Studies in Philosophy, Politics and Economics

Vol 1  2019
Introduction to *Studies in Philosophy, Politics and Economics*

Alex Voorhoeve*

* Philosophy, Logic and Scientific Method, London School of Economics, Houghton Street, London, WC2A 2AE, UK. Email: a.e.voorhoeve@lse.ac.uk; web: [http://personal.lse.ac.uk/voorhoev](http://personal.lse.ac.uk/voorhoev)

“It may well be that the chemist or physiologist is right when he decides that he will become a better chemist or physiologist if he concentrates on his subject at the expense of his general education. But in the study of society exclusive concentration on a speciality has a peculiarly baneful effect: it will not merely prevent us from being attractive company or good citizens but may impair our competence in our proper field (…) Nobody can be a great economist who is only an economist—and I am even tempted to add that the economist who is only an economist is likely to become a nuisance if not a positive danger.”


The publication of the first volume of LSE’s new student journal, *Studies in Philosophy, Politics, and Economics*, marks the graduation of the first cohort of students on the School’s novel, four-year BSc in Philosophy, Politics and Economics (PPE). The journal’s name is a salute to a collection of essays by Friedrich Hayek, the famous philosopher, political thinker and Nobel laureate in economics who taught at LSE from 1931 to 1950. Hayek’s conviction that social scientists and social reformers stand to benefit from interdisciplinary training is the guiding idea of this degree. It provides rigorous training in all three disciplines, including the formal methods they use, up to an advanced undergraduate level. (In this, it stands apart from many other PPE degrees, which permit students to drop one of the three disciplines after the first year.) It also aims to help students become imaginative, independent, wide-ranging thinkers, who know how to draw on social science and moral principles to make compelling arguments. To this end, it includes a number of tri-disciplinary courses, including a Research Seminar taught by scholars with extensive policy-making experience; a Capstone Project, which includes research in teams on topics set by partners in international organizations, government, NGOs and the private sector; and PPE Applications, which covers a series of topical policy questions that can be answered effectively only by drawing on all three disciplines. This journal aims to publish some of the best work on these courses.
This issue is devoted to work from PPE Applications. This course made the production of sharp, engaging papers, improved by feedback from other students, part of ordinary coursework. Each student wrote three short formative papers on which they received comments from staff. They then chose one paper on which to receive a further round of double-blinded, constructively critical comments from at least two student reviewers. Finally, they submitted a revised version of this essay, along with a response to reviewers, as a summative assignment. Here, we collect some of the papers which stood out for their originality, interdisciplinarity, and careful reasoning.

The contributors draw on insights from economics, behavioural science, political science and political philosophy to address a range of fascinating questions. Carlo Bellomo and Alessandro Luciano discuss the case for so-called “sin taxes”; Cheryl Ting examines under which conditions participants in medical research are exploited; Emike Ahmed considers how to incentivise innovative research; Preeti Pasricha weighs Britain’s obligations to compensate India for colonial plunder; Ella Creamer and Valentin Wiesner engage with principles for making policy decisions under uncertainty; and Stefan Nielsen criticizes Martin Luther King’s criteria for deciding when laws are just. The final part of this issue, with contributions by Andrea Villalba Cuesta, Edgar Akopyan, Qiang Hua, Florence Roughton, and Lynne Sakr, is devoted to a symposium on a topic which sparked heated debate in our seminars: Leif Wenar’s proposal to ban sales of natural resources by highly repressive governments and to institute tariffs on countries who violate this ban.

I am grateful to the colleagues involved in bringing this journal to fruition: Liam Kofi Bright, Lucy Lambe, Ewan Rodgers and Lichelle Wolmarans. We hope it provides a showcase for our students’ work and thereby contributes to scholarly and public debate on matters of importance.
A paternalistic justification for excise taxes on tobacco, alcohol and sugar

Carlo Bellomo

Abstract

This paper evaluates the Task Force on Fiscal Policy for Health's proposal to implement high excise taxes on tobacco, alcohol and sugar. It argues that the Task Force's proposed justification of preventing negative externalities due to the consumption of these goods is inadequate, but that the policy can be justified on soft paternalistic grounds. Moreover, it argues that excise taxes should be combined with policies which do not rely on price in order to positively affect the consumption habits of price-insensitive individuals.

Keywords: Paternalism; Taxes; Tobacco; Sugar; Alcohol

A central policy aim of the United Nations' 2015 Sustainable Development Goals (SDGs) is reducing mortality from Non-communicable Diseases (NCDs). The Task Force on Fiscal Policy for Health has advised governments to implement large excise taxes (taxes on goods levied at the moment of manufacture) on the leading causes of NCDs (tobacco, alcohol and sugar), or to increase these taxes if already in place, in order to curb consumption. In this essay I evaluate this recommendation by first outlining the Task Force's suggestions and motivations and then examining three different lines of criticism. One critique focusses on the lack of empirical evidence to back the negative externality justification for the proposed policy, another on the policy's implications for individual autonomy as a paternalistic measure, and the last centres on its distributional impact. My conclusion is that the Task Force’s recommendation is an acceptable policy as I deem the form of paternalism engaged in to be permissible. However, I argue that the excise taxes proposed should be complemented with other policies to ensure that less price sensitive consumers are positively affected too, and that more evidence regarding negative third-party effects must be presented to make the proposal more compelling.

In their report, the Task Force begins by identifying tobacco, alcohol and obesity, which is linked to the consumption of highly processed foods and sugary beverages, as the main causes of over 40 million deaths per year due to NCDs. These, they claim, impose "enormous costs on society" (Task Force 2019: 5). Furthermore, consumption of tobacco, alcohol and sugar entails "negative externalities" (Task Force 2019: 17). The Task Force thus justifies efforts to modify consumers' behaviour and curb consumption of these three products by appealing to: the generally accepted moral principle that individual freedom can
be limited to prevent harm to third parties and the corresponding economic rule that negative externalities should be corrected to ensure the optimal working of the free market. Having laid out and motivated the need to reduce consumption of tobacco, alcohol and sugar, the Task Force then provides evidence of the efficacy of measures affecting price in reducing consumption. Finally, they argue that from the range of price-affecting policies levying excise taxes is the most cost-effective policy (WHO 2017), where cost-effectiveness is measured in terms of cost per healthy life-year gained. Indeed, they argue for large specific excise taxes (based on quantity sold) since these reduce the relative price gaps between product tiers and so reduce the incentive to trade down (Task Force 2019: 12).

One of the most notable aspects of the Task Force’s proposal and supporting arguments is the absence of any reference to paternalism. Instead the authors appeal to the prevention of harms to others, a principle which is much more politically palatable. In doing so, however, they make their reasoning vulnerable to two criticisms. Firstly, if the harm we are trying to prevent is a direct consequence of the action, for example second-hand smoke, maternal smoking, drinking and obesity during pregnancy (Task Force 2019: 17), why should we implement a policy like excise taxes that also reduces consumption in cases in which no direct harm is caused? It would seem more reasonable to simply ban consumption when it causes direct harms to others, as is already done in many countries with smoking in public areas. One could argue that enforcing such policies in the private sphere is more difficult than implementing taxes which have a universal reach, including homes, and thus excise taxes can be justified in terms of prevention of direct harm. Nevertheless, difficulty in enforcement hardly justifies adopting a strategy that impacts so many outside the target group (those causing direct harm). If it were hard to identify speeders, surely we would not increase taxes on sports cars but rather take measures to better detect offenders.

I believe the Task Force is aware of the problems associated with using the prevention of direct harm as a justification for its policy. Indeed it predominantly focuses on indirect harm, which is a by-product rather than a direct consequence of the action, such as fiscal effects (e.g. non-consumers in countries with public health systems bearing part of the costs for treating illnesses due to consumption of tobacco, alcohol and sugar). However, the relevance of these harms can be subject to a second criticism of a more empirical nature. As Voorhoeve (2013) points out, one must be careful to consider not only cross-sectional fiscal effects, fiscal effects at a specific point in time, but also total lifetime effects. The Task Force focuses on the negative effect of tobacco, alcohol and sugar consumption on the former, which is supported by a significant amount of empirical evidence (Wang et al. 2011; Lehnert et al. 2013; Dall et al. 2007). It is an established fact that consumption of these substances causes various health problems. It is therefore obvious that, at any specific point in time, health expenditure for consumers will be higher than for non-consumers. By contrast, evidence about lifetime fiscal effects is lacking due to the complexity of calculating these. In one study, Van Baal et al. (2008) show that lifetime health spending was higher for healthy people than for the obese. Whilst this single study is not enough to conclude that tobacco, alcohol and sugar consumption does not have negative lifetime fiscal effects, the Task Force should present evidence in support of the claim that it does. Overall, I believe these two criticisms significantly weaken the strength of the Task Force’s arguments. Unless further hard evidence is presented showing indirect fiscal effects to be negative, a different approach must be taken to justify the proposed excise taxes.

In light of the failure of the harms-to-others approach in justifying the Task Force’s proposal, I suggest that instead of justifying the policy in spite of its paternalistic nature, an approach which validates it because of this nature should be adopted. A key criticism that can be levied against high excise taxes on tobacco, alcohol and sugar is grounded in the fact that these measures impose limits on the free choices of individuals for their own good without
their consent. In other words, they are paternalistic. When evaluating these critiques it is important to note the distinction between hard and soft paternalism. Policies are a case of hard paternalism when they limit the choices of individuals who are in possession of sufficient information and capable of thinking and deciding rationally. Such policies are justifiable only in terms of the purported benefits to the people whom they constrain. Policies are softly paternalistic if the individuals they constrain for their own good are either insufficiently informed or incapable of rational decision making (Dworkin 2007). Hard paternalism is objectively difficult to justify given its impact on the autonomy and sovereignty of individuals, as well as the fallibility of the paternalist who very well may not have the paternalised party’s benefit at heart or simply may have evaluated it incorrectly. I believe soft paternalism is a more acceptable principle, as it essentially consists of preventing individuals from making decisions which they, upon obtaining sufficient information or regaining their rational capacities, would be likely to regret. Soft paternalism is consistent with recognizing the personal sovereignty of a well-informed, rational individual (Feinberg 1986).

Are the excise taxes proposed by the Task Force a hard paternalistic measure, therefore unjustifiable, or an acceptable soft paternalistic measure? In other words, are the majority of the people affected acting rationally and with adequate information? Rabin (2013: 122) identifies three common irrational behaviours which increase the likelihood of the formation of bad habits. The first is addictive behaviour: when one engages in an activity today it increases the marginal utility of the activity and lowers the utility level of engaging in the activity tomorrow (such as the consumption of tobacco, excessive alcohol and sugar\(^1\)). Second is projection bias: a tendency to project current tastes into the future even when we know they will change. Third is present bias in which an individual “gives extra weight to well-being now over any future moment, but discounts all future moments by the same discount factor”. An individual may also be naive, not recognizing their own susceptibility to present bias.

These biases seem likely to play a role in the decision making process of most people and a large amount of empirical evidence suggest the existence of other apparently irrational behaviours when choosing (Sunstein and Thaler 2006: 238). This leads me to conclude that for a majority of individuals the excise taxes would indeed be a soft paternalistic, and therefore acceptable, measure. Taxes would be effective in counteracting these irrational behaviours. By increasing the cost of engaging in consumption, taxes would ‘correct’ the irrational underestimation of this cost by consumers due to the addictive nature of tobacco, sugar and alcohol and the biases described by Rabin. But what about those who effectively are choosing rationally and informedly? In this case they are not suffering a hard paternalistic, unacceptable policy, but rather their choices are constrained for others’ wellbeing. I believe this is a necessary sacrifice because other policies that would not affect them, such as information campaigns or self-binding facilitation, have either been proven to be less cost-effective\(^2\) or are difficult to apply in this context.

Another criticism of policies such as those advocated by the Task Force is their economically regressive nature due to the higher impact they have on lower-income groups who tend to display higher price sensitivity. Nevertheless, this critique is grounded in the idea that being under greater influence of a paternalistic measure is a negative thing. This can therefore be refuted in light of the justification of the paternalistic nature of the excise taxes: once it is accepted that a paternalistically induced change in behaviour is beneficial

---

\(^1\) There is scientific consensus around the addictive nature of nicotine and alcohol (National Institute on Drug Abuse, 2018; National Institute on Alcohol Abuse and Alcoholism, 2019). Sugar addiction is more controversial, but there is a vast amount of animal evidence and substantial human evidence which compares its effects to those, for example, of nicotine (Wiss et al., 2018).

\(^2\) Information campaigns are as cost-effective as taxes for tobacco, but less for alcohol and sugar (WHO, 2017).
for individuals in this context, more price sensitive, lower-income individuals are in effect benefiting more than others from the policy rather than being negatively affected. This conclusion leads however to a different and more powerful distributional critique: less price sensitive individuals who continue to purchase the products despite higher prices do not enjoy the positive effect of the policy, namely the reduction in harmful consumption, and are left with a lower budget for consumption of other products. In order to ameliorate this differential distributional impact, other cost-effective policies to reduce consumption of tobacco, alcohol and sugar that do not rely on prices and therefore also benefit price insensitive consumers, such as limiting advertising (WHO 2017), should be implemented.

In conclusion, imposing high excise taxes on tobacco, alcohol and sugar in order to reduce their consumption, as suggested by the Task Force, is an acceptable measure that should be enacted due to its soft paternalistic nature and the fact that it is more cost-effective and easier to implement than other solutions (WHO 2017). However, additional policies that do not rely on price to reduce consumption should also be considered in order to decrease the negative impact excise taxes would have on price insensitive consumers. Moreover, clearer evidence of the negative indirect fiscal effects of the consumption of these products should be presented in order to strengthen the argument for this policy.
A paternalistic justification for excise taxes

References


Intuitive Decisions are not Substantially Non-Voluntary: A Response to Voorhoeve and Rabin

Alessandro Luciano*

* Philosophy, Logic and Scientific Method, London School of Economics, Houghton Street, London, WC2A 2AE, UK. Email: alessandro.p.luciano@gmail.com

Abstract

I argue that Matthew Rabin's (2013) proposal for excise taxes on tobacco, alcohol and sugary drinks cannot appeal to soft paternalism. I summarise how projection bias and habit-forming behaviours cause the consumption of these goods. I then outline Alex Voorhoeve's (2013) argument that excise taxes can appeal to soft paternalism because they limit the “substantially non-voluntary” decisions of the “intuitive” system. I build on research by Mark Price and Elisabeth Norman (2008) to argue that “intuitive system” decisions are “partially voluntary” not “substantially non-voluntary” and conclude that excise taxes would limit autonomy.

