Economics & Philosophy (2021), **37**, 260–283
doi:10.1017/S0266267120000310

CAMBRIDGE UNIVERSITY PRESS

ARTICLE

Contractualism and risk preferences

Tobey K. Scharding

Rutgers Business School – Newark and New Brunswick, Piscataway, New Jersey, USA
Email: tscharding@business.rutgers.edu

(Received 29 September 2017; revised 4 August 2020; accepted 7 August 2020; first published online 26 October 2020)

Abstract

I evaluate two contractualist approaches to the ethics of risk: mutual constraint and the probabilistic, *ex ante* approach. After explaining how these approaches address problems in earlier interpretations of contractualism, I object that they fail to respond to diverse risk preferences in populations. Some people could reasonably reject the risk thresholds associated with these approaches. A strategy for addressing this objection is considering individual risk preferences, similar to those Buchak discusses concerning expected-utility approaches to risk. I defend the risk-preferences-adjusted (RISPREAD) contractualist approach, which calculates a population's average risk preference and permits risk thresholds below that preference, only.

Keywords: risk; contractualism; Scanlon; risk-weighted expected utility (REU) theory

Introduction

This paper applies and challenges two recent, very promising contractualist approaches to the ethics of risk: what I call the *mutual constraint* (MC) approach (Lenman 2008) and what I call the *probabilistic, ex ante* (PEA) approach (James 2012; Frick 2015).[1] These approaches seem promising because contractualism is often regarded as the most fruitful strategy for addressing ethical questions about risk[2] and the MC and PEA approaches escape problems associated with previous interpretations of contractualism.[3] The point of this paper is to test the approaches' suitability for practical ethics: in particular, evaluating to how much risk economically motivated agents (e.g. businesses) may, ethically

[1]Frick's paper develops the view more fully than James's, as in his discussion of 'statistical' versus 'identified' lives (2015: 181, 212–218). The basic account as I discuss it herein, though, is common to the theorists.

[2]In addition to the defences of contractualism offered in the papers cited above, Hayenhjelm and Wolff's (2011) overview of the literature on the ethics of risk offers trenchant criticisms of utilitarian and deontological approaches and a (guarded) defence of contractualist perspectives on risk.

[3]In particular, Frick distinguishes the view from the accounts of Scanlon (1998), Reibetanz (1998) and Fried (2012*a*). In addition to James (2012), Frick believes that his approach resembles Kumar (2015) and Lenman (2008) (especially, to my eye, Lenman 2008: 116–117).

speaking, expose diverse populations as the agents work to produce benefits for the same populations.

The scope of the ethical questions that I address, then, is limited to risky actions that impose small probabilities of serious harm in exchange for widespread social benefits. My investigation excludes, in turn, risky actions that expose some people to risk so as to benefit others. Contractualism typically prohibits the latter actions (e.g. Scanlon 1998: 235; James 2012: 269–270); many risks of harm that secure benefits for the population exposed to the risks, though, seem permissible (e.g. Scanlon 1998: 209; James 2012: 273–274; Frick 2015: 187). In particular, contractualism permits risks (roughly) to which everyone in the population exposed to the risk has reason to consent. For my purposes in this paper, a person has reason to consent to a risk when the risk comports with aspects of that person's way of life that the person has explicitly chosen (Scanlon 1998; Oberdiek 2004; James 2012; Frick 2015; Kumar 2015). Someone who chooses to participate in a risky activity such as drag racing, for example, would have reason to consent to risk exposures (such as those associated with very fast speed limits) that involve similar levels of risk.

Consider, for example, a municipality's risky activity of repairing a local highway. This activity slightly increases the probability that motorists will have accidents during the repairs. The payoff is safer, smoother roads once the repairs are completed. *Ex ante* contractualism requires the population subjected to this risk to have reason to consent to it before the repairs begin: e.g. when the municipality is deciding whether to undertake the repair project. At that time, people do not know who, in particular, will be harmed. This orientation means that even the people who will be harmed in accidents have reason – the social benefits of better roads – to consent to the repair work.

There seems to be a significant difference, though, between having reason to consent to a very small risk of harm and actually experiencing serious harm: devastating injury, disability, death. This difference challenges the claim that people's *ex ante* reason to consent renders permissible their serious *ex post* harm. This is where *ex post* contractualism comes in. Because *ex post* contractualism assesses people's reasons to consent after the risky activity's effects are manifest, the people who are harmed know who they are and, accordingly, have reason to withhold consent from the risky activity. Such reasons to withhold consent, though, entail that *ex post* contractualism can forbid risks that have widespread social benefits. In this sense, *ex post* contractualism seems prone to forbid rightful actions (such as in the highway-repair example) but *ex ante* contractualism can permit wrongful risky actions that seriously harm a few people such as (perhaps) in the very-high-speed-limit example. Neither reliably distinguishes wrongful from rightful risks.

The MC approach addresses this issue by allowing the *ex ante* and *ex post* perspectives to constrain one another. Whereas it is unreasonable to forbid risky actions that benefit many people, it is also unreasonable to permit risky actions that seriously harm people. The approaches must work with one another, according to this view, to produce fair evaluations of risky actions. The PEA approach, in turn, requires people to have reason to consent to risks from an *ex ante* perspective while fully aware of the probability of harm. Thus, if repairing the highway is expected to benefit 1,000,000 people while seriously

https://doi.org/10.1017/S0266267120000310 Published online by Cambridge University Press

harming 100 of them, PEA contractualism permits this risky activity only if the 1,000,000 people all have reason to consent to the 100-in-1,000,000 (0.01%) risk of serious harm in order to secure the benefit. The approaches thus appear well suited to permit rightful risks that benefit many people while prohibiting wrongful risks that seriously harm a few people.

To test and strengthen these approaches to the ethics of risk, I apply them to a group of risks involving economic decision-making: related to manufacturing and marketing cars. I object that the MC approach fails to offer a principled way to resolve disputes between the *ex ante* and *ex post* perspectives. The PEA approach, in turn, fails to account for diverse *risk-preferences* about risky products such as cars. People's risk-preferences, on the understanding employed herein, capture how much risk they have reason to consent to. My objection to PEA contractualism is based on the intuition that most people would consent to a 0.01% risk of harm to secure a (particular) benefit – but not necessarily everyone. The possibility of diverse risk-preferences undermines PEA contractualism's ability to resolve this risk in a manner to which all the affected people have reason to consent.

I argue that these objections are decisive against the contractualist approaches, strictly speaking: universal reason to consent, in the manner required by these interpretations of contractualism, is too high a standard for establishing which risks are permitted. I respond, though, that the dialectic suggests a strategy for incorporating contractualist ideals in economic decision-making about risk. Inspired by Buchak's (2013) recent, very promising amendments to the expected utility (EU) approach to risk, I propose amending PEA calculations so as to (a) capture the portions of a population that share various risk-preferences, (b) multiply those portions by the numeric risk-preferences and (c) sum the portion-adjusted risk-preferences over the population. In my approach, risk-preferences-adjusted (RISPREAD) contractualism, only risky activities whose harm-benefit ratios are lower than the population's portion-adjusted (average) risk-preference are permitted. The RISPREAD approach shows that formal EU theory (e.g. in Buchak's (2013) unorthodox recent version) can fruitfully influence contractualism, at least in cases evaluating risks and benefits to a single population, as noted above, to improve contractualism's ability to resolve questions about risk. I discuss the ways in which this approach meets certain contractualist ideals and conclude by considering possibilities for realizing contractualist ideals more fully.

1. Contractualist approaches to risk

Philosophers began developing ethical accounts of risk in the 1970s to address whether actions that impose probabilities of harm on populations of people (rather than directly harming particular people) are permitted (Nozick 1974: 73–78; Thomson 1986: 173–191; for a summary, see Hayenhjelm and Wolff 2011: e27–e29). Their work built on long-standing traditions in mathematics and economics, particularly expected utility (EU) theory (Savage 1954; Buchak 2013: 1–9). Before explaining contractualist contributions to this tradition,

https://doi.org/10.1017/S0266267120000310 Published online by Cambridge University Press

I set forth the basics of EU theory and discuss a problem with it that motivated contractualists to offer their own approaches to the problem of risk.

