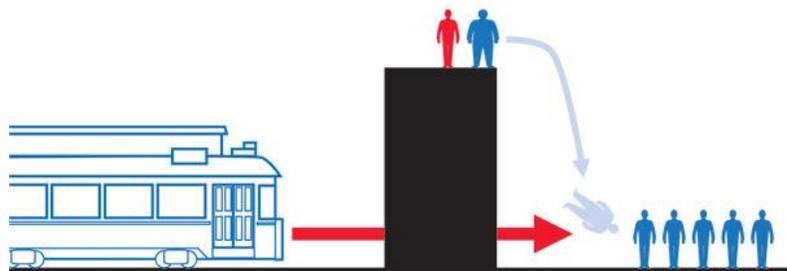
ANDREW BARR / NATIONAL POST

---

72. ~~Diagram from~~ Joseph Brean, *Explaining the 'Trolley Problem'*, NAT'IONAL POST
(Nov. 27, 2010), available at http://news.nationalpost.com/2010/11/27/explaining-the-trolley-
problem/.