Keywords: Paternalism; Intuitive Decision-making; Excise Taxes; Autonomy

In this paper, I rebut Alex Voorhoeve's (2013) argument that excise taxes on tobacco, alcohol and sugary drinks can appeal to soft paternalism in certain cases. Firstly, I outline Matthew Rabin's (2013) case for paternalistic intervention on the grounds that irrational biases lead to self-harming habits. I then analyse Voorhoeve's argument that, under circumstances where autonomy-enhancing strategies are ineffective, excise taxes have a soft paternalistic justification. Soft paternalistic policies constrain self-harming actions without consent, only when the action is ‘substantially non-voluntary’ (Feinberg 1986: 12). I oppose Voorhoeve's argument by appealing to a distinction between ‘partially voluntary’ decisions and ‘substantially non-voluntary’ decisions, leading to the conclusion that excise taxes do not have a soft paternalistic justification.

Rabin (2013) analyses a behavioural bias known as projection bias. Projection bias is the tendency to incorrectly predict the future utility of an activity by adjusting it in the direction of the current utility of that activity (Loewenstein et al. 2003). This is illustrated in Figure 1, in which t=1, etc. are periods, and the numbers are within-period utilities of various actions, both perceived and real. For example, when a consumer is calculating the utility of drinking Coca-Cola at t=1, because they enjoy it greatly at t=1, projection bias causes them to upwardly bias their utility estimates of drinking Coca-Cola at t=2 and t=3. This results in an overestimation of the total utility of drinking Coca-Cola across the three periods.
Rabin argues that projection bias is irrational and causes self-harm in the presence of bad habit-forming activities (Rabin 2013). He appeals to Becker and Murphy’s (1988) definition of an activity as habit-forming if the present marginal expected utility of consuming a good is higher when an individual has consumed more of it in the past. A habit is bad if and only if past and current consumption has the additional effect of lowering future expected utility levels. By Rabin’s and Becker and Murphy’s definitions, drinking Coca-Cola is a bad habit because drinking Coca-Cola at $t=1$ increases the within-period marginal utility (MU) of drinking Coca-Cola at $t=2$, while lowering the utility level at $t=2$ both of drinking Coca-Cola and giving up in that period.

<table>
<thead>
<tr>
<th>Never Drink Coca-Cola</th>
<th>t=1</th>
<th>t=2</th>
<th>t=3</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Always Drink Coca-Cola (Projection Bias)</td>
<td>10</td>
<td>9</td>
<td>7</td>
<td>26</td>
</tr>
<tr>
<td>Always Drink Coca-Cola (real)</td>
<td>10</td>
<td>7</td>
<td>5</td>
<td>22</td>
</tr>
<tr>
<td>Give up $t=2$</td>
<td></td>
<td></td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>Give up $t=3$</td>
<td></td>
<td></td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>MU</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td></td>
</tr>
</tbody>
</table>

_Figure 1._ Projection bias and habit-forming activities (adapted from Voorhoeve 2013: 143).

Rabin argues that we consume products that lower our expected utility because of projection bias. We then continue to consume these products because they are habit forming. This results in self-harm. Rabin concludes that governments should prevent this harm by taxing behaviour that leads to unhealthy habits, referred to as excise or sin taxes (Rabin 2013: 137).

This proposal invites the charge of paternalism. Paternalistic policies limit an individual’s choices against their will. A policy is paternalistic if the purpose of the policy is to benefit the constrained individual, and cannot be fully justified without counting the benefits to those constrained (Arneson 2005). Governments are reluctant to implement paternalistic policies because these policies have many objections. One of the prominent concerns with paternalism is that it undermines the goods of autonomy. The goods of autonomy are the benefits individuals derive from the ability to live their lives according to their own independent and authentic desires, free from distorting external forces (Christman 2012). Aside from being a good in and of itself, it is argued that autonomy facilitates moral obligation and responsibility (Dworkin 1988: 121–29). Moreover, Mill (1991 [1859]) famously argued that individual autonomy is necessary to achieve human excellence. Concepts of autonomy focus on a propensity for self-rule or self-governance. The idea of self-rule contains two components: the independence of one’s desires and choices from manipulation by others, and the capacity to rule oneself (Dworkin 1988: 61; Arneson 2005). By limiting individual free choice, paternalistic policies reduce autonomy since they either distort our authentic desires or prevent us from acting upon them.

Soft paternalistic policies similarly constrain self-harming actions without consent, but only do so “when conduct is substantially non-voluntary” (Feinberg 1986: 12). Voorhoeve argues that actions are substantially non-voluntary when agents are “insufficiently informed or inadequately capable of rational self-governance” (Voorhoeve 2013: 141). By limiting only non-voluntary actions, soft paternalistic policies are considerably less contentious since the goods of autonomy are not diminished. In this way, soft paternalism balances a concern for...
Intuitive Decisions are not Substantially Non-Voluntary

individual well-being with a respect for autonomy. Voorhoeve (2013) also argues that some excise taxes can appeal to soft paternalism if autonomy enhancing techniques are not effective at combatting projection bias. If projection bias can be corrected through autonomy enhancing techniques, this is preferable to taxation since it improves consumer autonomy rather than limiting it. For example, if the cause of projection bias is a lack of information (evidence on the declining utility of Coca-Cola consumption over time), then the best policy for well-being would be education. This policy would reduce 'irrational' action and enhance autonomy (Voorhoeve 2013: 142).

However, Rabin (2013) and Voorhoeve (2013) argue that some forms of projection bias cannot be effectively corrected through autonomy enhancing techniques. For example, projection bias can still be prevalent when individuals are adequately informed. Rabin refers to several studies which show that consumers’ current preferences influence their judgements of the future despite the fact that consumers possess abundant evidence that their tastes will change (Read and Van Leeuwen 1998; Lowenstein et al. 2003). If consumers are still subject to projection bias despite being sufficiently informed, then their actions could be classified as substantially non-voluntary. One explanation of consumer behaviour in these circumstances is that projection bias is the result of decision-making by the intuitive system (Kahneman 2002). The decisions of this system are fast, automatic, and easy. The intuitive system is often contrasted with the reasoning system whose decisions are slow, deliberate, controlled and effortful. Intuitive decisions are strongly influenced by current stimuli meaning that projection bias could be the result of reasoning faults by this system. Consequently, Voorhoeve (2013: 144) argues that these intuitive system decisions are substantially non-voluntary.

Voorhoeve (2013: 144) justifies his appeal to soft paternalism by arguing that intuitive system decisions cannot be corrected with autonomy enhancing techniques. Kahneman (2002) notes that we rely on the intuitive system for a variety of decisions because we have limited cognitive resources which must be used sparingly. It is therefore sensible and necessary to make many decisions in an automated and quick manner. Consequently, it would not be prudent to encourage consumers to deliberate over these consumption decisions. Voorhoeve (2013: 145) concludes that, since autonomy enhancing techniques are not effective in these circumstances, excise taxes have a soft paternalistic justification.

I disagree with this conclusion on the grounds that the automatic decisions of the intuitive system are not substantially non-voluntary. That intuitive system decisions do not involve effortful deliberation does not render them substantially non-voluntary. Voorhoeve argues that actions are substantially non-voluntary when agents are “insufficiently informed or inadequately capable of rational self-governance” (Voorhoeve 2013: 141). Projection bias caused by a lack of information does not have a soft paternalistic justification because it can be corrected with autonomy enhancing techniques. Thus, Voorhoeve is effectively arguing that the intuitive system is incapable of rational self-governance when exposed to current stimuli.

Price and Norman (2008: 29) argue that within the two systems framework (Epstein 1994; Kahneman 2002) there is a tendency to associate non-conscious processing with the intuitive system and conscious processing with the reasoning system. However, there is considerable ambiguity regarding how to map the notion of intuitive processing onto the distinction between conscious and non-conscious processes. They argue that the hallmark of consciousness is “control over the behavioural influence of information, and the ability to integrate the information in a flexible manner with changing contextual demands and executive goals” (2008: 30). Price and Norman’s use of conscious behaviour is extremely similar to autonomous behaviour in that both emphasise control and self-governance. This quality is supported by considerable empirical data (Jacoby et al. 1993; Merikle and
Daneman 1998). The authors conclude that intuitive processes can be conscious processes. Taking Price and Norman’s definition of conscious processing, their argument implies that we have some form of control over our intuitive system. For example, intuitive decisions can often be shortcuts to conclusions we have repeatedly made through the reasoning system (Koriat 2000). If we exercise conscious control over intuitive decisions then we are capable of self-rule when making consumption choices regarding tobacco, alcohol and sugary drinks. This would imply that excise taxes do not have a soft paternalistic justification since our decisions are not substantially non-voluntary.

Shiffrin’s (1997) work also supports the conclusions of Price and Norman (2008). He argues that automatic processes can yield representational outcomes that we are consciously aware of. If, according to these authors, we can have some form of control over the intuitive system, then the decisions of this system are not substantially non-voluntary but are partially voluntary. Limiting these decisions undermines the goods of autonomy, albeit less than limiting the choices of the reasoning system. This view is supported by evidence that the reasoning system regulates the intuitive system (Kahneman 2002). If the intuitive system reaches a conclusion that is in conflict with our reasoning system then the reasoning system overrules that decision. This evidence gives weight to the notion that intuitive system decisions are somewhat autonomous and that we possess a capacity for self-governance even when using this system.

Voorhoeve (2013) and Rabin (2013) portray the intuitive and reasoning systems as dichotomous, separate entities. However, these systems should be seen as interdependent, mutually reinforcing partners. To argue that one system is fully conscious while the other is fully non-conscious does not consider the myriad of ways autonomy can be exercised. Intuitive feelings can be used to guide behaviour in a controlled and contextually sensitive manner even when biases are present (Price and Norman 2008). For example, imagine an individual who sees a child drowning in a river. This individual chooses to dive into the river to save the child. This decision was taken by the intuitive system as it was under extreme time pressure (Sunstein 2015). Luckily, they are able to save the child and return to land where they receive praise for their noble act (Wilkinson 2013). However, this individual also suffers from the overconfidence effect. This well-established bias causes a person's subjective confidence in his or her ability to be greater than the objective assessment of that ability (Pallier et al. 2002). The overconfidence bias caused this individual to overstate the expected utility of saving the child because they underestimated the probability of drowning. The effect of overconfidence bias on this decision is the same as projection bias on consumption decisions, both cause an individual to exaggerate the expected utility of an action. I argue that this intuitive decision is partially voluntary and autonomous despite being subject to overconfidence bias. Most people would agree that the decision to save the child was autonomous and worthy of praise. If the decision taken by the intuitive system was substantially non-voluntary it would be inappropriate to praise this individual’s selfless courage. The individual’s intuitive decision reflects the values and beliefs held by their reasoning system (Sunstein 2015). This example shows that we regard intuitive system decisions to be autonomous and voluntary even if they are subject to bias. Analogously, intuitive system decisions which are subject to projection bias should be seen as partially voluntary.

This example rebuts Voorhoeve’s (2013) soft paternalistic justification, by showing that partially autonomous decisions can be made by the intuitive system even in the presence of bias. Consequently, excise taxes do not have a soft paternalistic justification since they inhibit partially voluntary action. Instead, excise taxes must be justified in a context wherein they diminish individual autonomy. To conclude, I have argued that justifications for excise taxes on tobacco, alcohol and sugary drinks cannot appeal to soft paternalism. Whilst I
agree with Rabin's model of how we acquire and sustain bad habits, I disagree that these decisions should be classified as substantially non-voluntary. Instead intuitive system decisions are partially voluntary since we can exercise some degree of conscious control over the intuitive system and regulate these decisions through the reasoning system. Thus, Rabin's policies undermine the goods of autonomy and cannot appeal to soft paternalism.
References


Can Voluntary Participation in Medical Research Be Exploitative?

Cheryl Ting*

* Philosophy, Logic and Scientific Method, London School of Economics, Houghton Street, London WC2A 2AE, UK. Email: c.s.ting1@lse.ac.uk

Abstract

This paper critiques the claim that medical research is not exploitative if participants volunteer to undergo it. I argue that consent cannot be a sufficient basis for non-exploitation since agents can reasonably consent to unfair distributional outcomes given a grossly unjust status quo. Hence, any evaluation of the fairness of a research programme must be able to account for normatively relevant structural sources of injustice. This may be done by incorporating an inter-transactional parity condition into the conventional transactional account of exploitation, and supplementing it with a framework capable of accounting for wider, community-based effects.

Keywords: Distributive Justice; Exploitation; Medical Research Ethics

In this paper, I present a critique of the libertarian claim that medical research is not exploitative whenever individuals volunteer to participate. My argument is twofold. First, I argue that the moral wrong of exploitation can be present even where valid consent is obtained, as agents can reasonably consent to unfair distributional outcomes if doing so improves on a status quo that is highly unjust to begin with. Second, following Vida Panitch (2013) and Danielle Wenner (2018), I argue that any conceptual framework for evaluating the fairness of some distribution of benefits must therefore be able to account for normatively relevant structural sources of injustice. I conclude by acknowledging that the relationship between exploitation and moral permissibility in a non-ideal world is not clear cut. Therefore, conclusions for policy and the regulation of medical research in practice must transcend the question of exploitation.

It is useful to begin by defining exploitation. A transaction is commonly understood to be exploitative if one party A takes unfair advantage of another party B in the transaction. This may arise from a failure in either: a) the transactional process, i.e. from a lack of valid consent on B’s behalf; or b) the transactional outcome, i.e. where B does not receive her fair share of the benefits and burdens generated by the transaction. Taking voluntariness to mean informed consent, the claim that medical research cannot be exploitative if participation is voluntary amounts to the claim that satisfying a) is a sufficient condition for
Can Voluntary Participation in Medical Research Be Exploitative?

Can Voluntary Participation in Medical Research Be Exploitative?

In other words, as long as valid consent is obtained, the outcome of a transaction cannot be deemed unfair or wrongful in a way that justifies intervention.

What are the grounds for this claim? Operationalising the concept of fairness in the definition of exploitation requires an underlying theory or standard of distributional justice. Our aforementioned claim is likely grounded in the libertarian view of distributive justice as being whatever a free market, characterised by voluntary market relationships between informed, competent agents, dictates. On this view, the fact that A gains more than B is not in itself morally troubling as long as B receives a fair share of the benefits; wherein fairness is established not by any objective measure but by B's own subjective determination of the matter, as captured in the informed consent condition.

However, this account of justice cannot hold given the imperfections of markets in the real world. The notion that fairness can be adequately captured by informed consent relies fundamentally on the assumption that agents have the ability to negotiate fair terms and turn down transactions that they deem unfair. However, this is unlikely to obtain in situations where the baseline positions and bargaining power of agents are highly unequal. Indeed, these are the very conditions that enable the taking of unfair advantage, since in the absence of choice disadvantaged agents can rationally consent to inequitable agreements as long as doing so improves on their expected outcomes.