In risky conditions, decision makers do not know which outcomes (x_1–x_n) will result from a selection of possible actions (*f*, *g*, *h*, …). They do know the probabilities (p_1–p_n) that different outcomes will result. According to EU theory, decision makers should choose the action that has the highest expected utility, where the expected utility is calculated by multiplying the utility of each possible outcome (x_1–x_n) by the probability that each outcome will occur, such that the sum of all probabilities ($p_1 + p_2 + \ldots + p_n$) is one. If *f* has an expected utility of 10 and *g* has an expected utility of 20 (and *f* and *g* are the only two possible actions), then decision makers should choose to do action *g*, according to EU theory.

The EU-theoretical outlook thus privileges quantifiable utilities over other values. Such other values include human life, political rights and the environment: whose worth does not appear to be limited to their utility and/or whose utility is difficult to calculate (e.g. Nussbaum and Sen 1995: 9–66). This privileging can be justified to the extent that quantifiable utilities offer a degree of precision and objectivity that the alternative values, listed above, appear to lack. Quantifiable utilities can, moreover, be manipulated using mathematical tools that allow them to be easily compared and ranked. It appears very difficult, in this sense, to use any methodology other than EU theory – and any values other than quantifiable utilities – to evaluate questions that involve risk (Hansson 1993, 2003; Fried 2012*a*; for a review of this issue, see Hayenhjelm and Wolff 2011).

Despite these advantages for making decisions about risky activities, however, EU theory is less adept at evaluating decisions from the standpoint of ethics. EU theory is not well suited to make ethical evaluations for the reason that it does not include ethically important (but, as noted above, difficult to quantify) values, such as the need to respect every person's humanity (e.g. Hill 2005). This concept entails, among other things, the reasoning capability that enables people to give or withhold consent to risk exposures. Respecting people's humanity, then, appears to require decision makers to refrain from exposing people to any risks except those to which they have reason to consent.

EU theory, though, can endorse risky actions even when some of the people whom the actions expose to (third-party) risks of harm lack reason to consent to the risks. Assume, for example, that the possible actions, *f* and *g*, discussed above, expose third parties to risks of harm. Even if the people whom *f* exposes to risk have reason to consent to the risk and the people whom action *g* exposes to risk do not have reason to consent, EU theory still recommends *g* over *f*: on the ground that *g*'s overall utility is twice that of *f*. Third parties' reasons to consent do not, in this sense, affect the EU-theoretical evaluation. Because people's reasons to give or withhold consent are connected to an ethically important value, respecting their humanity, ethicists have searched for an approach to risk that includes this value.

One ethical outlook that appears to demonstrate respect for humanity, while also offering resources to capture the advantages of quantifiable utilities for risk calculations, is contractualism. Contractualists began offering approaches to risk in the 1990s. The foundational view is Scanlon (1998). He permits risky actions only when they follow from principles that no one (who is subject to the

https://doi.org/10.1017/S0266267120000310 Published online by Cambridge University Press

principle) can 'reasonably reject' (Scanlon 1998: 153), a version of the contractualist claim that actions are permitted only when the people affected by the action have reason to consent to the action.[4] In order to determine whether people affected by an action can reasonably reject the principle that underlies that action, Scanlon (1998) instructs decision makers to reflect on the action from standpoints other than their own.[5] In such standpoints, people are to consider 'the weightiness of the burdens' and 'the importance of the benefits' that the risky action presents. They are not to consider, however, the probability that they would benefit from, or be burdened by, the risky action (Scanlon 1998: 208). If the burdens seem unreasonable in light of the benefits, people may reject the principle. If the burdens seem reasonable to gain the benefits, though, a person may not reject the principle – even if the person would bear a burden. In this manner, contractualism secures the advantages for risk calculations associated with quantifiable utilities while preserving respect for humanity. This ethical view, at its bones, offers an impartial way to think about risks in terms of possible burdens and benefits: focusing on the magnitude of the benefits and burdens rather than who, in particular, experiences them.

In contemplating the weightiness of the burdens and the importance of the benefits of a risky action, decision makers may consider the relative probabilities that the benefits and the burdens will result – so long as they take 'reasonable precautions' to reduce burdens. For example, decision makers should strive to ensure that any burdens result only 'by accident' and are never 'directly inflicted' by the action upon which the decision maker resolves. Scanlon notes, though, that 'the cost of avoiding all behavior that involves a risk of harm would be unacceptable' (Scanlon 1998: 209). In this sense, it seems reasonable to draw two (informal) principles from his account of contractualism's approach to the ethics of risk: decision makers (1) must be sure to take seriously the harm that their risky actions can cause but (2) should not be paralysed by the concern to avoid all harm.

These informal principles seem prima facie plausible. In ethically evaluating a risky action, $\boldsymbol{f}$, decision makers clearly should grapple with the harm $\boldsymbol{f}$ could cause to third parties. This principle follows naturally from respect for humanity. Respecting humanity, as defined above, requires decision makers to refrain from carelessly harming human beings in the sense that no person has reason to consent to be carelessly harmed so that others can benefit. EU theory also supports taking seriously the harm that $\boldsymbol{f}$ could cause. The action's expected utility, that is to say, offers information about its harms (and benefits): at least insofar as they affect the decision makers using EU theory to resolve what action to take.

The informal principle of avoiding paralysis, in turn, seems to follow from a concern to register the utilities associated with particular actions. As defined

[4]As the difference between these standards does not affect my argument, I use 'reason to consent' as well as 'reasonable rejection' in my discussion of the Scanlonian position.

[5]As an early commenter notes, Scanlon views moral reasoning as a 'hypothetical conversation, where all parties to the discussion share a common goal of reaching a consensus on principles ... for the general regulation of behavior' (Kumar 1999: 277).

https://doi.org/10.1017/S0266267120000310 Published online by Cambridge University Press

above, respect for humanity does not require decision makers to maximize utilities or even strive to increase them greatly. Human beings can benefit from increased utilities, however, and contractualism strives to capture these possible benefits. Whereas an exclusive focus on showing respect for humanity could incapacitate decision makers when an action promises great benefits while imposing a low probability of a terrible burden, the principle of avoiding paralysis permits decision makers to choose actions that risk harm.

Though helpful as rules of thumb, these informal principles clearly fail to provide firm guidance for distinguishing permitted from forbidden risks in actual cases. Since Scanlon developed his account, however, several commentators have investigated strategies for making contractualism's risk evaluations more precise. I offer a brief survey of efforts to distinguish permitted from forbidden risks. These include the (1.1) *actual harms* (AH), (1.2) *mutual constraint* (MC) and (1.3) *probabilistic ex ante* (PEA) approaches.

According to AH, contractualism should focus on the burdens and benefits that actually result from a risky action rather than those that are expected to result (Reibetanz 1998: 301–304). Contractualism's requirement that everyone have reason to consent to the risky action encourages this approach to the extent that AH gives voice to the people who have least reason to consent to the risky action (because they experience the harm associated with the action). Unless these people have reason to consent to the risky action, contractualism forbids the action. In this sense, the AH approach satisfies the first informal principle discussed above: taking seriously the harms that risky actions actually bring about. The approach faces a major explanatory burden, however, in explaining how it can justify life as most human beings currently experience it, much of which risks harm. In this sense, AH does not appear to satisfy the second informal principle.

The MC approach goes some distance to address this concern. It adopts the distinction between *ex ante* and *ex post* approaches to contractualism, broached above (Lenman 2008: 115–117). Unlike purely *ex post* approaches, such as AH, MC contractualism allows both *ex ante* and *ex post* perspectives to constrain what people have reason to consent to. People lack, for example, reason to consent to risky actions that will devastate (some of) the people implicated by them: on the ground that they would not choose personal devastation for themselves. Similarly, though, the viewpoint holds that people do have reason to consent to actions that have widespread benefits – even if the beneficial action harms the person. The justification for this claim is that most people regularly risk some personal harm as they pursue their interests. In this sense, the two viewpoints are to be used in cooperation, or friendly competition, with one another: working to establish and constrain the principles and actions that contractualism permits (Lenman 2008: 116).

There is only so far that the friendliness of this competition can extend, though. Under MC contractualism, the *ex post* perspective may prohibit *ex ante* endorsement of certain risky actions; it is unclear, though, how far *ex post* may constrain *ex ante*. Controversial cases effect a tug of war between the perspectives. Some compromise appears to be the correct

https://doi.org/10.1017/S0266267120000310 Published online by Cambridge University Press

solution but mutual constraint does not guide the perspectives in reaching it.[6] This is not necessarily an objection to the MC approach, of course: not if that is the best one can plausibly do. In truly difficult ethical problems, the solution reached using MC contractualism might be the correct one. But if there are ways of giving further guidance, they should be considered. In this spirit, I set forth an alternative strategy for including both *ex ante* and *ex post* perspectives on ethical questions about risk: probabilistic, *ex ante* (PEA) contractualism.

The PEA approach evaluates actions in light of their actual harms and benefits to various parties, as in the *ex post* perspective. It considers these effects, though, from an *ex ante* perspective: in terms of the probabilities of harm or benefit to which they expose various parties (James 2012: 267–268). In this sense, actions with low harm-benefit ratios – offering significant benefits in exchange for a small risk of death spread over many people – are permitted even if it is very likely that a few people will die (James 2012: 273–274). The PEA perspective can permit risky actions with widespread social benefits, even when they have deadly side effects.

The PEA approach thus offers an alternative to MC contractualism for combining *ex ante* and *ex post* viewpoints: preserving the benefits of full information while avoiding the perils of paralysis. Ideally, PEA contractualism offers a way to 'count the numbers without aggregating' (Frick 2015: 201). The approach 'counts the numbers' by including the number of people affected by a decision in the probabilities of harms and benefits it uses. It does not, though, aggregate those numbers: ethically evaluating a decision based on how many people the decision harms or benefits.[7] The moral importance of people's reasons to consent remains paramount. Probabilistic, *ex ante* contractualism thus preserves contractualism's distinctive normative characteristics while allowing it leeway to evaluate risks.

2. Contractualism and economic risks

Consider the PEA approach's facility in addressing the following example, modelled on a well known business ethics case study:

> *Gas-Tank Risks*: A car company is deciding between two gas tanks for its new line of inexpensive cars. One of the tanks can lose fuel in low-impact collisions, contributing to 150 serious injuries or deaths over 5 years. The other tank does not lose fuel but would require the company to reduce trunk size in order to keep sticker price constant. The company's market research shows that its customers (1,500,000 purchases over 5 years) strongly favour a larger trunk. May the company contribute to 150 deaths so as to prevent 1,500,000 people's frustration with a too small trunk?[8]

[6]In noting this point, I draw upon Fried's (2012*a*: 56–57) excellent discussion of a toxic waste dump that can never be made completely safe. The point also resembles James's (2012: 264–265) discussion of bridge, school and opera-house constructions.

[7]Aggregation is a feature of EU theory, a popular foil for contractualism. For a summary of the debate between these views, see Frick (2015: 175–176).

[8]Adapted from Birsch (1994: 5–10).

https://doi.org/10.1017/S0266267120000310 Published online by Cambridge University Press

Many people intuit that it is unethical for the company to choose the fuel-losing gas tank. Sacrificing 150 people's health or lives to save 1,500,000 people from (mere) frustration seems wrong. Actual-harms contractualism can support that reasoning. Each of the people who would be seriously injured or die in accidents caused by the fuel-losing gas tank have reason to reject the principle licensing the car company to use that gas tank, according to the AH approach.

Actual harms seems unable, though, to give voice to the other parties affected by the action: the 1,500,000 people whose interests would be frustrated, repeatedly, by their cars' small cargo spaces. They would be unable to transport the equipment they need for work and leisure, the quantity of groceries they need to sustain their lives (and/or those of friends and family members), and so on. Such interests appear to give them reason to reject the principle that selects for the smaller cargo space.[9]

Mutual constraint offers a way of addressing this issue. By applying *ex post* constraints to the reasoning discussed above, mutual constraint can argue that it is not reasonable for the 150 people to reject the principle permitting the car company to spare 1,500,000 people frequent aggravation. It is not reasonable, according to this argument, to forbid an innovation that will benefit many people – on purely self-interested grounds. This application of the MC approach suggests that, when rightfully interpreted, contractualism should not serve as a bludgeon by which a few people can impose a principle, which benefits them at the expense of many people, upon the population at large.

Using mutual constraint, though, it is unclear how to resolve the dispute between the parties. Clearly the small harm, frustration, suffered by the 1,500,000 people is less burdensome than the big harm, serious injury or death, suffered by the 150 people (call these harms H_2 and H_1, respectively). Still, avoiding H_2 seems like a reasonable goal, especially when so many people are affected. Contractualism, moreover, does not prohibit actions out of a concern to avoid all harm – at least on Scanlon's interpretation as defended above. The MC approach seems unable to resolve this conflict. Again, I note that this need not constitute an objection to the view, not if that is the best one can plausibly do. But other strategies should be considered.

Towards that end, the PEA approach appears to offer a strategy for avoiding the conflict entirely. According to this viewpoint, the harms and benefits are combined, along with the numbers of people who will actually be harmed and benefited (*ex post*) by the actions, into probabilistic harm-benefit ratios. Decision makers consider whether all of the people who would experience those *ex post* harms and benefits have reason to consent *ex ante* to the probabilistic harm-benefit ratio (i.e. without knowing whether they would experience the harm or the benefit). In the *Gas-Tank Risks* case, then, PEA contractualism's question is whether 1,500,000 people could reasonably reject an action that subjected them to a 150-in-1,500,000 (0.01%) chance of a big harm (H_1) in exchange for a small benefit (call it B_1) over 5 years.

[9]Alternatively, if consumers were willing to spend more money, they could (theoretically) enjoy both a safe gas tank and a large trunk. Consumers who prefer a larger trunk over a safer gas tank view the probability-adjusted values of each feature as equal to one other; the value of both, though, is not equal (in their eyes) to the money they would have to pay to get both.

https://doi.org/10.1017/S0266267120000310 Published online by Cambridge University Press

Probabilistic, *ex ante* contractualism thus offers a procedure for evaluating ethical questions about risk that is quantitative but not limited to quantitative values (as is EU theory). The procedure is quantitative in the sense that the ethical evaluation turns on whether the people affected by the principle have reason to consent to a particular (numeric) harm-benefit ratio. It is not limited to quantitative values in the sense that the harms and benefits that underlie this ratio are not utilities. In the *Gas-Tank Risks* case, B_1 and H_1 refer to the reasons underwriting people's interests in a larger cargo space: carrying (or not caring to carry) equipment, groceries, and other items that help them to pursue the goals they choose to, from place to place.[10] Although the extent of these benefits and harms can be estimated – e.g. B_1 is known to be smaller as a benefit than H_1 is big as a harm – PEA contractualism does not assign them a numeric utility.[11]

Because the relevant factors are not utilities, PEA contractualism does not offer strict evaluations. That is to say, the evaluations depend to some extent on how evaluators understand the reasons underwriting the harms and benefits. This lack of (numeric) precision in assessing the relative harms and benefits does not undermine the rigour of the PEA contractualism's evaluations, however. Rather, PEA contractualism's evaluation hangs on the permissibility of the harm-benefit ratios, e.g. whether the value of B_1 gives people reason to consent to a 0.01% risk of incurring H_1 over 5 years. It is in this sense that the evaluation is not limited to quantitative values (as is EU theory) but remains largely quantitative: focused on the (extent of the) numeric risk of harm.

Perhaps because of this characteristic of the view – that evaluators must consider the reasons underwriting the harms and benefits associated with a particular harm-benefit ratio under consideration and not just the utilities – PEA contractualism does not assign an absolute (numeric) threshold separating permitted from prohibited harm-benefit ratios. The expectation, though, seems to be that many principles imposing a small chance of serious harm in exchange for a valuable benefit will be permitted. I infer this from Frick's claim that 'For contractualists, the rightness of an action is a function of each individual's personal reasons for rejecting a principle that licenses the action. An act is wrong if and only if there is someone who can complain that we failed to treat her in a way that was justifiable to her' (Frick 2015: 187). As exposure to a small risk of harm does not typically make an action unjustifiable, it appears that PEA contractualism will permit many small risks of harm that secure benefits.

Although the PEA standpoint will often accord with EU-theoretical evaluations of risky actions – in endorsing actions that benefit many – it provides a distinctive rationale for the risky actions it permits. That rationale is: permitted actions are

[10]These reasons are 'generic' in Scanlon's (1998: 203) sense. As such, the PEA approach can avoid some of the epistemic problems that come with determining utility for a specific person.

[11]Indeed, in Frick (2015) concerns about the utilities of harms and benefits do not arise: for the reason that Frick contrasts benefits that are clearly different from one another (such as losing use of a leg as opposed to losing one's life) and making the difference part of the evaluation (Frick 2015: 181–183). The benefits are not expressed quantitatively on PEA contractualism and the difference between differently sized benefits is taken to be intuitive.

https://doi.org/10.1017/S0266267120000310 Published online by Cambridge University Press

underwritten by a principle to which all parties affected by the action have reason to consent (i.e. could not reasonably reject) from an informed, *ex ante* position.