For instance, concerns about exploitation in medical research usually centre on clinical trials conducted by first world pharmaceuticals in low-income countries, where local standards of care and healthcare provision are often significantly poorer. In these settings, there is usually a surplus of willing participants for whom participation in such trials is their only means of obtaining any treatment, and therefore have virtually no bargaining power. This allows research sponsors from high-income countries to profit by utilising low-income settings for riskier research, or to reduce costs by offering far lower standards of care than would be required in an equivalent trial in a high-income setting (Wenner 2018). In such cases, wrongful exploitation occurs even if valid consent is obtained, as the better-situated party, A, takes inappropriate advantage of structural inequalities and B's disadvantaged position within them to offer B a worse distribution of benefits and risks than she might have otherwise demanded as fair. Under such conditions, a more appropriate baseline for fairness might be an assessment of what an individual would consent to in a counterfactual transaction in which the exploitable disadvantage is removed (Mayer 2007).

This claim, however, may be challenged by those who question the salience of structural or background inequalities to the fairness of specific transactions. According to the non-worseness principle, for example, it cannot be morally worse for A to interact with B than not to interact with B if i) the interaction is better for B than non-interaction, ii) B consents to the interaction, and iii) the interaction has no negative effects on others (Wertheimer 1996). Appealing to a principle like this, one could argue that offering poorer terms of contract in low compared to high-income settings, whilst not ideal, is not necessarily inappropriate or exploitative where doing so does not deprive low-income participants of any benefits they would otherwise be entitled to, and in fact still benefits them.

However, such a position provides an unsatisfactory basis for understanding exploitation, particularly in the context of international medical research. There are strong reasons to regard background conditions as normatively relevant to assessments of fairness, since ex-ante inequalities and power differentials can greatly influence the ex-post distributive justice of individual transactions. This is particularly so given the injustice of existing structural inequalities. Taking the luck egalitarian perspective that inequality in a distributional outcome is only just insofar as it is grounded in differences in morally relevant features such as choice
and desert (Arneson 1989), the fact that some agents are significantly disadvantaged on grounds of morally arbitrary characteristics, such as birthplace or geographical location, is surely unjust. Hence, principles like non-worseness that take an unjust global institutional order as an acceptable baseline, and thereby grounds for denying an agent a fair share of the benefits she helped produce, cannot be an adequate basis for understanding exploitation.

Indeed, considering the historical and sociological origins of global structural inequalities, a case can be made that members of the developed world owe a collective duty of redress to the global poor for their part in bringing about this injustice. This would lend support to the adoption of a more demanding account of the duty of non-exploitation held by medical researchers conducting trials in low-income countries. However, it can be argued that even if such a collective duty of redress exists, it is not immediately clear how it would translate into duties on the part of specific individuals or institutions. On the one hand, appealing to such a collective duty as justification for strengthening the duty of non-exploitation in medical research seems to place an undue and disproportionate moral burden on researchers, considering that most developed-world citizens do nothing to benefit the global poor to no great moral admonition.

On the other hand, it can be argued that researchers do owe special obligations to their subjects that other developed-world citizens do not, due to the nature of the relationship between researcher and research subject. According to the interaction principle, one has special responsibilities to those with whom one interacts which one would not have if one had chosen not to interact with them at all (Wertheimer 1996). This accords with a respect-based view of exploitation: if exploitation consists in deliberately interacting with another in a way that degrades or fails to respect their inherent value, then it is in some respect worse than neglect, even if neglect can sometimes have worse consequences (Sample 2003). Under this view, mutually beneficial studies can still be exploitative if they employ ethical double standards because the moral wrong of exploitation must be understood in deontological rather than purely consequentialist terms. According to egalitarianism, moral agents identical in non-arbitrary respects are deserving of the same concern and respect; hence to deny a research subject the share of benefits that would be enjoyed by another who made the same contribution, on the basis of arbitrary differences, is to deny them the same concern and respect. Consequently, if developed-world researchers choose to conduct research in vulnerable populations, they must treat the wellbeing of their subjects with the same concern and respect as they would equivalent subjects in non-vulnerable populations, particularly where they share in a collective duty to redress the injustices that birthed these vulnerabilities. Where they not only fail to do this, but deliberately seek to profit from the injustices, there should be little doubt that they have committed the gross moral wrong of exploitation (Panitch 2013).

The upshot of this is that any conceptual framework for assessing fairness and exploitation in international medical research must go beyond purely intra-transactional features, such as consent, and account for the structural and background inequalities that affect it. What might such a framework look like? Panitch (2013) argues that this may be achieved by incorporating a requirement for inter-transactional parity into the conventional transactional account of exploitation, so that the fairness of the benefit share enjoyed by developing world research subjects is judged not relative to their own meagre starting points, but to those of their less vulnerable counterparts in the developed world. This ensures that any unfairness in the ex ante status quo is not carried forward into the distributional outcomes of transactions. Thereby, equal concern and respect for the wellbeing of all research participants can be shown.
Lending greater consideration to structural inequalities may also require us to broaden our conceptualisation of exploitation. Wenner (2018) argues that medical research programmes cannot be reduced to isolated transactions between researchers, sponsors and participants, because they produce benefits and burdens that apply to entire communities rather than to individuals alone. For instance, international clinical trials may benefit host communities by facilitating skill and resource transfers from high to low-income countries. Conversely, host communities may incur costs if scarce local resources are diverted to research that would ultimately benefit only developed world patients. As community-wide benefits or costs can significantly impact the overall size and distribution of the benefits of a programme, they can exert some influence on global structural inequalities, and are therefore of normative significance to broader assessments of fairness and exploitation. Perhaps unsurprisingly, many of the solutions currently proposed as means of ameliorating exploitation in international clinical research focus on ensuring that research delivers sufficient benefits to host communities, such as being responsive to the health needs of the local community, rather than to individual participants. This suggests that the purely transactional account of exploitation currently assumed in research ethics is insufficient, and a conceptual framework that can address the exploitation of communities as distinct from the exploitation of individuals is required as well (Wenner 2018).

In this paper, I have argued that consensual medical research can nonetheless be exploitative by showing that consent cannot be a sufficient condition for fairness in situations where existing inequalities significantly constrain one party’s ability to bargain on terms. In such situations, any approach for evaluating fairness or exploitation in a particular transaction or programme must be able to account for any wider, structural sources of injustice that are normatively relevant. However, it should be noted that whilst all exploitation is morally reprehensible, the relationship between exploitation and its permissibility in various contexts in the real and non-ideal world is not clear. Arguably the biggest challenge to introducing more demanding regulatory standards for fairness is the fact that it could result not in more equitable benefit sharing but in non-compliance or withdrawal, which could lead to worse outcomes if the implicated research was nonetheless Pareto optimal. Therefore, conclusions for policy and regulation – such as whether specific trials should be approved or specific interventions to ameliorate risk of exploitation be introduced – should transcend normative considerations alone and be informed by empirical assessments of expected outcomes.
References

**Arneson, R. J.** 1989. “Equality and equal opportunity for welfare.” *Philosophical Studies* 56(1): 77-93. URL: [https://doi.org/10.1007/BF00646210](https://doi.org/10.1007/BF00646210)


How to Make Innovative Research the Norm

Emike Ahmed*

* Philosophy, Logic and Scientific Method, London School of Economics, Houghton Street, London, WC2A 2AE, UK. Email: emikepahmed@gmail.com. URL: www.linkedin.com/in/emike-ahmed-a47a96105

Abstract

Scientific mavericks go against the grain to conduct innovative, higher-risk research. However, a conservative bias against such unorthodox research in contemporary science causes researchers to conform to safer research. Adrian Currie (2017) argues that to rectify this bias, maverick thinking should become the norm, instead of simply encouraging more mavericks. This paper presents and evaluates Currie’s case for making maverick thinking the norm through changing the incentives in scientific research, focusing on Currie’s proposal to change the funding system from meritocratic peer review to Shahar Avin’s (2015) lottery system, which randomly allocates funding. This paper argues that the lottery considerably reduces conformity, within limits.

Keywords: Science; Maverick; Funding; Lottery; Diversity

This paper presents and evaluates Adrian Currie’s (2017) argument that maverick thinking should become the status quo in scientific research. The term ‘maverick’ refers to an individual who acts in an independent way, operating outside the status quo, shaping history (Currie 2017). Maverick thinking refers to the innovative thinking that mavericks exhibit. In scientific research, these mavericks shape history through their willingness to go against the grain with their innovative research. Currie explains that the status quo of scientific research is susceptible to bias against innovative projects. Instead of simply encouraging more brave mavericks to go against the scientific status quo, Currie argues that we should address the issue by promoting the innovative thinking we associate with mavericks as integral to scientific research. Innovation should no longer be seen as ‘maverick’, it should be the norm.

First, I explain why the current state of scientific research is not as accommodating of maverick research as it previously was. Following this, I expand on Currie’s argument that the current allocation of incentives in scientific research should be altered to make maverick thinking the norm and explain why Currie argues for this solution. I shall evaluate Currie’s proposal for changing the current incentives through using Shahar Avin’s (2015) lottery funding allocation system, by critically analysing this system’s ability to make innovative research the norm.

Currie argues that the current state of science is not as accommodating of mavericks as it previously was, by comparing the contemporary scientific status quo to previous eras in
which mavericks, such as Newton and Buffon, adopted unorthodox methods to reach their conclusions. In the 1760s, Buffon deduced the time required for minerals to cool to room temperature using his sense of touch. Centuries before that, Newton developed his theory of optics by slotting a thick needle between his skull and eye to alter the shape of his eyeball (Currie 2017). In that era, anyone with money and time on their hands could conduct research. By contrast, to conduct successful scientific research today, one must have a degree, be published, peer reviewed, and granted funding (Currie 2017). Despite the value of these processes in upholding good scientific conduct and thereby enhancing the validity of research, Currie argues that, due to the conservative nature of science, these processes can lead to conformity and the standardisation of research output.

To illustrate this conformity, I shall outline the peer review system Currie mentions. Peer review consists of a panel of scientific peers who review research proposals and allocate funding on a meritocratic basis (Avin 2017: 2). Brezis (2007: 693) expands on the conservatism in science which Currie highlights by explaining that, due to funding resources being finite, a conservative funding bias can occur where reviewers favour safer, more conventional research over higher-risk innovative research. Under this system researchers are incentivised to conform to the preferred, safer research design standards to secure funding. The national Research Excellence Framework (REF) exemplifies this, using the peer review system to rate the quality of research UK universities produce, and allocating funding correspondingly. McCulloch’s (2017) study of academic researchers’ experiences under the REF found that some researchers face pressure from their employer universities to confine their research projects to more conventional topics, in order to receive funding, illustrating how the conservative funding bias leads to conformity. Additionally, in Murphy and Sage’s (2015: 33) study of the implications of REF, some researchers note that their departments dictate what they research, turning down projects that are more interesting and relevant, but take longer to conduct. This behaviour by departments facilitates the production of more research within the REF timeframe to maximise funding. This evidence supports Currie’s claims that systems such as peer review encourage conformity. This system differs from Newton and Buffon’s era in which mavericks could operate without being susceptible to the conservative funding bias that exists in today’s peer review system. Hence, conducting maverick research is currently more difficult than in the past.

I shall now expand on Currie’s argument that maverick research should become the norm. Currie argues that we should not tackle the difficulty of being a maverick under the current system by simply encouraging more individual researchers to become mavericks. Instead, the entire system should change such that riskier, unorthodox research becomes the status quo (Currie 2017). Currie urges us to move away from romanticizing mavericks, such as Newton and Buffon, as solitary fighters moving against the status quo. Rather than encouraging more mavericks to go against the tide, the tide itself should change – higher-risk innovative research should be integral to science. Currie’s key reason for arguing that the entire system should change is the issue of diversity, which will persist if innovative research is left solely to mavericks.

The conservative funding bias causes the diversity problem because wealthy, white men tend to have greater emotional, financial and physical support than other demographics. This support is required to pursue high-risk research, thereby making these men less constrained by the conservative bias (Currie 2017). Additionally, institutional bias may occur within peer review where panellists may prefer research from prestigious institutions or individuals (Brezis 2007: 693). This bias also disproportionately affects people from more disadvantaged socioeconomic backgrounds due to the correlation between educational and economic inequality. For example, the proportion of pupils on free school meals that attend university is 21%, compared to the 85% who attend from private schools (Teachfirst 2019).
This foundational inequality causes a disparity in the institutions students from these lower socioeconomic backgrounds can access, compared to their wealthier counterparts. Consequently, the institutional bias in peer review contributes to the diversity issue as it disproportionately affects researchers from less advantaged backgrounds. The diversity problem will persist if we simply encourage more scientific mavericks without changing the norm.

Not only is this a moral issue of injustice against disadvantaged demographics, it also leads to less innovative outcomes. Diversity proves useful in reaching optimal outcomes, due to the breadth and quality of output that accompanies diverse perspectives and ideas in research (Currie 2017). The benefits of diversity are not exclusive to science - they are also evident in business. According to McKinsey, a management consultancy firm, companies in the top 25% of executive-board ethnic diversity were 35% more likely to outperform their peers financially (Hunt et. al. 2015). The diversity issue in scientific research leads to suboptimal innovation outcomes, because the lost value of having multiple perspectives and ideas hampers the creativity of research output. Furthermore, this issue places a wedge between the helper and the helped. Currie argues that these frontrunners are not accountable to the wider population, whom their research affects. They are at liberty to apply their resources to conduct whatever research they see fit, which may not be in the general public's interests. In other words, entrusting innovative research to individuals from a certain demographic leaves the general population vulnerable to the whims of these mavericks (Currie 2017). Thus, Currie advocates for making innovation the norm within scientific research by making maverick thinking more accessible, instead of simply encouraging more mavericks which reinforces the diversity issue.

To make scientific research more accommodating to maverick thinking, Currie proposes to change the way incentives are allocated within science in order to remove conformity, which results from the conservative bias against innovative research projects (Currie 2017). To this end, he suggests that science should diversify the indicators that determine scientific success thereby incentivising researchers to be more innovative, preventing conformity. Additionally, more institutions should be dedicated to conducting exploratory research. Currie draws on Avin’s (2015) lottery proposal, in which experts filter out the worst research proposals, select the best research, and randomly allocate funding to the remaining proposals through a lottery. Avin recognises that this filtering process may still contain bias. However, he claims that more innovative proposals will get to the lottery stage, thus funding is allocated through a less biased mechanism than the current system allows (Avin 2015: 176). Currie argues that using a lottery reduces the peer reviewers’ power over researchers, thus reducing the current conformity that occurs when researchers strive to please the panel to gain funding.

For innovative scientific research to become the norm, the ability to conduct maverick research must be accessible to all, rather than to predominantly rich, white, male scientists. Avin’s lottery is undoubtedly an improvement on the peer review system in this regard. The randomized element of the lottery contributes fairness to the funding allocation, limiting the conservative bias against innovative projects. This system also, to an extent, considers the institutional constraints that prevent certain demographics from even entering the lottery. Certain conditions are required of funding applicants, such as having a PhD from certain institutions. Avin recognises the diversity issue that arises from such criteria, and therefore calls for a reformation of the indicators that make up the conditions for applying for funding (Avin 2015: 174). Avin argues that indicators should include the minimal possible level of qualification required for research, to create fairness and representation by increasing the pool of applicants accepted to participate in the lottery. This way, more diverse, innovative
research can be considered, and entered into the lottery, thereby altering researchers’ current incentives to choose safer research topics.