In *Gas-Tank Risks*, the chance of injury or death is quite small; thus, people seem likely to agree to it. In this sense, the PEA approach offers a more straightforward way to evaluate *Gas-Tank Risks* than AH or MC. It permits risks associated with the smaller gas tank on the grounds that people seem likely to accept those risks from the *ex ante* perspective, while fully aware of the probability of harm.

This evaluation might not be accepted, however. In particular, people might counter that the risk of death, though small, is unreasonable in light of the insignificance of the benefit, B_1, for which it is exchanged. The case is taken, moreover, from the Pinto business-ethics case, in which the Ford Motor Company is widely regarded as having acted outrageously in making the tradeoff (safety for storage space) that is considered here.

Responses to these objections are available. First, outrage in the Pinto case typically attaches to the allegation that Ford continued selling cars with fuel-leaking tanks based on (a) a cost-benefit analysis that included a dollar figure for the human lives lost and (b) a calculation that it would be cheaper to lose the lives than to fix the tanks (Dowie 1977). This calculation also involves a trade-off but it is cruder (and, as such, appears more ethically questionable) than the one between safety and storage space. The safety-storage space trade-off does not seem unethical according to PEA contractualism in the sense that people could agree to the harm-benefit ratio from the *ex ante* perspective while fully informed of the probability of harm. Second, most people regularly make trade-offs like the one contemplated here: their personal safety for material gain. One of the things that makes PEA contractualism so powerful is that it can avail itself of this widespread tendency. It allows PEA contractualism to respect humanity while still capturing the benefits associated with many risks.

3. Objection: diverse risk-preferences

In order to test and strengthen PEA contractualism, I challenge it with a third, more difficult objection: concerning people's diverse risk-preferences. Some very risk-averse people might reasonably reject even a 0.01% chance over 5 years of a large harm to secure a small benefit, that is, on the basis of generic reasons (Scanlon 1998: 203) that many people share. Because of this possibility, the PEA approach fails to meet contractualism's demanding standard that every person affected by an action can agree to the principle that underlies it. To investigate whether PEA contractualism's failure to address diverse risk-preferences is significant, I offer accounts of (3.1) the reasonableness of diverse risk-preferences and (3.2) how diverse risk-preferences stymie PEA contractualism.

3.1. Individual risk-preferences

Individual risk-preferences concern the harm-benefit ratios associated with actions that people freely undertake. A risk-loving person might relish her daily motorcycle commute, which imposes an approximately 100 in 100,000 (0.1%) risk of death over 1 year (which is a 0.4% risk of death over 5 years, the time frame in the

https://doi.org/10.1017/S0266267120000310 Published online by Cambridge University Press

gas-tank-risks case[12]) in order to secure benefit, B_2, completing her/his daily commute. A more risk-averse person might prefer to drive, experiencing an approximately 10 in 100,000 (0.01%) risk of death in a year (or 0.05% over 5 years) to gain B_2.[13] A very risk-averse person could employ defensive-driving techniques or, perhaps, telecommute. I assume that very risk-averse people have risk-preferences of 0.001% risk of death in a year (or 0.005% over 5 years) to secure benefit B_2.

A person's individual risk-preference, as I understand it, is the highest harm-benefit ratio of any activity the individual freely undertakes. I define risk preferences in terms of individuals' highest harm-benefit ratios (rather than, e.g. the lowest or average harm-benefit ratios) on the assumption that people provide tacit consent to higher risk levels by freely participating in riskier activities. By freely participating in riskier activities, that is to say, individuals indicate that such risk levels do not exceed their risk preferences. Indeed, their actual risk preferences could be higher than the riskiest activities in which they freely participate.[14]

In order to understand whether extreme risk-aversion is reasonable (i.e. whether it provides reasonable grounds to reject a principle imposing risks with greater frequency than the extremely risk-averse harm-benefit ratio) under PEA contractualism, decision makers must consider the reasons underlying this viewpoint.[15] Imagine, for example, a person who is very sensitive to the challenges that threaten her basic existence. This sensitivity might result from a natural predisposition to nervousness or from training in a discipline such as accounting, law or actuarial science. Such a person could reasonably respond to her sensitivity by trying to minimize risk wherever she can: keeping money in a safe rather than in the bank, staying close to home rather than travelling about, making the products she needs rather than purchasing them in stores. Considering this standpoint empathetically, the extreme risk-aversion of such people appears reasonable. Thus, evaluators may conclude that people whose safety is risked in the *Gas-Tank Risks* case could reasonably reject the 0.01% risk of harm (per year) to secure benefit B_1.

3.2. *Probabilistic,* ex ante *contractualism and diverse risk-preferences*

Probabilistic, *ex ante* contractualism has no obvious resources to accommodate such diverse risk-preferences. In the *Gas-Tank Risks* case, for example, the PEA

[12]If r = risk per year, r_5 (risk over 5 years) = $r + (1-r)\, r + (1-r)^2\, r + (1-r)^3\, r + (1-r)^4\, r$.

[13]Based on National Highway Safety Administration (2007: 3). According to this data review, the fatality rate for motorcyclists was 72.34 per 100,000 registered vehicles in 2007; the fatality rate for passenger cars was 13.10 per 100,000 registered vehicles.

[14]In crafting investment portfolios, people generally treat risk differently from my understanding: balancing riskier with safer investments such that the overall risk profile of their profile (which reasonably expresses their investment risk-preferences) is lower than the riskiest investments included therein. I assume that people do not generally hedge the risky activities in which they participate in this manner, however.

[15]For a discussion of the problems with using individual preferences to motivate moral claims, see Scanlon (1975: 655–669).

https://doi.org/10.1017/S0266267120000310 Published online by Cambridge University Press

evaluation requires people to find reasonable a 0.01% risk of harm that secures a benefit. Absent the assumption that people actually do or must have reason to consent to this risk, PEA contractualism seems to be in the same position as AH contractualism: bound to reject the risk as unethical if even one person lacks reason to consent to it. Because one of the PEA approach's advantages over AH contractualism appears to be its ability to endorse more (socially beneficial) risks, however, this result would weaken the view. In light of this objection, the only way PEA contractualism could endorse socially beneficial risks like those associated with some gas tanks is by compelling people to accept whatever risk-preference the decision maker finds reasonable. Because contractualism aims to respect what risks people accept from their own perspectives – rather than compelling them to accept risks that others find reasonable – this aspect of PEA contractualism invites contractualists to seek a better characterization of when risks are reasonable.

4. Including individual risk-preferences: REU theory

Though this problem challenges PEA contractualism, diverse risk-preferences do not defeat the view, in my estimation. Indeed, the problem of diverse risk-preferences appears to relate to one of contractualism's morally laudable aspects: its ability to include and value a variety of standpoints, in particular (a) preserving respect for the humanity of individuals even as it (b) considers how actions benefit populations. Contractualism takes seriously the concerns of people on the margins of societies, such as those with very risk-averse risk-preferences. At the same time, though, Scanlon formulated his viewpoint with the explicit requirement that it must permit (some) socially beneficial risks. Contractualism should not, in this sense, require decisions to be so risk averse that they become unable to benefit societies.

Around the time James (2012) and Frick (2015) introduced (what I call) PEA contractualism, Buchak (2013) proposed a revision to EU theory to accommodate decision makers' personal attitudes towards risk. She calls her revision *risk-weighted expected utility* (REU) theory (Buchak 2013). Although REU theory does not involve either contractualism or risk preferences as I have defined them (as the highest harm-benefit ratio of the activities a person willingly undertakes), its strategy in amending a decision-making theory to include individual attitudes vis-à-vis risk is well suited to guide my efforts in making contractualism more responsive to diverse risk-preferences.

Buchak (2013) seeks to include agents' personal attitudes towards risk (whether they are risk-avoidant or risk-inclined) in rational (i.e. EU-style) decision-making. To do so, she formulates them as people's *risk functions* (Buchak 2013: 49–51). Risk-avoidant risk functions give more weight to worse outcomes (risks) associated with a gamble; risk-inclined risk functions give less weight to these risks, thereby focusing more on the better outcomes. A risk-avoidant agent's REU value for a risky act, then, is likely to be lower than her EU value for the same act; a risk-inclined agent's REU value for a risky act is likely to be higher than her EU value for that act. Using this theory of rational decision-making (sensibly) makes it more

https://doi.org/10.1017/S0266267120000310 Published online by Cambridge University Press

likely both that a risk-inclined agent will decide on a risky act and that a risk-avoidant agent will decide against a risky act.

Buchak's (2013) risk-weights can be helpfully contrasted with the harm-benefit ratios by which PEA contractualism conceived of various levels of risk. Harm-benefit ratios are particular numbers, e.g. 1/1,000 or 1/10,000. Risk-weights, by contrast, magnify or diminish the impact of risks (or the worse outcomes) on decision makers' options. Risk-weights thus enjoy an advantage over harm-benefit ratios with respect to decision making about risk to the extent that it seems intuitively more plausible to conceive of individual risk-preferences as weights on decision-making rather than as particular numbers. An advantage of harm-benefit ratios, on the flip side, is that they are more explicit and, as such, more readily criticized (and thus more readily made more precise).

Buchak's (2013) innovation, in this sense, engages with the concern I noted about PEA contractualism: by allowing decisions to be modified by individuals' preferences vis-à-vis risk. It does not resolve the concern, though, to the extent that REU theory focuses on individual decision-making involving risk (Buchak 2013: 1–4) whereas the PEA approach addresses risks affecting populations. Though attractive as a way of conceiving of individual risk-preferences, moreover, risk-weights would be difficult to use in modifying PEA contractualism. As PEA evaluates risk levels as particular numbers, in assessing whether those risk levels correspond to individual preferences it seems plausible that decision makers should employ numeric individual risk-preferences.

Consider a version of *Gas-Tank Risks*, for example. A potential consumer could use REU theory to decide whether to buy an inexpensive car with a large trunk and risky gas tank or an inexpensive car with a small trunk and a safe gas tank, based on the following assumptions concerning probabilities and utilities. If the consumer buys the risky car, she will experience one of three outcomes: no accident (0.5 probability), accident in which the car does not lose fuel (0.4 probability), and accident in which the car loses fuel (0.1 probability). The no-accident and the no-fuel-loss-accident outcomes have a utility of 5. The fuel-loss-accident outcome has a utility of 0. If the consumer buys the safe car, by contrast, s/he will either have no accident (0.5 probability) or an accident with no fuel loss (0.5 probability). In either case, the small-trunk car offers the consumer a utility of 4. As such, EU theory recommends:

$$\mathrm{EU}(\text{risky car}) = 0.5(5) + 0.4(5) + 0.1(0) = 4.5$$

$$\mathrm{EU}(\text{safe car}) = 0.5(4) + 0.5(4) = 4$$

To allow the consumers' risk-preferences to play a role, REU theory reconceptualizes the utility measures in terms of *incremental utility*, or how much utility increases among the various options. With respect to the risky car, there is a probability of 1 that the consumer will experience at least a utility of 0 and a probability of 0.9 that she will experience a utility of 5 more than that. REU theory will decrease the weights of the probabilities for a risk-avoidant consumer (for example, by squaring them) and increase them (for example, by taking their square root) for a risk-inclined consumer:

https://doi.org/10.1017/S0266267120000310 Published online by Cambridge University Press

Risk-avoidant consumer:

$$\text{REU(risky car)} = (1-0.9^2)(0) + (0.9^2)(5) = 4.05$$
$$\text{REU(safe car)} = (1-0.5^2)(4) + (0.5^2)(4) = 4$$

Risk-inclined consumer:

$$\text{REU(risky car)} = (1-0.9^{0.5})(0) + (0.9^{0.5})(5) = 4.74$$
$$\text{REU(safe car)} = (1-0.5^{0.5})(4) + (0.5^{0.5})(4) = 4$$

With these risk-weights, REU theory recommends the risky car for both the risk-avoidant- and the risk-inclined consumer; a more extreme risk-weight, though, would cause REU theory to recommend the safe car for the risk-avoidant consumer (thus diverging in its recommendations from EU theory).[16] It is natural to assume here that the risk-inclined (motorcycle riding) consumer corresponds to the square-root calculation and the risk-avoidant (defensive driving) consumer corresponds to the squaring calculation. Because of the differences in the ways in which REU theory and PEA contractualism assign individual risk-preferences, however, I am unable to assess whether squaring the probabilities of each outcome corresponds well to a 0.1 risk of harm in exchange for benefit B_2.[17] In this sense, REU theory provides guidance about what cars particular agents should buy but not about the overall permissibility of exposing a population of people (who have various risk-preferences) to the risks associated with the risky car.

5. Including individual risk-preferences: the *RISPREAD* approach

Although I seek a way of addressing the latter question rather than the former, Buchak's theory is nonetheless instructive as regards the problem in contractualism that this paper has identified. In principle, REU theory could be extended to address questions about populations;[18] although promising, I do not attempt such an extension in this paper for the reason that it would not readily develop PEA contractualism, as I aim to do herein. Rather, I follow a Buchakian model of modifying a canonical decision-making theory (in her case, EU theory;

[16]This discussion of *Gas-Tank Risks* is modelled on the example in Buchak (2017*a*).

[17]There might be a way of translating individual harm-benefit ratios into Buchakian risk functions; I do not pursue such translations here, though.

[18]In her recent work on REU theory, Buchak (2013) has developed the view in ways that show its potential for broader application. She has used REU theory to support a Rawlsian claim of distributive justice (that the interests of the relatively worse off should bear greater importance in decision-making than the interests of the relatively better off) in Buchak (2017*b*). She has used insights from REU theory to argue that medical researchers may enrol willing participants in risky medical studies so long as the risks correspond to the patients' risk-preferences in Buchak (2016). The present paper contributes to the broader application of REU theory to the extent that it offers a strategy for adding the individual risk-preferences, whose importance Buchak (2013) has highlighted, to ethical decision-making about populations.

https://doi.org/10.1017/S0266267120000310 Published online by Cambridge University Press

in mine, contractualism) to make it more responsive to individual attitudes towards risk. Whereas Buchak (2013) conceives of individual attitudes towards risk in terms of risk functions that increase or decrease the weight of risk in EU decision-making, my focus on PEA contractualism encourages me to conceive of individual attitudes toward risks in terms of harm-benefit ratios. This is because PEA contractualism's (inspired, on my view) advance over other forms of contractualism is its ability to conceive of population-wide risks in terms that address individuals: as ratios that weigh the total harm to the population over the total benefit to the population. Individuals can reflect on these ratios as risks to themselves, not others, allowing individuals to decide for themselves whether they consent to the risks (i.e. from an *ex ante* perspective while fully informed of the risks of harm). And whereas PEA contractualism considers whether the number is intuitively compelling, I seek to provide additional structure for contractualist evaluations of risk: in terms of whether the risk level is reasonable in light of individual risk-preferences in the population exposed to the risk.

I thus offer the risk-preferences-adjusted (RISPREAD) approach, which holds that the members of a population of people exposed to a risk have reason to consent to that risk whenever the risk is less than, or equal to, the average of their individual risk-preferences. The approach thus includes each member's individual risk-preference in the harm-benefit ratio it endorses as reasonable for a population. To capture and apply this ratio, RISPREAD offers three steps. First, decision makers establish (5.1) the numeric risk-preferences of the individual members of the population exposed to risk. Next, they calculate (5.2) the portions of the population holding each risk-preference (for a variable number of population groups, such as three, depending on what is most straightforward for a particular decision). Then, decision makers multiply (5.3) each numeric risk-preference by the portion of the population holding the risk-preference. These products are summed and the result is the population's portion-adjusted (i.e. average) risk-preference. According to RISPREAD, only risks that have lower (more risk-averse) harm-benefit ratios than this are permitted. I describe the procedures by which decision makers can carry out these stages in turn.