Despite the merits of Avin’s proposal in fostering innovation as a norm, more can be done to make maverick status attainable for a diverse demographic of researchers so that innovation is the status quo rather than only attainable for a select few. For example, to bridge the educational inequality at a foundational stage, more diversity and inclusion initiatives should be run to encourage students from disadvantaged socioeconomic backgrounds to pursue careers in scientific research. This should bring a broader variety of perspectives into scientific research, thus fuelling innovation. Additionally, diversity should be enforced within the panel of peers who conduct the pre-lottery research proposal filtering. The panel should be made up of scientists from varied academic backgrounds, who specialize in a range of disciplines and have attended a broad selection of universities. There should also be gender and ethnic diversity within the panel to prevent the maverick status from being exclusive to rich, white men. This will provide multiple perspectives on the panel, ensuring that decisions in the pre-lottery process are based on merit, and not on a unilateral view amongst peers of how innovative research should look.

In conclusion, Avin’s lottery proposal is an improvement on the current meritocratic peer review system, as it helps to reduce the conservative bias against innovative research, thereby weakening conformity. However, I have argued that this solution does not completely alter the status quo, as educational inequality at foundational stages cause the diversity issue to persist. Thus, the panel of peers who filter through the research proposals should be more diverse, and initiatives to encourage students from disadvantaged socioeconomic backgrounds to pursue scientific careers should be implemented. This would be a further step towards making maverick thinking the status quo in scientific research.
References

Avin, S. 2015. Breaking the grant cycle: on the rational allocation of public resources to scientific research projects (Doctoral thesis). University of Cambridge. URL: https://doi.org/10.17863/CAM.16172

Avin, S. 2017. “Science funding is a gamble so let’s give out money by lottery.” Aeon. URL: https://aeon.co/ideas/science-funding-is-a-gamble-so-lets-give-out-money-by-lottery


Currie, A. 2017. “Does science need mavericks or are they part of the problem?” Aeon. URL: https://aeon.co/essays/does-science-need-mavericks-or-are-they-part-of-the-problem


Identifying Colonial Harms and Reparations: A Study of Britain and India

Preetbir (Preeti) Pasricha*

* Philosophy, Logic and Scientific Method, London School of Economics, Houghton Street, London, WC2A 2AE, UK. Email: ppasricha6@gmail.com. URL: https://www.linkedin.com/in/preetbir-pasricha/

Abstract

This paper evaluates A. John Simmons’ (1995) theory of fair shares to determine how due reparations are calculated. Reparations are understood to be the effort taken to repair a harm that resulted from a wrong. Using the case study of Britain’s colonisation of India, this paper assesses three conditions assumed in Simmons’ theory, and proposes two alternative conditions. It argues that Simmons’ theory is narrow and only identifies partial due reparations. Finally, this paper concludes with a suggested alternative reparation: to implement a gesture that sincerely and publicly recognises past colonial injustices by teaching colonial history in the school curriculum.

Keywords: Colonialism; Historical Injustices; Reparations; Britain; India

Using the example of Britain’s colonisation of India, this paper argues that Simmons’ (1995) theory is narrow and only works to identify partial reparations that are due after historical harms. I make this argument by first presenting Simmons’ theory of fair shares and the concepts of ‘harm’ and ‘reparations’. Next, I apply Simmons’ theory to the example of Britain’s colonisation of India. I consider a situation in which Simmons’ theory does identify a reparation, before presenting an example which shows that Simmons’ theory only identifies partial reparations that are due. I then explain the argument for reparations, demonstrating that there exists a possible reparation that is morally significant; here, my suggested reparation is to teach colonial history in the school curriculum. Finally, I conclude that Simmons’ theory does not fully work to identify which reparations are due in my case study; it only identifies partial reparations. This conclusion speaks against Simmons’ theory as a means of identifying which reparations are due, but not against the argument for reparations overall.

Firstly, I outline Simmons’ theory of fair shares. To begin, an actor X owes reparations to an actor Y if and only if three conditions hold. The first necessary condition is that X has benefited from the injustice that Y suffered. The second necessary condition is that there are identifiable goods that one can confidently claim to be connected to, and are sufficiently comparable to, what was lost due to the injustice. This relies on making conservative
Identifying Colonial Harms and Reparations

assumptions of the potential counterfactual outcomes, such that the reparations due can be fairly determined (Simmons 1995: 158). The third necessary condition is that X must be able to transfer these goods to Y. Simmons then concludes that, under these conditions, X owes reparations to Y (Simmons 1995). In my case study, X refers to the people of Britain who committed the injustice (“Britain”) and Y refers to the people in India who were affected by the injustice (“India”). In this essay, reparations are understood to be the effort taken to repair a harm that resulted from a wrong, which includes transferring goods intended to compensate for those goods lost during the wrong (Radzik and Murphy 2015). I understand harm to be the consequence of an action that leaves an individual worse off than they otherwise would have been, had the action not happened, by setting back their interests (Mill 1859). I choose to focus on the impact on individuals and not the nation, as the effect is ultimately a summation of what is felt by individual people. I refer to the actors in question as “Britain” and “India” for simplicity.

Simmons considers an example of a stolen bicycle across different possible counterfactual outcomes. If one’s bicycle is stolen, then one has a right for this wrong to be rectified, where rectification after an injustice entails returning the share of goods, adjusted by size to the ones stolen (Simmons 1995: 160). Simmons recognises that after the passing of time or the destruction of property, the judgement of the counterfactual claim is complicated. For example, goods could be destroyed, or the affected group could change, which makes determining the counterfactual claim difficult. He thinks that the judgement of the counterfactual claim should be “conservative” (Simmons 1995: 158), with fair determination of which reparations are due. He goes on to claim that “it will certainly turn out that some past property injustices are simply unrectifiable” (Simmons 1995: 176).

Secondly, I apply Simmons’ theory to my case study to argue that his theory only identifies partial due reparations. Consider the first condition that X (Britain) has benefited from the injustice that Y (India) suffered. This seems to be true, as there is evidence that Britain’s growth occurred at the expense of India’s. British civil servants found highly paid employment under colonisation efforts and therefore benefitted unjustly from the harm (Tharoor 2015). The second condition, the identification of goods that are confidently connected to the injustice, holds as well. In this case, the Koh-i-Noor diamond is a good stolen and confidently connected to the injustice (Tharoor 2015). The third condition also holds, since it is possible for Britain to transfer this good to India as it is an undamaged, physical object. Therefore, under Simmons’ theory of fair shares, a reparation that Britain owes India has been identified. Since it has been recognised as a reparation, it should be returned under Simmons’ theory. However, this is only a partial reparation and does not encompass all reparations that are due.

To extend this point, the Koh-i-Noor diamond is not the only loss. Consider the millions of lives lost and displaced during the partition of India, the famines and destruction of industries such as the Indian textile industry (Tharoor 2015). The first condition still holds as per the previous example: Britain benefited from the injustice that India suffered. The second condition fails, because, although these are goods that can be identified and confidently said to be connected to the injustice, they are not ‘sufficiently comparable’. To fairly compensate for something, its value must be evaluated. Lives cannot be confidently assigned a value, unlike material objects (Bjorndahl et al. 2017). Consider the perspective of the individual: each individual could be someone’s child, someone’s parent, someone’s friend. These personal relationships are important and incomparable with each other; it seems morally uncomfortable to assign them a value. This argument extends the Kantian notion of dignity (Bjorndahl et al. 2017) such that if a being has dignity, its replacement necessarily incurs a cost. Thus, without a clear value, lives are not compensable, so the second condition fails. The third condition also fails: due to the passing of time, the fact that lives cannot be
replaced, and the fact that there was a change in the cast of those affected, Britain is not able to transfer these goods to India. Hence, due to the failure of the second and third conditions, Simmons’ theory seems to argue that Britain does not owe reparations to India for the losses outlined.

I argue against Simmons’ second and third conditions, by explaining that there exist reparations that Simmons’ theory does not recognise but that are still due. Consider the first premise that a harm has still been done, even if it is not measurable or identifiable. There are non-material harms, such as psychic harms, like those that are identity-related (Cohen 2009). These are unaccounted for under the second condition when identifying ‘sufficiently comparable’ goods to the injustice. For example, partitioning India harmed community relations and those who previously lived peacefully side-by-side became hostile. Consider a second premise that reparations are understood to be the effort taken to repair a harm that resulted from a wrong (as per the definition). I introduce a third premise: where harms are done, they need to be repented for. Although it is impossible to identify a reparation of equivalent value to the harm, which requires putting a value on community relations, it seems morally problematic to just leave these communities in unrest.

A critic might argue that it is relatively easy to restore these community relations to their previous level. However, this is next to impossible in this example and this holds true across cases of those impacted by ‘divide and rule’. For a start, India was partitioned resulting in two new national identities, and this is not something that can be undone. Additionally, many of the tensions that originated in the time of the British Empire still exist today: religious groups that used to live side-by-side no longer could, and some of these sentiments were passed down the generations. Communities are so vital to an individual’s identity that, by neglecting the community, the individual is also insulted (Waldron 1902: 6). Hence, if we accept this third premise that harms need to be repented for, then, following the harmful consequence of this injustice, a reparation is due. Simmons’ theory does not recognise this since it is not possible to assign value to, and hence compensate for, this harm. As a reparation is due that Simmons’ theory has not recognised this speaks against Simmons’ proposal as a method for identifying due reparations, specifically the second and third conditions outlined above, and not against the argument for reparations in general.

Thirdly, a potential rebuttal to my proposal from a defender of Simmons might contest why the third premise I proposed should be accepted, on the grounds that this is a historical injustice and the people who committed the harm, the British colonialists, are no longer alive. The rebuttal might argue that British people today who have not committed the harm should not repay for something they did not do, where they are not directly responsible. It is their ancestors who might not be alive today who hold responsibility for the harm to Indian people instead. If those who are alive today did not commit the harm, then it can be contended that Britain need not repay the harm caused to India.

I respond to this rebuttal by arguing that, even if those who committed the harm are no longer alive, this does not remove the fact that the harm occurred and is morally significant (Waldron 1992). The moral significance arises since, although those who committed the harm are no longer alive, those who are affected by the harm still are (Cohen 2009). Independence was only granted in India seventy-one years ago; thus, at least three generations are alive today. It appears to be in good faith to these people to recognise that a harm has been committed that impacted them and their family. Keeping this acknowledgement in memory is a ‘symbolic gesture’ (Waldron 1992: 7), not intended to compensate for harms, but at least to address the significance of their occurrence. This approach recognises that since a reparation is possible, it is morally relevant to believe a reparation is morally required.
Identifying Colonial Harms and Reparations

If reparations are “efforts taken to repair a harm that resulted from a wrong”, as stated earlier, one possible effort is including colonial history in the school curriculum. This restricts the emergence of historical amnesia and does not aim to directly repay for the injustice but addresses that it happened through a sincere public recognition. Being included in the school curriculum also means that lessons can be learned to prevent repeating the injustice. This argument for education was also recently rekindled. It has been recognised that through education, the story of the British Empire can be made impartial and factual, rather than a story understood through the “simplistic notions” (Singh 2019) of it being a good or bad thing. Educating about the Empire can allow potentially difficult conversations to emerge, so that the young generation, regardless of where their ancestors were, can participate in this conversation and learn from it.

In conclusion, Simmons’ theory is too narrow, failing to recognise the moral significance of all harms resulting from historical injustices. He considers whether goods are sufficiently comparable and identifiable, but by failing to consider all harms and their moral significance, his theory fails to identify all types of reparations that are due. By recognising that harms occurred, and that there exist possible reparations, such as education in the school curriculum, I argued that Simmons’ method of identifying reparations is narrow and therefore only identifies partial reparations in my case study of Britain and India. I first outlined Simmons’ theory of fair shares and the concepts of harm and reparations. Then I applied his theory to my case study, showing how it identifies reparations in the case of the Koh-i-Noor diamond. I then explained where Simmons’ theory fails to recognise all harms, in particular psychic harms. I addressed the rebuttal to my argument that those who committed the harms are no longer alive, so no-one today is responsible for the harm. I responded to this rebuttal by arguing that Simmons’ theory is too simplistic and does not consider the moral significance of harms. I extended this argument by demonstrating that there exists a morally significant reparation of including colonial history in the school curriculum. Finally, I conclude that although Simmons’ theory only identifies partial reparations, this does not speak against reparations in general, only against the conditions in Simmons’ theory.

Acknowledgements

I thank Professor Alex Voorhoeve and Dr. Liam Kofi Bright for their comments and guidance. I am also grateful to the two anonymous reviewers who provided comments that greatly improved the paper.
References


Simmons, A. J. 1995. “Historical rights and fair shares.” Law and Philosophy 14:149-184. URL: https://doi.org/10.1007/BF01001042


A Defence of Ambiguity Aversion in Policymaking

Ella Creamer*

* Philosophy, Logic and Scientific Method, London School of Economics, Houghton Street, London, WC2A 2AE, UK

Abstract

Some have argued that ambiguity aversion implies information aversion (refusing free information) and is therefore irrational (see, e.g., Al-Najjar and Weinstein 2009 and Fleurbaey 2018). However, using a simple model involving climate policy, this paper shows that such supposed information avoidance merely demonstrates a perfectly reasonable preference for time-consistency and/or cost-avoidance.

Keywords: Ambiguity Aversion; Information Aversion; Policymaking

Ambiguity aversion entails that a decision maker prefers to bet on known risks than on unknown risks (Montesano and Giovannoni 1996: 133). Objections to the rationality of ambiguity aversion have incorporated appeals to information aversion: the refusal of free information (Fleurbaey 2018: 31). However, I shall demonstrate that these arguments are misguided as they confuse information aversion with a mere desire for time-consistent and/or cost-saving preferences. I begin by outlining ambiguity-averse preferences and Nabil Al-Najjar and Jonathan Weinstein’s (2009) objection to the rationality of such preferences. Following Siniscalchi (2009), I argue that this objection is misguided. The debate will be contextualised using a climate policy scenario based on Marc Fleurbaey’s (2018: 30) health example, leading to a critique of Fleurbaey’s neglect of cost in such scenarios. I will then remodel the scenario, incorporating the cost factor, to demonstrate why being information averse in this context is not irrational and is in fact akin to being cost averse. I conclude that information aversion cannot be employed as an objection to the rationality of ambiguity aversion in policymaking.

Imagine the ambiguity-averse decision maker is faced with an urn containing 120 balls, precisely 40 of which are black. The remaining balls are red or yellow in an unknown proportion. The decision-maker chooses bets among the options listed in Table 1.
Table 1 four options.

<table>
<thead>
<tr>
<th>Bet</th>
<th>b</th>
<th>r</th>
<th>y</th>
</tr>
</thead>
<tbody>
<tr>
<td>f₁</td>
<td>10</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>f₂</td>
<td>0</td>
<td>10</td>
<td>0</td>
</tr>
<tr>
<td>f₃</td>
<td>10</td>
<td>0</td>
<td>10</td>
</tr>
<tr>
<td>f₄</td>
<td>0</td>
<td>10</td>
<td>10</td>
</tr>
</tbody>
</table>

The cell values correspond to payoffs, for example, f₁ corresponds with payoff = 10 if the ball is black. An ambiguity-averse decision-maker would display the following preferences: f₁ > f₂ (as there are definitely 40 black balls and they prefer the bet on known risks) and f₄ > f₃ (as there are definitely 80 red or yellow balls). However, if the preferences expressed are based on an assignment of precise probabilities, these preferences are paradoxical: f₁ > f₂ implies that P(b) > P(r) and f₄ > f₃ implies P(r) + P(y) > P(b) + P(y) which simplifies to P(r) > P(b). We are left with P(b) > P(r) & P(r) > P(b), which is a contradiction. This is known as the Ellsberg (1961: 651) paradox.

Al-Najjar and Weinstein (2009: 266) attempt to refute the rationalisation of ambiguity aversion by highlighting the undesirable implications of such preferences, one of which is information aversion. They cite a scenario, outlined in a decision tree, to illustrate their argument. I have redrawn their diagram as follows:

Figure 1. A Decision Tree (Al-Najjar and Weinstein 2009).

Note that the payoffs on each of the left and right branches are identical at time 1 (viz. corresponding to f₃ and f₄ in the ambiguity-averse preferences table). At Time 1, the ambiguity-averse decision-maker has the choice to take what I will call the status quo path (L) or the information gain path (R). Taking the status quo path, they will then face a choice at Time 2 between bets u and d. Alternatively, if they take the information-gain path at Time 1, they will find out whether the ball chosen is yellow or not immediately after their Time 1 decision. If the ball is yellow, they obtain a payoff of 10 and the game is finished. If not, they progress to Time 2 and have the option to bet u or d (on the same ball first selected).

I will now break down the dynamic elements of the scenario and, following Siniscalchi (2009), explain why the implications that Al-Najjar and Weinstein draw are misguided. I will represent the choice of R at Time 1 and u at Time 2 as (R,u) and so on. I define commitment preferences to be the preferences of the ambiguity-averse decision-maker at Time 1 and the actual preferences as the preferences of this decision-maker at Time 2. Dynamic consistency occurs when commitment preferences and actual preferences coincide.
At Time 1, the decision-maker's commitment preferences are \((L, d) = (R, d) > (L, u) = (R, u)\). If the status quo path is taken, the actual preferences at Time 2 map to the commitment preferences, as the payoffs are the same: \((10, 0, 10)\) for \((L, u)\) and \((0, 10, 10)\) for \((L, d)\). Therefore, in the status quo case, dynamic consistency is maintained. However, for the information-gain path the case is not so simple. At Time 2, (after learning that the ball is not yellow), the payoffs for \(u\) and \(d\) respectively are \((10, 0, *)\) and \((0, 10, *)\). The decision-maker will now strictly prefer \(u\) (as black balls are guaranteed to be in the urn). Thus, we face the following situation: their initial commitment preferences are \((R, d) > (R, u)\), however, their actual preferences at Time 2 are \(u > d\): dynamic consistency is violated. Because the sophisticated ambiguity-averse decision-maker is aware of their future actual preferences, they know that if they choose \(R\), it is as if they were to choose \((R, u)\). Given that their commitment preferences are \((L, d) > (R, u)\), they rationally choose the status-quo path \((L, d)\) to which they can commit.

Al-Najjar and Weinstein argue that because the decision-maker strictly prefers the status quo path, they can be said to be information averse as they knowingly take the path which does not grant them extra information. The authors conclude that this is irrational, since “a statistician who finds himself choosing not to look at all available data should feel rather embarrassed” (Al-Najjar and Weinstein 2009: 266). The crux of the problem with their argument is that, whilst they infer that the scenario signals information-averse tendencies, the scenario in fact signals a trade-off between information aversion and an inability to commit to preferences and simply signals the decision-maker’s desire for commitment over information. As Siniscalchi (2009: 345) argues, the scenario says little about the decision-maker’s attitude to information. By choosing the status quo path, they can be certain that they will act with dynamically consistent. We should not take this as evidence of information aversion – they simply consider commitment to be of higher importance than information gain.

To further my objection to Al-Najjar and Weinstein, I will contextualise the debate with a climate policy problem, similar to Fleurbaey’s (2018: 30) application of their argument to health policy. I will then criticise Fleurbaey’s approach to ambiguity aversion in such policy problems. In my proposed scenario, there is a climate crisis. Green technologies are being developed and it must be decided whether they are worth the risk of investing in. A new technology has been developed that harnesses the sun for electricity. It is cheaper than existing technology, thus more homes and businesses would have access to it. A policymaker working for a climate charity must decide between donating to the use of the new or the existing technology. They do not yet know whether the use would take place in country A or B, which have different climates. The key variables for the decision-maker are: (a) the chance of the technology being a success or not; and (b) cost. The chance of the existing technology being successful is known and constant across countries. By contrast, the chance of the new technology being a success is ambiguous and depends on the country in the manner displayed in Table 2.

<table>
<thead>
<tr>
<th></th>
<th>Country A</th>
<th>Country B</th>
</tr>
</thead>
<tbody>
<tr>
<td>Existing technology</td>
<td>50%</td>
<td>50%</td>
</tr>
<tr>
<td>New technology</td>
<td>(p)</td>
<td>100% - (p)</td>
</tr>
</tbody>
</table>

Table 2. Chance of success of two technologies, with \(0 < p < 100\%\).

The decision maker can choose to discover the assigned country or not. Suppose they opt not to find out the country and that the likelihood of the project taking place in country A or B is equal. Thus, the expected chance of success under the new technology is \(0.5p + 0.5\)
(100% - p) = 50%. Since the new technology is cheaper than the existing one, it is the preferred option under such ignorance.

Suppose, by contrast, that they opt to find out the country. If the country is revealed to be A, the chance of success of the new technology is ambiguous, ranging from 0 to 100%. (The same is true if the country is revealed to be B.) An ambiguity-averse policymaker may then choose the existing technology despite the higher cost. If faced with the decision whether to find out the country or not, the policymaker will likely choose not to test for it, and simply choose the new technology, as they would not then have to change their choice to a costlier option after discovering the country assigned.

Fleurbaey (2018: 31) argues that this “refusal of free and useful information” is difficult to justify when the information introduces no disadvantage. However, drawing on my prior argument, it may be justified by the fact that this choice to refuse to discover the country does not mean that the policy-maker is information averse, but rather that they rank time-consistent preferences over the information gain as they know that they cannot commit to their initial cheaper preference (of investing in the new technology) if they choose to discover the country.

A critic may object to the argument that the preference for commitment can dominate that of information, questioning why commitment is inherently valuable. In response, I would claim that information is also not inherently valuable: in this policy scenario, choosing the information-gain path (viz. finding out the assigned country) actually carries a cost. The reason why commitment is valuable is because it guarantees the avoidance of this cost. Fleurbaey (2018: 31) argues that the information induces “no strategic disadvantage”, yet there are few examples of disadvantage more explicit than a direct cost being induced. If we assume that taking the cheapest option according to a cost-benefit analysis is rational, then it is rational to choose the status quo path and refuse the extra information. The preferences expressed here are not concerned with avoiding information but with avoiding cost. This is of particular importance in the climate policy scenario as the extra money could be invested in other technologies.

In sum, my abstract and applied scenarios demonstrate that we cannot take the ambiguity-averse decision-maker’s choices as evidence of information aversion, but rather as a preference for commitment and/or cost aversion. Thus, Al-Najjar and Weinstein and Fleurbaey’s claims do not amount to rejecting the rationality of ambiguity aversion in policymaking – if it is to be rejected, a different appeal must be made. Their assumption that information is always valuable is false: it may carry a cost which must be accounted for in decision modelling. These costs are especially important for policymaking scenarios where some resources could be better distributed elsewhere without sacrificing policy effectiveness.
References


Risk Aversion is Grounds for Rejecting the Ex Ante Pareto Principle

Valentin Wiesner*

* Philosophy, Logic and Scientific Method, London School of Economics, Houghton Street, London, WC2A 2AE, UK. Email: valentin.wiesner@gmx.net. URL: https://www.linkedin.com/in/valentin-wiesner-3b04a3147/

Abstract

This paper argues against applying the ex ante Pareto principle universally as a decisive criterion for policy decisions under risk. The ex ante Pareto principle relies on the measure of expected utility in order to assess people’s prospects under each policy. Expected utility theory, however, does not sufficiently account for people’s risk attitudes as distinct psychological attitudes. Therefore, policies might expose people to risks which they would rationally not want to take. It concludes that the ex ante Pareto principle cannot be justified from a contractualist standpoint and loses its appeal.

Keywords: Allais Paradox; Expected Utility Theory; Risk Aversion; ex ante Pareto Principle

In this paper, I argue against the universal applicability of the ex ante Pareto principle (hereafter EAP) as a decisive criterion in cases of policy decision making under risk. I argue that the necessary measure of well-being standardly employed by the EAP, expected utility (EU), does not adequately capture the risk preferences of individuals. Consequently, certain policies may subject individuals to risks they would not willingly take, leaving the EAP unjustifiable from a contractualist standpoint in some cases. First, I outline the EAP and explain its appeal for decision makers. I then motivate my critique by using a stylized case of a policy decision. In order to assess the validity of EU as a criterion for decision making, I delineate what it in fact captures, arguing that the independence axiom in expected utility theory (EUT) is inadequate for measuring attitudes to risk. This point is illustrated using the Allais paradox, which systematically violates the independence axiom in individual decision making. Concrete cases are taken from experiments in health outcomes in similar contexts to exemplify the argument’s direct relevance for policy. I conclude that decision makers must take this shortcoming of the EAP into account when making policy decisions concerning risk.

Policy decisions often have to be made regarding uncertain states of the world. The EAP is a principle that is often invoked to make such policy decisions involving risk. The principle "holds that if an alternative has higher expected utility for every person than every other alternative, then this alternative should be chosen" (Fleurbaey and Voorhoeve 2013: 114). This principle is very appealing to decision makers as it is often claimed, it can be justified...
Risk aversion and the ex ante Pareto principle

from a contractualist standpoint of moral rightness, meaning that “an action is right if and only if it is justifiable to each person” (Frick 2013: 132). As far as EU forms basis of decision making, it follows that it is justifiable to choose a policy through which EU is higher for every person than in any other alternative policy. EU maximising individuals are not reasonably expected to reject an alternative that gives them the highest EU. The decision maker would in this manner determine what is in people’s best interest, given the available information.

To illustrate the principle through a stylized case: A policymaker might be faced with a decision about compulsory cancer screening. There are minor side effects to being screened, which means that for many individuals the outcome utility if screened is slightly lower than if the test is not conducted. If a person is not screened, they have a 1% chance of dying from cancer, but are otherwise slightly better off. The EUs are expressed in Table 1.

<table>
<thead>
<tr>
<th>Probability of state of the world</th>
<th>Individuals’ utility of outcome for policy X (no screening)</th>
<th>Individuals’ utility of outcome for policy Y (screening)</th>
</tr>
</thead>
<tbody>
<tr>
<td>No cancer (P=0.99)</td>
<td>0.86</td>
<td>0.85</td>
</tr>
<tr>
<td>Cancer (P=0.01)</td>
<td>0</td>
<td>0.85</td>
</tr>
<tr>
<td>Expected Utility</td>
<td>0.8514</td>
<td>0.85</td>
</tr>
</tbody>
</table>

Table 1. Expected utilities of cancer screening policies.

To calculate EU in the bottom row, we multiply the probability of a world state with an individual’s utility of outcome for that world state. As we see above, every individual’s EU is higher without a vaccination under policy X. The EAP therefore dictates that we should choose policy X, since the EU is higher for every person. Importantly, this measure of utility is assumed to be derived from the three rationality axioms of the von Neumann-Morgenstern utility theorem (Fleurbaey and Voorhoeve 2013: 115). The theorem is the basis of EUT and prevails in orthodox decision theory and economics (Stefansson and Bradley 2017: 2). It allows us to measure and calculate EU as above.

Should this measure of EU, which is central to the EAP, really be a decisive criterion in decisions concerning risk? Both policies yield relatively desirable outcomes. The EAP dictates that we choose X, precisely because it yields a higher EU. But would I forego an outcome of 0.85 and risk losing everything for a marginal gain in EU (to 0.8514)? Intuitively, I do not think I would. I would feel disinclined to take this, admittedly small, risk. However, as a utility maximising agent, rationality dictates that I choose policy X. So, what accounts for the EAP’s misalignment with the intuition that I should choose Y?

The problem lies with EUT, which is used to determine the ex ante superior alternative. It does not independently take into account the attitude I hold towards risk taking. In EUT, my risk attitude is derived from the choices I make in certain gambles, not the other way around. I am said to be risk averse with respect to a certain good if I prefer to take the certain outcome of quantity q over a gamble which may yield the expected outcome of q (Stefansson & Bradley 2017: 1). This is only the case if I get diminishing marginal utility from the good. In EUT, risk aversion is a property of an agent’s utility curve of a good and arises as a consequence of diminishing marginal returns to the good in question – no independent attitudes towards risk taking are taken into account. The guaranteed advantage of this approach is that we can easily and elegantly calculate the EU with only two inputs, as we did in Table 1. However, it does not capture our attitudes towards risk in a manner that is useful for decision making purposes. Hansson (1988) expresses this problem as one of conjoint measurement. Our measurement of utility conjointly measures risk attitudes as well as the desirability of a certain outcome. This is where the above dissonance between intuition and
rationality arises. Of course, under certainty, I prefer an outcome of 0.8514 over an outcome of 0.85. But at the same time I may not be willing to risk a certain outcome of 0.85 for a slightly higher expected utility carrying with it a small chance of losing everything. Desirability of an outcome is one thing – my attitude towards risk taking is another. This distinction is not accounted for in EUT.

One of the most significant examples of the shortcomings of this measure of utility and risk is the Allais (1953) paradox. In this paradox we compare the choices of individuals in two experiments, which in turn consist of two gambles with monetary payoffs (see Table 2). The expected payoffs of these gambles can be seen in the last row of the table. When presented with the choices between gambles 1A or 1B and gambles 2A or 2B, EU maximising agents are supposed to choose either 1B over 1A and 2B over 2A, or 1A over 1B and 2A over 2B, depending on their utility curves. Either of the two choice patterns would be valid under EUT. Agents choosing the latter are simply less risk averse with respect to losing money than agents choosing the former. Very often, however, agents choose 1A over 1B but choose 2B over 2A (Allais 1953). This choice pattern is not compatible with EUT. If agents are more risk averse and choose 1A over 1B, they must also choose 2A over 2B, since the chance of winning is higher in both gambles, and vice versa. This result is significant in the sense that the choice pattern indicates a systematic violation of EU maximization.