5.1. Calculating individual risk preferences

Decision makers begin by calculating the individual risk-preferences of the members of the population whom they consider exposing to a risk. Following PEA contractualism, I conceive of risk-preferences as harm-benefit ratios;[19]

[19]It is worth pointing out that there are other ways in which contractualism could formulate individual risk-preferences. In addition to the risk-weights of Buchak's (2013) REU theory, as discussed above, contractualism could conceive of risk-preferences in terms of average *benefit-harm* ratios (among other strategies). As noted by a helpful reviewer for *Economics and Philosophy*, if A has a harm-benefit ratio of 0.1 and B has a harm-benefit ratio of 0.01, the average harm-benefit ratio is 0.055. Given these risk-preferences, A has a benefit-harm ratio of 10 and B has a benefit-harm ratio of 100, producing an average benefit-harm ratio of 55, which is equivalent to a harm-benefit ratio of 0.018. In the RISPREAD approach, I follow PEA contractualism and therefore use harm-benefit ratios. In the broader debate about individual risk-preferences and ethics, though, it is worth bearing in mind that scholars face a variety of options and must justify our ultimate choices carefully.

https://doi.org/10.1017/S0266267120000310 Published online by Cambridge University Press

as defended above, an individual's risk-preference is the harm-benefit ratio of the riskiest activity in which the individual willingly participates (Oberdiek 2004). When risky activity k is the riskiest activity in which a person participates, that person's risk-preference is formulated as:

$$\frac{\text{Number of people (expected to be) harmed by risky activity } k}{\text{Number of people who benefit from risky activity } k}$$

Individual risk-preferences are thus based in a quantitative understanding of risk (as are EU and REU theory) but are not limited to that quantitative understanding. People's risk-preferences can also be used to express how much risk they have reason to consent to, such as in the sense discussed above with respect to PEA contractualism. Because RISPREAD calculates risk-preferences as the numbers of people expected to be harmed by a risky activity divided by the numbers of people expected to benefit from a risky activity – rather than, e.g. in terms of the utilities of the harms and benefits associated with the risk – it is important to use risk-preferences assessed from activities that involve similar benefits and harms.

This requirement may appear to be difficult to fulfil. Consider, as illustration, the example of motorcyclists and car drivers offered above. I claimed that, if motorcycle riding is an individual's riskiest activity, then that person's risk-preference is a 0.1% risk of death per year in order to secure benefit B_2. Motorcyclists probably experience additional benefits (call them B_3, the thrill of the wind racing through their hair, and so on) which makes them willing to take on the additional risk associated with their preferred means of transport. Stripped of B_3, though, the risk-preference that RISPREAD assigns them is too great (i.e. too risk-loving).

In response, I grant that this is a serious worry about the RISPREAD approach. The worry can be mitigated, though, by ensuring that the harms and benefits (as utilities) associated with RISPREAD's harm-benefit ratios meet a certain standard. In order to use people's (individual-risk-preference constituting) harm-benefit ratios in other situations, then, decision makers must be able to establish that the harms and the benefits are comparable, that is, that the harm associated with the greatest harm-benefit ratio of any activity in which a person willingly participates, H_R, is greater than or equal to the harm associated with the risk under examination, H_E, and that the benefit associated with this harm-benefit ratio, B_R, is less than or equal to the benefit associated with the risk under examination, B_E: $H_E \leq H_R$ & $B_E \geq B_R$.[20]

According to RISPREAD, then, motorcyclists have reason to consent to other risks that impose a 0.1% risk of a serious harm (such as devastating bodily injury or death) in order to secure a benefit greater-than-or-equal-to $B_2 + B_3$ (which confers significant convenience, at a minimum, being able to travel to the places they want to go). Car drivers, by contrast, have risk-preferences of 0.01% (risk of death per year in exchange for B_2). Most people can be assumed

[20]Ensuring that harms and benefits meet this standard allows decision makers to evaluate risks when they can establish an ordering of harm-benefit ratios, even if they cannot establish the precise values of the harm-benefit ratios, as discussed further below.

https://doi.org/10.1017/S0266267120000310 Published online by Cambridge University Press

to share the risk-preference of car drivers to the extent that, even if they do not drive themselves, they purchase goods that have been transported via automobile and rely on the services of car-driving police officers and ambulances (Oberdiek 2004: 203). As broached above, I assume that very risk-averse people (e.g. who neither drive nor choose to avail themselves of services associated with automobiles) have risk-preferences of 0.001% (risk of death per year in order to receive benefit B_2).

5.2. Calculating population-portion-adjusted (average) risk-preferences

The next step is determining which portions of the risk-exposed population hold which risk preferences. For purposes of simplicity, I use the three risk-preferences groups mentioned above: 0.1%, 0.01% and 0.001% (risk of harm in exchange for a benefit, where the benefit is assumed to be constant). The quantitative risk-preference associated with each group is approximate. Although more precise risk-preferences would allow RISPREAD's evaluations to calculate more accurately when people may reasonably reject a risk to which they are exposed, there are advantages to approximating risk preferences. In particular, numeric risk-preferences under the RISPREAD approach apply to all risks, even when they are qualitatively quite different from the activity that gives rise to a person's individual risk-preference. As a largely non-quantitative decision-making procedure – which focuses on reasons rather than utilities – contractualism has more flexibility to consider these qualitative differences than strictly quantitative decision-making procedures like EU theory. Relying on approximate (rather than exact) risk-preferences thus supports contractualism's evaluative process of considering the reasons underlying people's individual risk-preferences.

In this stage of its evaluative process, RISPREAD relies on empirical information about the frequencies of various risk-preferences, such as those available in the US Census. Take the state of California. According to the US Census, California's population was 38.99 million in 2015 (US Census Bureau 2015). There were 828,883 motorcycle registrations in California (Wagner 2020*b*) and 25,532,920 licensed Californian drivers that year (Wagner 2020*a*). Based on these numbers, I infer that 2% of Californians have risk-preferences of 0.1% (risk of death over one year in exchange for benefit B_2). By contrast, 65% have risk-preferences of 0.01% (risk of death over one year in exchange for B_2).

I attribute (more speculatively) a risk-preference of 0.001% (risk of death over one year in exchange for B_2) to the remaining 33% of the California population. This number might seem surprisingly high. Part of the surprise can be attributed to the fact that the California population included approximately 10 million children (Lucile Packard Foundation for Children's Health 2015) (25% of the population) and 7.3 million seniors (18% of the population) in 2015, many of whom are not drivers. As many of the seniors do not drive and may not be competent to consent to being driven, it seems reasonable to attribute the more risk-averse preference to that population groups.[21] Children's risk-preferences can be assessed using similar standards – they do not drive and may not

[21]A metric other than driving might produce different risk-preferences for the population and this is an issue to which decision makers should pay heed.

https://doi.org/10.1017/S0266267120000310 Published online by Cambridge University Press

consent to being driven – but may be more controversial for the reason that they will grow up into the driving world with all of its benefits. Because of this characteristic, their risk-preferences might be construed as mirroring those of adults. I interpret child non-drivers as belonging to the more risk-averse group for the reason that they constitute non-drivers who might not be competent to consent to being driven at the time that the risk calculation is made – regardless of what the future might hold for their risk-preferences.[22]

5.3. Summing portion-adjusted risk-preferences over a population

Based on these numbers, RISPREAD calculates California's population-portion-adjusted (average) risk-preference: $(0.1 \times 0.02) + (0.01 \times 0.65) + (0.001 \times 0.33) = 0.00883\%$ risk of death over one year in order to receive benefit B_2 or 0.0434% risk of death over 5 years to receive B_2.[23] As noted above, the RISPREAD approach permits risk exposures when they are lower (less risky) than the population's greatest common harm-benefit ratio (over the same amount of time) and prohibits them when they are higher (riskier) than that. Because California's 2015 population-portion-adjusted (average) risk-preference (based on benefit B_2) is thus lower than the *Gas-Tanks Risks* preference of 0.05% (risk of harm over 5 years in exchange for B_1), RISPREAD prohibits the car company in *Gas-Tanks Risk* from adopting the riskier tank in California in 2015, assuming that two conditions hold ($H_E \leq H_R$ & $B_E \geq B_R$) as defined above. In this case, H_E and H_R are the same: death or serious injury, thus satisfying the first condition. The benefit under examination (B_1), though, is not plausibly greater than the benefit associated with the harm-benefit ratio (B_2), in the sense that B_2, being able to travel from place to place, permits people to carry out basic aspects of their lives but B_1, larger trunk space, seems like a mere convenience. This difference does not entail that the *Gas-Tanks Risks* are not prohibited, after all; rather, it demonstrates that harm-benefit ratios based on B_2, can never justify risks for the sake of B_1. Different data, then, would need to be used to establish H_R and B_R for the purpose of evaluating risks that produce benefits similar to those in *Gas-Tanks Risks*.

The *Gas-Tanks Risks* case does not explicitly engage with another concern about RISPREAD, which I raise in order to clarify RISPREAD's different standard for permitted risks from PEA contractualism. In the *Gas-Tanks Risks* case, RISPREAD prohibited the decision maker (the car company) from choosing the riskier (larger-trunk-space) alternative. This prohibition might frustrate the members of the population who prefer the larger-trunk-space alternative; such frustration would not, however, give risk-loving members of the population reason to reject the car company's choice according to PEA contractualism. Members of a population with more risk-averse risk-preferences than the population's average risk-preference would, however, have reasonable ground to reject RISPREAD-endorsed risks that are higher than their individual risk-preferences according to the former interpretation of contractualism.