<table>
<thead>
<tr>
<th>Experiment 1</th>
<th>Experiment 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gamble 1A</td>
<td>Gamble 2A</td>
</tr>
<tr>
<td>Winnings</td>
<td>Winnings</td>
</tr>
<tr>
<td>$1 million</td>
<td>Nothing</td>
</tr>
<tr>
<td>89%</td>
<td>89%</td>
</tr>
<tr>
<td>$1 million</td>
<td>Nothing</td>
</tr>
<tr>
<td>11%</td>
<td>1%</td>
</tr>
<tr>
<td>$5 million</td>
<td>$1 million</td>
</tr>
<tr>
<td>10%</td>
<td>11%</td>
</tr>
<tr>
<td>$1 million</td>
<td>$5 million</td>
</tr>
<tr>
<td>$1.39 million</td>
<td>$110k</td>
</tr>
<tr>
<td>$110k</td>
<td>$500k</td>
</tr>
</tbody>
</table>

Table 2. The two gambles of the Allais paradox.

When agents demonstrate this choice pattern, they effectively violate the independence axiom of EUT, which is one of the three fundamental rationality axioms we need to measure and calculate EU. This axiom states that if two gambles are mixed with an irrelevant third one, then agents should still make the same choice as before. Formally put, for all gambles R, P and Q, if R ≥ P, then RxQ ≥ PxQ, where RxQ means gamble R is mixed with gamble Q. In Table 2 above, a violation of this axiom can clearly be seen. Comparing experiments 1 and 2, we observe that 89% of the gambles gave the same outcome respectively (row shaded in grey), making these parts of the gambles irrelevant. If we disregard the row shaded in grey, however, all the gambles are identical, which should give us a consistent outcome in both experiments. But, as we know, once we mix in the irrelevant part, the choice pattern is often inconsistent, which violates the independence axiom.

It is essential to my argument to explain why this violation matters. The violation of the independence axiom highlights risk attitude as a distinct factor in decision making. The choice pattern is only inconsistent due to a significant change in risk, not due to a significant change in the expected outcome. The significant difference between Experiments 1 and 2 are not the outcomes, B is dominating in both. The significant change lies in the probabilities. Consequently, it seems that having risk preferences as distinct psychological attitude should not be seen as irrational or wrong. If this is true, then the EAP as defined by Fleurbaey and Voorhoeve (2013) and defended by Frick (2013) is in trouble. The reason we employ the EAP is because it can be justified from a contractualist standpoint. Policy decisions under the EAP are seen as correct because they are supposedly justifiable to each person. And yet – they might not be. Decisions made under the EAP are not
necessarily justifiable to each person because people may hold psychologically distinct risk attitudes which are not represented by the EU. Individuals might, therefore, reasonably reject policy X in Table 1, despite it yielding a higher EU.

Though the issue may seem technical, it is important for policymakers not to ignore it. The Allais experiment has been replicated many times, including, relatively recently, using health outcomes. Adam Oliver (2003) found a pattern violating the independence axiom for 14 out of 38 participants of his study. The experiment used “living x years in full health” rather than money as payoffs, but maintained the same chance profile (Oliver 2003: 43). Oliver explains that the pattern is the result of different cognitive processes and argues that “we cannot conclude that these cognitive processes are inconsistent, irrational or wrong. They are merely different” (Oliver 2003: 45). If this is true, using EU as a basis for decision making may not be justifiable to each person. Plausibly, a significant part of the group affected by the policy would be subjected to risks they do not want to take. The example in Table 1 is one instance where this might be the case.

I have argued that the EAP should be used with caution as a decisive criterion for decisions involving risk. The ex ante measurement that is employed, EU, does not capture all the relevant factors of individual decision making – particularly attitudes to risk. Experiments in Allais-type contexts show this to be the case. If it is true that there are distinct psychological attitudes towards risk, decision makers need to take them into account, potentially through theories such as Risk-Weighted Expected Utility Theory (Buchak 201). Otherwise, policy decisions may not be justifiable to each person, as individuals could be subjected to risks they would not take. If policies are not justifiable to each person, then the EAP cannot be justified from a contractualist standpoint of moral rightness, which is a principal basis of its appeal in the first place.
References


Assessing the Justice of Laws: A Critique of King’s Criteria

Stefan Nielsen*

* Philosophy, Logic and Scientific Method, London School of Economics, Houghton Street, London, WC2A 2AE, UK. Email: stpn21@gmail.com

Abstract

This paper argues that Martin Luther King Jr.’s (1992) second and third criteria for assessing the justice of a law are severely limited in their usefulness. I argue that the second criterion is too often inconclusive, because it requires conjecture on the preferences of groups not affected by the law. This process of conjecture can generate inconsistent answers for laws regarding economic equity. I argue that the third criterion is overly broad, since it essentially requires all groups to have real power to affect decision-making. This almost never happens, leading to the absurd conclusion that almost all laws are unjust.

Keywords: Martin Luther King; Justice; John Rawls; Law

1. Introduction

This paper evaluates Martin Luther King Jr.’s (1992) second and third criteria for assessing the justice of laws. I argue that the second criterion is useful to a limited extent, but proves indeterminate for many laws. I also argue that the third criterion deems an undesirably large set of laws unjust. I begin by examining the second criterion. I argue that it must be understood substantively if it is to be applicable in a broad range of cases. Under a substantive reading, however, it follows that laws are rarely, if ever, universally binding. In these instances, a formulation of the preferences of groups not bound by a given law is required to determine if that law is just. I demonstrate that this process of formulating group preferences is inconclusive when questions of economic equity are concerned. Therefore the second criterion is typically inconclusive. Next, I consider the third criterion and argue that we should not place excessive emphasis on the appendage of universal suffrage, but understand the criterion as a requirement that all groups have real power in affecting decisions. I argue that this interpretation leads us to conclude that almost all laws are unjust. This conclusion renders the third criterion overly broad and therefore not useful. I conclude that the second and third criterion are limited in their usefulness for assessing the justice of a law.
King’s first criterion for the justice of a law is that it complies with the law of God (1992: 3). I set King’s first criterion aside for reasons of scope, in order to deal more thoroughly with the second and third criteria. The second criterion states that a law cannot be enacted by a majority on a minority, where that law binds the minority and not the majority and where the majority would not will that law on themselves (1992: 3). The problem with this criterion stems from how we conceive of a law as binding and how we conjecture what a majority would will. The third criterion states that a law cannot be inflicted upon a minority where that minority had no part in enacting or creating that law due to being denied the ability to vote (1992: 3). Essentially, this criterion requires that all groups have a role in the decision making process through their ability to vote. The difficulty with this criterion stems from how we conceive of a group partaking in the decision making process. I assume that a useful criterion is not one that deems all laws, or an extremely large subset of laws, unjust. I acknowledge that one could equally bite the bullet, so to speak, and accept the conclusion that unjust law-making is the norm almost everywhere, when that is the conclusion that follows from applying King’s criteria. For the purposes of this paper, I simply stipulate that such a conclusion renders the criterion overly broad and not useful.

2. Discussion

Starting with King’s second criterion, one first needs to consider who is bound by a given law. Furthermore, one needs to consider, in cases where some are bound by a law and others not, what those not bound would will if they were bound by that law. I refer to this as the role-reversal requirement. That is, in cases where a law binds one group and not another, there is no injustice provided the unbound group would will themselves to be bound by that law. Hence, were roles reversed, the unbound group would consent to that law. As for the first consideration, there are two ways to formulate a law’s scope. Either we interpret the law formally or substantively. I argue that the formal approach is too weak and does not preclude the possibility of targeted injustice. Laws can be formulated to be universally applicable but, in practice, target groups. Waldron (2016) points out that “even a norm like ‘A person who is of African American descent must sit in the back of any public bus they ride on’ applies, universalizably, to everyone. A formal requirement of generality does not guarantee justice”. The issue is that laws can have a conditional form where it will just so happen that only some groups will meet the antecedent condition, but the law is still universally applicable. The law “if you are African American, you must sit at the back of bus” applies to everyone, but in reality only African Americans must endure the law. Therefore, I discard this weak, formal reading of King’s second criterion and consider the substantive reading.

Under a substantive reading, we consider who the de facto target of a law is rather than simply de jure. That is, we go beyond what is formally stated in the law and consider how that law’s implementation will affect people. With a substantive reading, we avoid obvious injustices such as Waldron’s example. However, the issue is that the de facto target of a law is rarely the whole population. Almost every law will have some implicit conditionality to it that will lead to that law affecting groups unevenly. Inheritance tax affects disproportionately the wealthy whereas VAT regresses more poor people more. A tax hike on gasoline and diesel will disproportionately affect consumers of these fuels, which, as the Gilet Jaunes protests in France in 2018 and 2019 have exposed, is predominantly poorer, rural populations and long-distance truck drivers. The point is that differentials across people - be that in terms of income, geography, culture or some other factor - inevitably lead to laws affecting people differently. If it is the case that laws, by and large, de facto bind some and not others, then by implication of the second criterion, almost all laws are unjust. This then means we need to defer to the role-reversal requirement in each of these cases if we wish to
avoid the conclusion that practically all laws are unjust. I argue, however, that this is not a successful system for determining justice, since we are rarely guaranteed a unique answer to the role-reversal requirement. Hence, in most cases the justice of a law is not discoverable by King’s second criterion.

To meet the role-reversal requirement, we need to formulate the hypothetical preferences of groups. There are several approaches to discerning what a group or person hypothetically desires, however I will consider just one approach - the Rawlsian (1971) approach. The Rawlsian approach places us behind a veil of ignorance and asks if we would will ourselves to be bound by that law without knowing if, on entering the world, we fall into the affected or the unaffected group. This approach delivers some desirable outcomes. We would not will that African Americans must sit at the back of a bus, since we would not know if when entering into the world we would be African American and so would not want to constrain our own rights. For cases such as this, the Rawlsian approach yields desirable answers to the role-reversal requirement. But in cases with distributive consequences, reaching a unique answer from this position may not be possible. Sen (2009: 57) argues that multiple, inconsistent answers may emerge. The outcome will necessarily be contingent on "utility, economic equity and distributional fairness, and the entitlement to the fruits of one's unaided efforts". Continuing the example of inheritance tax, arguably the best way to treat people impartially and unbiasedly would be to protect the property they accrue via their efforts and not tax wealth on the basis of that person dying. Alternatively, behind a veil of ignorance we might argue that treating people impartially means trying to correct for arbitrary benefits accruing to individuals for which they cannot claim credit, such as limiting arbitrary economic advantages that exist by virtue of lineage. Both answers aim towards impartiality but start from different theories of economic equity. Since we are not guaranteed a unique answer when considering the role-reversal requirement, it then follows that the justice of a law is rarely discoverable. Under King’s second criterion, we constantly need to defer to the role-reversal requirement, since almost all laws de facto bind some and not others, and this system of formulating preferences of unaffected groups contains no guarantee of a conclusive result. Therefore, King’s second criterion for justice does not return an answer for an extremely broad class of laws.

The third criterion relates to minority participation in the law-making process. There is some ambiguity in how to understand the appendage of unhindered voting rights as a necessary or sufficient condition to satisfying this criterion. I consider three interpretations: firstly, I take universal suffrage as a necessary condition. Secondly, I take it as a sufficient condition. Thirdly, I de-emphasise this appendage and consider a more general reading of the third criterion in which it is read as a requirement that all groups hold real power to affect decisions. The immediate consequence of understanding universal suffrage as a necessary condition for just law making is that any non-democracy is ruled out by default. This means that every law formulated in Singapore, for example, is unjust. Although King’s first principle of justice – that a law agree with the law of God – is not the focus of this paper, I will use it here to show how this interpretation can generate a contradiction for King. Given King’s assertion that “one has the moral responsibility to disobey unjust laws” (1992: 3), one can conclude that we are morally obliged, for example, to refuse to testify in a Singaporean court to convict a murderer who we knew to be guilty. This contradicts King’s first principle, which would almost certainly require cooperation in the case of convicting murderers. Therefore, I argue we should discard the necessary interpretation due to such problems generated with regard to non-democracies.

Understanding universal suffrage as a sufficient condition is also problematic. Universal suffrage is ultimately a formal institution. It tells us nothing of the real power held by different groups in affecting or creating laws. Indigenous populations throughout the Americas can
vote yet cannot be said to hold any significant, real influence in the law-making process. There also exist hegemonic democracies, Russia for example, where all people can vote but the governing party engages in voter suppression, strategic bribes, and other tactics to ensure they retain power. Injustices are prevalent against indigenous populations in functioning democracies and regular citizens in hegemonic democracies. Yet under a sufficient reading of the third criterion we need to defer to some other criterion to argue that these laws are unjust.

One may argue that perhaps we are improperly focussing on this universal suffrage appendage. We can redeem King’s criterion by noting that its motivation is inclusion of those affected by laws in the decision-making process and in the power to enact and change laws. I therefore employ a more substantive reading. There are some benefits to this interpretation. Firstly, we escape the conclusion that all laws enacted in non-democracies are unjust, as citizens in non-democracies may have alternate ways of affecting change in their governments. A vote is not the only way to influence change. Secondly, we avoid brushing over hegemonic democracies and marginalised, but technically enfranchised groups. Despite the advantages of this reading, we run into similar issues as we did with the second criterion. Just when, exactly, do all citizens have real, significant power to create and change laws? This can only really be said to occur under very direct forms of local democracy. Indeed, this is most likely true to King’s intention. Practically however, no country operates this way, and so we end up concluding that almost every law is unjust everywhere because it is almost impossible to have all citizens holding legitimate sway in the policy arena.

3. Conclusion

I have argued that King’s second and third criteria are limited in their usefulness for assessing the justice of a law. I have shown that King’s second criterion must be understood substantively, but that a substantive reading necessitates constant deferment to the role-reversal requirement. The role-reversal requirement produces no clear answer as to what groups would or would not will when economic equity is involved. It then follows that the second criterion is inconclusive for a broad set of laws. I acknowledge, however, that this argument examined only the Rawlsian method for discerning the preferences of unaffected groups. One could also observe actual behaviour in order to forecast preferences of unaffected groups or various other empirical approaches. This is a suggested avenue for further research. I have also argued that King’s third criterion is flawed, because it is overly demanding. A requirement that all individuals have real power to affect decision-making would require highly localised direct democracy, which exists virtually nowhere in the world.
References


Consent and the Resource Curse

Andrea Lidia Villalba Cuesta*

* Philosophy, Logic and Scientific Method, London School of Economics, Houghton Street, London, WC2A 2AE, UK. Email: a.villalba-cuesta@lse.ac.uk

Abstract

The resource curse is paradoxical in the sense that an abundance of resource wealth typically hinders economic development for citizens of authoritarian regimes. Leif Wenar’s (2008, 2016) proposal bans the purchase of natural resources from such governments on the grounds that these resources are sold without citizens’ consent and thus violate property rights. In contrast, this paper argues that curse-ridden populations may have no reason to object to the trade in natural resources. They exhibit ‘low-order’ consent due to concerns that banning trade will make them worse off. I conclude that this undermines the case for Wenar’s proposal.

Keywords: Leif Wenar; Resource Curse; Consent

The resource curse refers to the failure of citizens to fully reap the benefits of their country’s natural resource wealth, as the wealth is controlled by the autocratic elite. Often, this adversely affects governance and development outcomes (Ross 2015: 239). Leif Wenar (2008, 2016) claims that by recognizing the property rights of autocratic regimes, other state governments become complicit in the autocratic governments’ theft of natural resources from their populations. Foreign states are breaking the “first rule of capitalism in transporting stolen goods” on a huge scale (Wenar 2008: 2). Wenar’s solution is to ban the import of resources sold by repressive governments which do not meet the minimal criteria necessary to authorize the sale (i.e. respect for basic civil and political liberties). I shall argue against Wenar’s positions, concluding that curse-ridden populations may have no reason to object to the trade of their natural resources. I shall argue that Wenar’s proposal cannot be justified because his conditions for authorisation do not track consent.

The resource curse is conditional (Caselli and Tesei 2016). In countries with inclusive institutions, natural resources are a blessing. Incentives for “participation by the great mass of people in economic activities” (Acemoglu and Robinson 2012: 144) are created by securing property rights, law and order, and markets. By contrast, extractive institutions concentrate power in the hands of a small self-serving elite (Robinson 2013: 2). Under these conditions, natural resource windfalls generate three overlapping curses (Wenar 2008: 3):

1. Increased likelihood and depth of authoritarianism;
2. Increased risk of civil conflict;
3. Reduced economic growth.

Here, I focus on so-called point source resources such as oil. Isham et al. (2005) argues that countries with such resources are doubly disadvantaged by virtue of exhibiting concentrated and capturable revenue patterns. These resources are often extractable without the population’s cooperation and flow directly to the government, increasing its repressive capacity. Accountability is thwarted as concealed income streams reduce the ability need for intrusive bureaucracies to raise tax revenue (Fearon and Laitin 2003: 81). Since natural resources yield transferable rents, opportunities for controlling primary commodity exports, referred to as “predation” (Collier and Hoeffler 2000), cause conflict. The prize value of capturing a state increases as greedy rebels sell future exploitation rights called “booty futures” (Ross 2002). Groups that gain income and funding directly from conflict (e.g. military personnel paid for security) may also prefer to fight than to win: they exhibit a “domestic conflict premium” (Humphreys 2005: 516). Poor governance leads to low investment in human development. Roberts and Robinson (2013: 73) classify entrepreneurs in two groups: grabbers maximise monopoly rents whilst producers use resources to generate additional wealth. Exclusive institutions produce grabber-heavy economies.

Wenar (2008: 3) claims that the resource curse is a “symptom of the violation of property rights.” A valid sale requires that a vendor have the right to sell: he must either be the owner or have the owner’s authorisation. Contracts facilitating theft or trade in stolen goods are void unless sellers or buyers acted in good faith, i.e. lacked awareness that the seller did not hold a clean title. Wenar extends legal rules governing private party conduct to the state. The principle of national ownership grounds his proposal: the people occupying a given political territory ‘own’, in the legal and moral sense, the resources within it. This principle conflicts with the “international resource privilege” principle (Pogge 2002: 22) which authorises rulers, regardless of how they attained power, to sell their domestic resources on the global market. Accordingly, the legal right to transfer property is acquired disproportionately by those who are lucky and strong; “might makes right” (Wenar 2008: 13). An inevitable by-product is that unaccountable governments are permitted to sell a territory’s natural resources without their citizens’ consent. By enforcing this right in courts, importing countries trade in stolen goods. The revenues generated provide means and incentives for oppression and conflict.

Wenar’s solution is two-fold. First, his proposed Clean Trade Act bans the import of natural resources that do not meet the minimal criteria necessary to authorise the sale. The minimal criteria are that the exporting country’s citizens must (Wenar 2008: 20):

i. Be able to find out about the sales.
ii. Be able to stop the sales without incurring severe costs.
iii. Not be subjected to extreme manipulation by the seller.

Wenar claims that these criteria are not satisfied by countries with a rating of 7 (the worst) on Freedom House civil and political rights rankings (Wenar 2008: 25). Underlying this claim is the issue of consent. Given that trade in natural resources may not be in citizens’ best interest, or of sufficient value to them, their preferences should be taken as inputs for collective decision making. The second part of Wenar’s solution is a Clean Hands Trust, which imposes equivalent-value duties on trade partners who buy stolen resources. Revenues are assigned to a trust held for the autocratic country’s population until the minimal conditions are met.

Objections typically centre on Wenar’s minimalism. His explicit focus on “depriving dictatorships” (Nili 2017: 317), violent regimes that rob their populations of the majority of all
resource rents, ignores that compelling moral challenges also affect dictatorships that are not the worst of the worst (with a rating less than 7). It has been argued that his criteria therefore demands too little. However, by acknowledging the limited scope of Wenar’s proposal, I arrive at the opposite conclusion: his minimal criteria demand too much. I identify two main justifications for Wenar’s limited scope. First, a “feasible” reform should be based on the most minimal premises possible in order to get a foot in the door (Wenar 2008: 17). This sets a legal precedent that can progressively be applied to more contested cases. Second, given severe hardship, it is difficult to see how reform could impoverish inhabitants of depriving dictatorships any further. Hence, Wenar assumes that:

(A) Autocrats cannot obtain actual consent from their people for continuing tainted trade.

On this basis, trade in natural resources violates property rights. The rest of my discussion concerns the validity of assumption (A). My point is clearest when applied to what the literature calls *distributive dictatorships* (Nili 2011: 105). A population can either ask the regime to sell (‘explicit’ consent) or acquiesce by remaining silent (‘tacit’ consent). Significantly, the ability to prevent the sale of natural resources like oil is valuable only if there is a reason to object to it. Schaber (2011: 188) argues that, unlike historical monuments, oil is not something you could wish to preserve in the same way. It becomes interesting only once it is sold. Distributive dictatorships often divide resource rents amongst their population, rarely employ the naked use of force and guarantee a baseline standard of living (Nili 2017: 317). In such countries, citizens can maintain that property rights violations occur but still credibly consent to tainted trade (hereinafter referred to as ‘low-order’ consent) because they fear that reform could make them worse off (Allen and Lektzian 2013: 122). Indeed, bans are likely to produce relative deprivation in the short-run (Nili 2017: 315). These points make Wenar’s assumption less plausible.

I extend this reasoning to *depriving dictatorships*. Under President Obiang, 77% of Equatorial Guinea’s population lives below the poverty line despite inhabiting Sub-Saharan Africa’s fourth largest crude-oil producing country (Schaber 2011: 185). Arguably, Equatorial Guineans may also have no reason to want their oil to remain unsold. Nili (2014: 7) constructs a hypothetical scenario in which citizens of curse-ridden nations face a choice between inflows or outflows of resource-derived wealth. Whereas the former leaves political conditions unchanged, the latter involves the dictator going into exile. Provided that exile leads to “a much more accountable, transparent and responsive government” (Nili 2014: 7), long-term poverty rates should fall significantly, making it the favoured option. In reality, it is difficult to convince citizens that the key task is initially to take resource wealth out. Firstly, conforming to Kuran’s (1991: 2) notion of the “predictable unpredictability of dictatorial breakdown,” uncertainty about how trade reform will affect a dictatorship’s internal power structures abounds. Secondly, it takes an estimated 59 years for a failed state to turn around and sustain itself (Chauvet and Collier 2005: 1). This means that revenues could sit in the Clean Hands Trust for much longer than an average citizen’s lifetime. Therefore, Nili’s choice is not so clear-cut; citizens are in fact choosing between maintaining the status quo or a transitional period marked by high uncertainty. Given this economic risk, citizens in depriving dictatorships may similarly exhibit ‘low-order’ consent.

It follows that Wenar’s argument suffers a fatal blow because his conditions for authorisation does not track consent; his minimal criteria are overly-demanding. This is worrisome given that his key premise (i.e. trade in natural resources violates property rights in depriving dictatorships) rests on (A). I use the following example to fully illustrate this point. Imagine that, on a whim, a benevolent dictator decides to launch a school-building programme. He engages in this secretly; without asking for consent or allowing the public to ever find out. Even if information were leaked, no one could stop the programme without incurring severe
costs. If we conform to Wenar’s minimal criteria, our Benevolent Dictator would not be authorised to build schools, despite our hunch that his country’s citizens would never object to him doing so.

A rebuttal could be that this is not the notion of property that Wenar deems worthy of protection. Ownership is the state of exclusive control over property. It is self-propagating to the extent that the owner also decides “what to do with any benefits that accrue” from it (Nili 2011: 106). Wenar implies that control, not only benefits, constitute legitimate ownership: “the curse results from a defect in the rules that allocate control over these resources” (Wenar 2008: 8). The examples above suggested that rather than focus exclusively on rights to sell (‘control’), one should be equally concerned with whether citizens’ interests are being served by the sale of natural resources (‘benefits’). Problematically, this can blur the distinction between promoting interests and respecting property rights. However, this rebuttal bolsters my argument against Wenar. Recall that the ability to block sales only becomes relevant once there is a reason to object to it. In making the relevance of his minimal criteria contingent on outright objection, Wenar unintentionally places more value on benefits than control.

Arguably, not every instance of actual consent is valid. According to Nili (2017: 320), consent might stem from duress. This should be understood as either having no choice or as having no acceptable choice (Wood 2014: 21). Ensuring that consent is valid becomes more important as the dangers to which it exposes the owner increase. In this particular case, the danger is substantial: continued subordination to the dictatorial regime. Whilst I agree with this point, Wenar does not sufficiently account for duress in criterion (iii). He indicates that to be able to authorise a transaction, an owner cannot be “hypnotised, brainwashed or subject to extraordinary psychological manipulation”, as in North Korea (Wenar 2008: 21). Nonetheless, this misidentifies the wrong committed: manipulation is the steering of others’ choices in morally problematic ways. Criterion (iii) should be weakened to incorporate the importance of citizens’ bargaining power vis-à-vis the state, including the ability to exit. In short, given that citizens have no acceptable alternative, they consent to tainted trade because reform could make them worse off. Once again, Wenar’s minimal criteria demand too much.

To conclude, Wenar’s proposal fails because his minimal criteria for the authorisation of sales do not track consent. I have argued that Wenar’s thesis rests on the mistaken assumption that autocrats cannot obtain consent from their people when engaging in resource trade. I showed that this was problematic insofar as the ability to prevent the sale of natural resources is only valuable if there is a reason to object to it. In distributive dictatorships this reason may not exist. Concerns that trade reform could make citizens worse off, particularly given the uncertainty of dictatorial breakdown and the pay-out timeframe of the Clean Hands Trust, explain why there may be no reason to object. I then examined the counterargument that granting the validity of low-order consent masks the distinction between promoting interests and respecting property rights. However, I concluded that by making the relevance of his minimal criteria contingent on a capacity for outright objection, Wenar is responsible for placing a higher value on ownership of benefits than control. In doing so, he ignores the role played by duress, which differs markedly from the manipulation mentioned in his third criterion.
References


Consent and the Resource Curse


Abstract

Leif Wenar (2008, 2016) argues that by purchasing goods produced from the natural resources of repressive regimes, Western consumers share responsibility for repression in those autocratic countries. To address this problem, he proposes a Clean Hands Act, which would restrict the purchase of natural resources from these states, and a Clean Hands Trust, which would tax third-party countries that buy natural resources from those regimes. This paper demonstrates how practical limitations, such as sanction avoidance and secret trade, make this proposal untenable. The paper further argues that Wenar’s framing of the problem as a ‘property rights violation’ is insufficient and it is more appropriate to frame the problem as a lack of legitimacy.

Keywords: Sanctions; Legitimacy; Property Rights; Consumerism; Autocracy
they use to oppress their populations, rig elections and to personally profit. Wenar’s chief contention is that by upholding demand for those goods, consumers, especially in the developed world, are morally responsible for the consequences of the resource curse. To alleviate the problem, Wenar (2008) proposes a Clean Hands Act, making it “illegal the purchase of resources from a disqualified country” (Wenar 2016: 284) where disqualified states are resource-cursed states. Wenar’s ban is not extended to other resources, applying only to natural resources. Wenar’s sentiments are shared by John Bolton, the current National Security Advisor of the U.S. who has urged people not to trade with Venezuela in gold and oil as these are “being stolen from the Venezuelan people” (Bolton 2019).

The Clean Hands Act is problematic. Wenar (2008: 30) notes that previous sanctions have not been effective, yet he introduces a proposal that is similar to sanctions in terms of the mechanism which is employed to produce an effect. While Wenar’s proposal is not an embargo, it is an example of a targeted sanction which focusses on specific segments of society and is meant to reduce collateral damage to other parts of society and to third-party countries (Hufbauer and Oegg 2000). Wenar’s proposal simply limits the ability to trade natural resources with specified countries, as usual sanctions do. As a result, the proposal is susceptible to the same shortcomings as sanctions.

One of the main concerns is the role which third-party countries can play in helping to alleviate the cost of sanctions, known as sanction-busting techniques, which allows sanctioned states to circumvent sanctions to some extent (Early 2015). One example of sanction busting is the use of an intermediary to conduct trade. For instance, after the Nobel Prize was given to a Chinese dissident, China imposed sanctions on Norwegian salmon. Private actors avoided these sanctions through rerouting, falsifying country-of-origin certification and smuggling (Chen and Garcia 2016). Another example is the sanctions imposed by the West on Iran. After Iran was disconnected from the EU-based SWIFT (Society for Worldwide Interbank Financial Telecommunication) financial transaction network, there was a surge of gold transportation from Turkey to the United Arab Emirates to Iran and vice versa (Early 2015). This way, Iran could still trade in gold without relying on the SWIFT system. In fact, imposing sanctions incentivises third-party countries to participate in sanction busting as they tend to profit, even when they are allies of the sanction imposing state. When the US imposed sanctions on South Africa, trade of the latter with West Germany, Great Britain, Japan and Italy increased (Early 2015). Thus, sanctions do affect the welfare of the sanctioned entity, in this case the autocrat, but the possibility of avoiding sanctions through diverting trade means that this impact is at best partial.