[22]Cf. Buchak (2017*b*, 2019).

[23]Using the formula cited in fn. 12.

https://doi.org/10.1017/S0266267120000310 Published online by Cambridge University Press

In response, I note that it was PEA contractualism's inflexibility on this issue that led me to propose RISPREAD contractualism: to offer an alternative characterization of when risks are reasonable. Although RISPREAD is vulnerable to this objection – some of the people affected by its population-portion-adjusted risk-preferences could have reason to reject the risk-preferences according to the standard set forth above – such vulnerability does not entail that the approach fails to satisfy contractualist standards. In particular, RISPREAD meets Scanlonian contractualism's requirements that decision makers reflect on risks from standpoints other than their own and consider the risk's burdens and benefits from these various standpoints (Scanlon 1998: 153). The average standpoint that RISPREAD targets is well suited as a reflection place to the extent that this standpoint incorporates all standpoints affected by the risk in the proportions in which they are represented in the population. As such, RISPREAD allows all standpoints to influence the standard (i.e. population risk-preference) used to determine which risks are reasonable. Risks are reasonable, according to RISPREAD, when they reflect all the standpoints affected by the risk in the proportions in which those standpoints are represented in the population.

6. Discussion

The RISPREAD approach seems promising in its ability to calculate a numeric risk-threshold for a population that responds to the individual risk-preferences of members of that population. In this section, I consider two worries about the approach: the approach seems unintuitive in populations with widely divergent risk-preferences and, relatedly, the approach seems well suited to evaluate only a narrow range of risks. I discuss these worries in turn.

First, consider a population with starkly different risk-preferences. In this population, 95% of people are extremely risk-averse, tolerating only 0.001% risk of death in a year (or 0.005% over 5 years) in order to secure benefit B_1 (or one of similar or greater magnitude). The remaining 5% of people, though, are so risk-loving as to be nearly suicidal; they prefer any risk at all, up to 1.0%, to secure B_1 (or a benefit of similar or greater magnitude). Their population-portion-adjusted risk-preference is thus 0.05% (risk of death to secure a benefit like B_1) over 1 year or 0.055% over 5 years. Despite the overwhelmingly risk-averse preferences of this population, then, the RISPREAD approach calculates that it permits both car driving and the risky gas-tanks discussed above.

In response, I concede that RISPREAD appears to be useful in establishing a population's risk-preference only when individual risk-preferences are fairly close knit. I am not certain how serious this problem is, as the polarized preferences discussed in the foregoing example seem unrealistic. Importantly, though, RISPREAD does provide a strategy for resolving extreme disputes, such as those in the population examined above; its strategy, moreover, takes all perspectives into account. In the above example, it is unfortunate for the risk-averse 95% that their preferences are not decisive; it is a virtue of the RISPREAD approach,

https://doi.org/10.1017/S0266267120000310 Published online by Cambridge University Press

though, that it can recognize minority preferences when they are very strong. Within the contractualist framework of requiring respect for the humanity (understood as the capacity for consent) of everyone affected by a decision, it seems plausible that risk-averse preferences should not automatically overwhelm risk-loving preferences (or vice versa).

In this sense, the RISPREAD approach represents a strategy for competition between different standpoints, not dissimilar from the MC approach. The RISPREAD approach thus incorporates aspects of both of the very promising approaches to risk that this paper has examined. Like the PEA approach, it evaluates actions that impose risks across populations from the perspectives of individuals within those populations. It does more than the PEA approach to include the perspectives of all individuals in its evaluations. As regards the MC approach, it does more to offer a compromise strategy between different perspectives.

Second, the RISPREAD approach's focus on numeric harm-benefit ratios appears to exclude other factors that are potentially relevant to whether a risky action is permitted. Consider the following case:

> *NO_x Emissions Risks*: A car company is deciding whether to add emissions controls to its cars' diesel engines in order to reduce the amount of nitrous oxide (NO_x) these cars emit. Adding the controls would reduce driving performance, decreasing utility for the cars' 500,000 young, middle-class drivers. Failing to reduce NO_x emissions, though, would increase fine particulate matter in the air, contributing to air pollution-related premature mortality for 50 impoverished or vulnerable people over 5 years.[24]

Whereas the *Gas-Tank Risks* case is an example of a *straightforward* problem for contractualism, *NO_x Emissions Risks* seems more complicated.[25] *Gas-Tank Risks* is straightforward in the sense that car consumers are subject to both the risks and the benefits of the fuel-losing tank: all customers can be both (a) frustrated by a small trunk or (b) seriously injured or killed as a result of leaking gas. Thus, the interests of these two groups are not directly pitted against one another. In *NO_x Emissions Risks*, by contrast, the interests of 500,000 young, middle-class drivers appear to conflict with the interests of 50 impoverished, vulnerable people.

Framed in this way, in which a car company configures its diesel engines so as to benefit one population by harming another, the NO_x emissions risks

[24]These figures are based on the 59 people who are expected to experience air pollution-related premature mortality as a result of fine particulate matter released by 482,000 Volkswagen cars with diesel engines between 2009–15 (Chossière *et al.* 2017). The claim that those who die prematurely as a result of the increased NO_x emissions are impoverished or vulnerable is based on a health analyst's argument that most emissions occur (a) as drivers enter highways and (b) in stop-and-go traffic: 'less desirable areas are where poorer people live' (Borenstein 2015).

[25]The difference between 'straightforward' and 'more troubling' or 'trickier' problems of risk is insightfully discussed by Lenman (2008: 101–102). A straightforward case imposes a risk on an entire population in order to benefit that population. In a 'super'-straightforward case, as Lenman (2008: 102) describes it, the benefit is a lower harm-benefit ratio than that which an alternative action would impose.

https://doi.org/10.1017/S0266267120000310 Published online by Cambridge University Press

appear prohibited.[26] They appear prohibited because decision makers cannot justify the risk (premature death) associated with the decision not to add emissions controls to diesel engines to the people who would be exposed to this risk. The risk seems unjustified, that is, for the reason that the people who would bear the risk do not benefit from it: as they do not own the cars that have the enhanced driving performance. In this frame, the RISPREAD approach (along with the PEA approach on which it is based) appears wrongheaded. Both approaches require the car company to evaluate the risk from the standpoint of each person affected by it. They instruct the company to consider whether it may impose a 50 in 500,050 (0.01%) chance of premature death on each individual over 5 years in exchange for increased driving pleasure.

In response, I grant that RISPREAD is, like the PEA approach, unable to resolve questions of this structure: in which the interests of one group are pitted against those of another. I reject, though, the idea that questions with this structure are questions about risks, for which people's individual risk-preferences are relevant. Questions about risks concern populations, not specific individuals or subdivisions of the populations. All questions about risk are thus straightforward in the sense defined above: concerning whether the population may be exposed to a risk, evaluated in terms of harm-benefit ratios relative to individuals. Questions such as that framed in the *NO_x Emissions Risks* case, by contrast, concern whether it is permitted to impose a harm on one person (or group) so as to benefit a different person (or group). As noted in the Introduction, contractualism prohibits actions of this form, generally.

7. Conclusion

In this paper, I discussed the development, and the promise, of the contractualist approach to risk: culminating in the recent promulgation of the mutual constraint (MC) and probabilistic *ex ante* (PEA) approaches. The MC approach evaluates risks from both *ex ante* and *ex post* perspectives; the PEA approach permits risks when everyone who would be affected by the risks could agree to them (a) from the *ex ante* perspective and (b) while fully informed of the probabilities of harm that the risk would impose on the population. I argued that these approaches escape problems associated with other recent interpretations of the contractualist approach to risk and meet Scanlon's core standards. First, the approaches consider seriously the harms associated with risky decisions. Second, they do not prohibit all risky actions even when serious harms could occur as side effects.

[26]The *NO_x Emissions Risks* case is modelled on the 2015 Volkswagen emissions fiasco. The Volkswagen case is often regarded as a prohibited business decision – but for a different reason. That is, what was clearly unethical in the Volkswagen Defeat Device scandals was the allegation that Volkswagen schemed deliberately to trick the Environmental Protection Agency's emissions tests (not the risk to which people were exposed). It was the intention to deceive that outraged regulators and the public. This back story does not play a role in my analysis of the case, which examines business decision-making rather than what level of emissions the government should permit or whether companies must obey the law.

https://doi.org/10.1017/S0266267120000310 Published online by Cambridge University Press

I then challenged the PEA approach with an objection. I noted that the PEA approach appeared to assume that all people affected by the risk would assess the risk's probability of harm in a uniform manner. If the risk presented a low probability of harm, PEA appeared to assume that all of the people affected by the risk would agree to it – that is, that none of them would have reason to reject it. I argued that people have diverse risk-preferences and that even extreme risk-aversion can be reasonable on an individual basis (if socially suboptimal). Based on people's diverse risk-preferences, I inferred that the PEA approach to risk did not do as complete a job of fulfilling contractualist principles as it initially seemed to. Because of diverse risk-preferences, some people in a community exposed to a risk could reasonably reject even risks that present very low probabilities of harm.

I argued that contractualists are bound to permit socially beneficial risks and proposed that they do so by offering a more precise characterization of when risks are reasonable, as in the risk-preferences-adjusted, probabilistic-*ex ante* (RISPREAD) approach. Under this contractualist decision-procedure, decision makers (a) establish the portions of a population that share various risk-preferences, (b) multiply those portions by the numeric risk-preferences and (c) sum the portion-adjusted risk-preferences over the population. Only those harm-benefit ratios that are below the population's portion-adjusted preference are considered reasonable according to this contractualist approach.

I considered two objections to RISPREAD: that the approach works well only in populations whose risk-preferences are close-knit and only in 'straightforward' cases of risks in which the population exposed to harm is the same as the population that benefits from a risk. Especially in populations whose risk-preferences are polarized (e.g. extremely risk-averse and extremely risk-loving portions with little middle ground), the harm-benefit ratio that RISPREAD recommends will not satisfy all members' preferences. In fact, the approach can recommend decisions that are counter to the risk-preferences of the majority of the population.

I responded that the flexibility suggested by this objection is actually a strength of RISPREAD, which helps to paint a fuller picture of it as an interpretation of contractualism. RISPREAD responds to the preference sets of actual people, meaning that the evaluations it offers are truly specific to the situations in which it offers them. Its attention to diversity and disagreement is a further strength of contractualism. In societies consisting of people with varied risk-preferences, RISPREAD does not insist on conformity in order to issue recommendations about what is the right thing to do. Rather, it offers principled procedures by which people can work out for themselves what is right for the population as a whole.

Acknowledgements. I am grateful to many people and institutions for comments and guidance as I worked on this article. I had the privilege of presenting earlier versions of these ideas at the 2016–17 Normative Business Ethics Workshop Series of the Carol and Lawrence Zicklin Center for Business Ethics Research at the Wharton School, University of Pennsylvania and the 2019 Society for Business Ethics Annual Meeting in Boston, Massachusetts. I would like to thank Brian Berkey, Ellen Susanna Cahn, Matthew Caulfield, David Dick, Robert Hughes, Julian Jonker, Eric Orts, Gaston de los Reyes, Amy Sepinwall, Michael Simonetti, Alan Strudler, Danielle Warren, Robert Wiener, and others for

constructive critical comments. Finally, I wish to thank Matthew Adler and two anonymous reviewers for extremely careful, incisive and constructive feedback on earlier versions of this article.

References

Birsch D. 1994. Introduction: The Pinto controversy. In *The Ford Pinto Case*, ed. D. Birsch and J.H. Fielder, 5–10. Albany, NY: State University of New York Press.

Borenstein S. 2015. AP analysis: Dozens of deaths likely from VW pollution dodge. *Associated Press*. http://bigstory.ap.org/article/a6925f0af82e44aaa1a1ed4b55d030f6/ap-analysis-dozens-deaths-likely-vw-pollution-dodge.

Buchak L. 2013. *Risk and Rationality*. New York, NY: Oxford University Press.

Buchak L. 2016. Why high-risk, non-expected-utility-maximizing gambles can be rational and beneficial: the case of HIV cure studies. *Journal of Medical Ethics* **43**, 90–95.

Buchak L. 2017a. Precis of *Risk and Rationality*. *Philosophical Studies* **174**, 2363–2368.

Buchak L. 2017b. Taking risks behind the veil of ignorance. *Ethics* **127**, 610–644.

Buchak L. 2019. Weighing the risks of climate change. *Monist* **102**, 66–83.

Chossière G.P., R. Malina, A. Ashok, I.C. Dedoussi, S.D. Eastham, R.L. Speth and S.R.H. Barrett 2017. Public health impacts of excess NOx emissions from Volkswagen diesel passenger vehicles in Germany. *Environmental Research Letters* **12**, 1–14.

Dowie M. 1977. Pinto madness. *Mother Jones*. https://www.motherjones.com/politics/1977/09/pinto-madness/.

Frick J. 2015. Contractualism and social risk. *Philosophy & Public Affairs* **43**, 175–223.

Fried B.H. 2012a. Can contractualism save us from aggregation? *Journal of Ethics* **16**, 39–66.

Fried B.H. 2012b. The limits of a nonconsequentialist approach to torts. *Legal Theory* **18**, 231–262.

Hansson S.O. 1993. The false promises of risk analysis. *Ratio* **6**, 16–26.

Hansson S.O. 2003. Ethical criteria for risk acceptance. *Erkenntnis* **59**, 291–309.

Hayenhjelm M. and J. Wolff 2011. The moral problem of risk impositions: a survey of the literature. *European Journal of Philosophy* **20**, e26–e51.

Hill T.E. Jr 2005. Assessing moral rules: Utilitarian and Kantian perspectives. *Philosophical Issues* **15**, 167–170.

James A. 2012. Contractualism's (not so) slippery slope. *Legal Theory* **18**, 263–292.

Kumar R. 1999. Defending the moral moderate: contractualism and common sense. *Philosophy & Public Affairs* **28**, 275–309.

Kumar R. 2015. Risking and wronging. *Philosophy & Public Affairs* **43**, 27–49.

Lenman J. 2008. Contractualism and risk imposition. *Politics, Philosophy & Economics* **7**, 99–122.

Lucile Packard Foundation for Children's Health 2015. Demographics of California's children. *Kidsdata*. http://www.kidsdata.org/export/pdf?cat=7.

National Highway Safety Administration 2007. Traffic safety facts: 2007 data. https://crashstats.nhtsa.dot.gov/Api/Public/ViewPublication/810993.

Nozick R. 1974. *Anarchy, State, and Utopia*. New York, NY: Basic Books.

Nussbaum M. and A. Sen 1995. *The Quality of Life*. New York, NY: Oxford University Press.

Oberdiek J. 2004. The ethics in risk regulation: towards a contractualist reorientation. *Rutgers Law Journal* **36**, 199–204.

Reibetanz S. 1998. Contractualism and aggregation. *Ethics* **108**, 296–311.

Savage L.J. 1954. *The Foundations of Statistics*. New York, NY: John Wiley and Sons.

Scanlon T.M. 1975. Preference and urgency. *Journal of Philosophy* **72**, 655–669.

Scanlon T.M. 1998. *What We Owe to Each Other*. Cambridge, MA: The Belknap Press of Harvard University Press.

Thomson J.J. 1986. Imposing risks. In *Rights, Restitution, and Risk: Essays in Moral Theory*, ed. W. Parent, 173–191. Cambridge, MA: Harvard University Press.

US Census Bureau 2015. QuickFacts: California. URL: https://www.census.gov/quickfacts/CA.

Wagner, I. 2020a. Licensed drivers in the U.S. – total number by state 2018. *Statista*. https://www.statista.com/statistics/198029/total-number-of-us-licensed-drivers-by-state.

Wagner, I. 2020b. U.S. motorcycle registrations by state 2018. *Statista*. https://www.statista.com/statistics/191002/number-of-registered-motorcycles-in-the-us-by-state/.

Tobey K. Scharding is an Assistant Professor at Rutgers Business School, where she has taught business ethics since 2017. She earned her PhD in philosophy at Stanford University. She specializes in ethical decision-making and ethics of risk. Her book, *This is Business Ethics*, was published by Wiley-Blackwell in 2018. A collection of articles co-edited with Joanne Ciulla, *Ethical Business Leadership in Troubling Times*, was published by Edward Elgar in 2019. Her articles have appeared in *Business Ethics Quarterly*, *Journal of Business Ethics*, *Business & Society Review* and *Rutgers Business Review*, among other journals. URL: https://business.rutgers.edu/faculty/tobey-karen-scharding

Cite this article: Scharding TK (2021). Contractualism and risk preferences. *Economics & Philosophy* **37**, 260–283. https://doi.org/10.1017/S0266267120000310