Wenar (2008) defends his proposal against this charge by presenting a case for the Clean Hands Trust, which he claims makes approach more efficient than traditional sanctions by correctly aligning incentives. The idea behind the Trust is that countries would tax third-party countries up to an amount equivalent to the illegal trade that the third-party country does with the sanctioned country. For instance, if China buys $2 billion worth oil from Equatorial Guinea then Chinese imports will be sanctioned by the amount of $2 billion. However, this defence is flawed for three reasons. Firstly, similar sanctions against other countries have harmed people in those countries more than their autocratic leaders (Van Der Ploeg 2017). Wenar could respond that this would be alleviated by the Trust as the money would be given back to those people after the country satisfies minimum requirements of civil liberties and political rights (Wenar 2016). In essence, the burden on people will be temporary as the money will be given back to them. The problem is that it is possible for the dictator to manipulate ratings to become marginally freer and then, after the money is received, slide back into a dictatorship.
Secondly, a greater concern is that in such a system intermediaries and other countries will be incentivised to hide their trade with sanctioned states. Historically this has been the case with many sanctions. For example, after Napoleon forced Russia to impose an embargo on Britain, Russians kept trading with the British, but the official government policy was that the government was unaware of the trade (Roberts 2016). A contemporary example is North Korea. China is said to be North Korea’s biggest trade partner, but Russia also conducts trade with this sanctioned country (Lukin and Zakharova 2018). This is conducted in a secretive manner due to the North Korean regime’s reputation. The willingness to support sanctioned states through trade is partly because trade between autocratic governments is motivated not only by material gains but also by the desire to maintain similar regimes against the pressures of the liberal West (Blackwill and Harris 2017). In an environment of secrecy it will be extremely challenging for countries to decide how much tax to levy on other countries, which makes the proposal unfeasible.

Thirdly, Wenar’s initial argument assumes institutions that work and that external incentives can drive countries towards less autocratic politics. This lies at the very core of Wenar’s proposal – to some extent the idea behind the Clean Hands Trust is to incentivise citizens in autocratic countries to fight for accountability from their governments and for governments to relax oppressive actions to allow trade. However, this may not be the case in low state-capacity countries that have weak legal and fiscal infrastructure and cannot properly ensure rule of law or collect taxes (Besley and Persson 2013). By virtue of being poor, having high political instability and high revenue from resources, those countries have a vicious circle: low capacity and low social cohesiveness ensure low tax revenue which in turn incentivises low investment in state capacities such as the ability to uphold the rule of law and low levels of development. Instead of investing in state institutions, governments will simply try to stay in power, often by authoritarian means, collecting as much money as possible through trading illegally acquired resources.

If one implements Wenar’s proposal, then one would expect lower revenue from resources. This decrease, however, will only be partial as other autocratic governments might step in to buy resources or be intermediaries. This might cut down the amount of repression that a population experiences as the autocrat will have less money to support an army, but it is not at all clear that this decrease would be sufficient to ensure that the government would try to accommodate civic and political rights. If our desire to alleviate resource curse is from current guilt - that we as consumers care only about our current continued contribution to the curse - then indeed cutting revenue from resources would help. However, as Besley and Persson’s (2013) model demonstrates, this would not help those countries alleviate the damage that has already been done and would be insufficient if our aim is to help populations to be free from resource curse and repression.

Finally, a philosophical challenge can be presented as well. Wenar (2008: 13) states that “might makes right” cannot be a good international principle for property rights, especially when it comes to natural resources, as it violates the commonly accepted principle that natural resources belong to the people stated in the Article 1 of the International Covenant on Civil and Political Rights (United Nations General Assembly 1966). Thus, he frames the problem as a “violation of property rights” (Wenar 2016: 3) since the people have not consented to the use of their resources. But it is not clear why this should not be presented as an issue of state legitimacy instead. Suppose Robert Mugabe, in his popular years, created a law approved by the public that made the country’s resources private, and acquired most of the shares in the company owning these resources. Suppose he went on to become a dictator after that. In such circumstances, he would have acquired resources legally and hence it would not count as trading stolen goods. Nevertheless, it is still the case that by buying resources from him, Western consumers would perpetuate the resource
curse that Zimbabwe has (Doro and Kufakurinani 2018). While property violation does not apply in this case, one could say that some form of Wenar’s argument of the harm inflicted by consumers of the developed world should still be applied to Zimbabwe.

One may appeal to the status quo legal principle of state sovereignty. This norm protects morally important interests such as the right for self-determination and formal equality of states in the international arena. One consequence of this norm is that property regimes should be determined domestically (Pavel 2015). But this would hardly be an acceptable response within Wenar’s framework. The status quo is not a valid justification as long as democratic governments and government by the people is upheld by many people across societies, and has been enshrined in international agreements such as United Nations Resolution 64/155 (United Nations General Assembly 2009), the Universal Declaration of Human Rights Article 29 (United Nations 1948) and the African Union’s Constitutive Act (African Union 2000) amongst others. This challenge is particularly relevant in what John Rawls (2003: 106) describes as the case of burdened societies: states that are not expansive or aggressive but “lack the political and cultural traditions, the human capital and know-how”. In those states, the people simply cannot establish a domestic property regime. Arguably, countries like Equatorial Guinea can be described as one of those countries. These are countries where government institutions are unable to provide basic services such as protection of a person’s life, as in many countries in Sub-Saharan Africa. Many of those countries have governments that are not merely corrupt but that also lack democratic legitimacy. Here lies the problem: what should be done if there is no legitimate entity to deal with?

In conclusion, Wenar’s proposal is not sufficient for dealing with the problem which he presented. The Clean Hands Act in being similar to sanctions, suffers from the same deficiencies. The Clean Hands Trust does not alter incentives sufficiently and will be circumvented by other autocratic states. Moreover, resource revenue is only part of the reason why autocrats remain in power and cutting revenue partially will not incentivise them to extend political and civic rights, as is the aim of Wenar’s proposal. Finally, I have also challenged the general framework Wenar uses and argued that the core issue is not a property rights violation but a lack of legitimacy of governments.
References


Bolton, J. 2019. Twitter Post. 30 Jan, 2019, 4.48PM. URL: https://twitter.com/AmbJohnBolton/status/1090652929508417537


Pavel, C. 2015. “Negative duties, the WTO and the harm argument.” Political Studies 63(2): 449-465. URL: https://doi.org/10.1111/1467-9248.12098


Why It Is Unjustifiable to Ban the Purchase of Natural Resources from “Resource-Cursed” Countries

Florence Roughton*

* Philosophy, Logic and Scientific Method, London School of Economics, Houghton Street, London, WC2A 2AE, UK. F.M.Roughton@lse.ac.uk

Abstract

This paper considers Leif Wenar’s (2008) proposal for banning the purchase of ‘stolen goods’ from resource-cursed countries, as he argues governments do not have the consent required to authorise these transactions. The benefits of the proposal are that it disincentivises the sale of stolen goods, incentivises citizens in these countries to overthrow their corrupt governments, and the proposal’s penalties track the stolen goods. However, his proposal could cause significant harm to citizens of resource-cursed countries, and the conditions he deems necessary for citizens to authorise the sale of their resources are not sufficient. In concludes that Wenar’s proposal is unjustifiable.

Keywords: Resource curse; Property rights; Economic sanctions

The ‘resource-curse’ occurs when sale of natural resources strengthens the hand of repressive regimes and thereby enables poor governance and causes lack of broad-based economic growth. In this paper, I shall analyse Wenar’s (2008) proposal, grounded in property rights, for banning the purchase of natural resources from extremely repressive regimes and for levying a tax on the exports of countries who continue to purchase such resources, with the proceeds being held in trust for the citizens of the repressive regime until they acquire minimally decent government. He outlines certain conditions necessary for citizens to consent to their government selling natural resources, which, if not fulfilled, means these goods are effectively stolen. This proposal has some apparent merits: it disincentivises the sale of stolen goods, incentivises people in resource cursed countries to overthrow their corrupt governments, and the proposal’s penalties track the stolen goods. However, I will criticise it with respect to two considerations. Firstly, his proposal could worsen the situation of the people in the resource-cursed countries, as I posit with reference to empirical evidence. Secondly, the conditions under which Wenar deems it acceptable for citizens to authorise the sale of their collectively owned resources are not stringent enough. Thus, I shall conclude that whilst Wenar’s proposal has certain merits, overall, its drawbacks mean it is unjustifiable.
Wenar’s argument is based on the premise that natural resources are collectively owned by the people of a state. He supports this with reference to international human rights law, as well as national laws entrenched in the constitutions of certain countries (Wenar 2008: 9-10). Thus, the property rights of citizens within a country are violated whenever natural resources are sold without their consent. Wenar (2008: 20) asserts the following minimal conditions necessary for consent to be given to these governments to sell the natural resources:

1. Citizens must be able to find out about the sales.
2. They must be able to prevent the sales “without incurring severe costs” (Wenar 2008: 20) i.e. they must be able to voice their dissent without fear of reprimand.
3. They must be able to make decisions of their own volition, without manipulation by the seller.

If these conditions are not fulfilled, the silence or acquiescence of the people does not indicate consent. In his proposal, Wenar (2008) uses the Freedom House ratings, which evaluate the political rights and civil liberties afforded to citizens, to identify countries in which the above conditions do not hold but governments are nevertheless selling natural resources. These ratings were devised by an independent NGO, are based on the Universal Declaration of Human Rights, and are used by the US government, giving them credibility and legitimacy as a standard. Countries with the worst Freedom House rating of 7 in terms of civil liberties have “virtually no freedom” and “[a]n overwhelming and justified fear of repression” (Wenar 2008: 24). Those with the worst rating for political rights find their rights “absent or virtually non-existent” due to the oppressive regime in power (Wenar 2008: 24). Wenar therefore concludes that any country with the worst rating of 7 on either civil liberties or political rights cannot have the above conditions fulfilled. Thus the governments in these countries are selling resources owned by the people without their consent, which violates the people’s property rights and makes the goods ‘stolen’.

Wenar proposes a two part solution to this problem. First he proposes enacting a ‘Clean Trade Act’ which requires countries to stop trading with such extremely oppressive regimes. Second, countries that trade with these authoritarian governments and purchase ‘stolen goods’ from them should face tariffs on their exports to other countries (Wenar 2008). He uses the example of China, Sudan and the US: if China buys Sudanese oil then, if the US follows his framework, it should place import tariffs on all Chinese goods. The tariff revenues will be stored in a ‘Clean Hands Trust’ that will be allowed to fill up to the point that the Trust money is equivalent to what China has paid for Sudanese oil. This trust money is retained by the US until Sudan experiences a regime change such that the above conditions are met, at which point the Trust money shall be given to the Sudanese people. Wenar (2008) defends his proposal on the grounds that it protects the property rights of Sudanese citizens, and will return to them the exact monetary value of the property that was stolen from them by tracking these resources and placing indirect taxes on their purchase. Furthermore, it will encourage the Sudanese people to overthrow their government. They will be highly incentivised to do so as they will receive the trust money as well as being relieved of a corrupt regime. Lastly, the tariff-and-trust mechanism will disincentivise countries like China from trading with Sudan because their exports to the US will face larger tariffs, taking power out of the hands of the Sudanese elite.

However, Wenar’s proposal could worsen the situation of the people in these resource-rich countries. The little income they do receive could be dependent on resource revenue, thus, taking it away by disincentivising the demand for resources could seriously increase dire poverty. For example, in Sudan where (on the most recent estimate available, which is for 2009) almost 50% of people live below the national poverty line (World Bank 2019), falling incomes could lead to (increased) malnutrition and more deaths. Pogge (2005) argues that
we have a negative duty to prevent harm to people, and that citizens of resource-cursed countries should not face the suffering they do at the hands of other countries. According to Pogge "someone is harmed when she is rendered worse off than she was at some earlier time, or than she would have been had some earlier arrangements continued undisturbed" (Pogge 2005: 4). Thus, if the sanctions make the populations of resource-cursed countries worse off than they would have been without the sanctions, then the proposal is unjust through its perpetuation of harm. Wenar (2008: 31) briefly considers this criticism, but dismisses the problem of harm as merely temporary. However, this assumes that the Sudanese people will quickly be able to overthrow their government, which is unlikely given their lack of resources and power. (While, at the time of writing, public pressure led to the fall of the Sudanese dictator Omar al-Bashir, the military junta which has taken over does not appear more democratic.) It becomes even less likely that they are able to do so without receiving whatever little benefit they may have gotten from resource revenue. Collier and Chauvet in fact estimate that failed states like those Wenar describes as scoring 7 on the Freedom House ratings, take, on average, 59 years to turn around and be capable of sustaining themselves (Collier 2007: 71).

Oechslin (2014) provides evidence for this argument, examining the case studies of economic sanctions on Haiti and Iraq. He finds that imposing sanctions on authoritarian regimes in these countries, such as the bans and tariffs proposed by Wenar, causes elites to withhold public goods from citizens, thereby worsening instead of ameliorating the effects of the sanction (Oechslin 2014: 25). He theorises that this is because the elites aim to weaken the citizens’ ability to revolt. Reducing the supply of public goods makes it more costly for citizens to challenge the regime by lowering their incomes, meaning citizens face a bigger loss in utility if they decide to revolt. In Haiti, for example, there is evidence that the government weakened the healthcare system (Oechslin 2014: 27) and destroyed some of the agricultural infrastructure to prevent production and investment (Werleigh 1995: 166-7). Oeschslin (2014) also argues that governments use the sanctions as an excuse for worsening conditions, often to mask their own involvement. This point is supported by interviews with Sudanese citizens after the lifting of sanctions by the US, one of whom said “The government used to say that US sanctions were responsible for all that was wrong in our country… [t]hey can’t hide behind that excuse anymore” (Vergee 2018: 2).

Evidence that conditions worsen under the imposition of sanctions has been found to persist in various other examples. Hufbauer et al (2007) reviewed 57 cases of sanctions that ended prematurely between 1914 and 2000. In 65% of cases sanctions were lifted despite a lack of regime change, which had been at least one of the aims of the boycotts (Hufbauer et al 2007: 80). Oechslin argues this is because the sanctioning bodies came to realise that the sanctions are in fact worsening the situation for the citizens of the country, rather than causing regime change, and thus lifted the sanctions to prevent further damage (Oechslin 2014: 27-8). Therefore, it is not possible to dismiss harms to citizens of resource-cursed countries as merely temporary, since empirical examples demonstrate that sanctions often cause significant harm without even leading to the long term benefit of regime change.

A further criticism is that the conditions Wenar sets out as necessary for resource sales to be permissible are not stringent enough. For here are cases that meet his conditions that involve a failure of consent. These cases should, by his logic, incur embargoes. Condition 2 dictates that citizens must have the power to stop the sales, but Wenar qualifies this by saying a regime must have “some effective mechanisms in place through which it acknowledges that people can dissent to the sales” (Wenar 2008: 21). This is, however, far too minimal. Citizens could dissent to resource sales whilst having no influence whatsoever because, crucially, the cases being examined here are of states with authoritarian governments who do not need the support of their people to stay in power. Rather, they
retain power through control of these resources, allowing them to buy security and buy out opposition, not through a democratic mandate. Governments of resource-cursed countries can and do ignore dissent, and other states often recognise them as legitimate despite awareness of this corruption (Wenar 2015). Thus even if all the conditions are fulfilled, i.e. the Sudanese population could have access to information about the sales without manipulation from the state and voice their dissent, this does not mean they have given their consent. The condition could be strengthened by reference to a stronger example of the prevention of sales, e.g. through a voting system. This criticism is supported by Nili (2011) who argues that Wenar ignores the larger moral complexities that arise from cases that are not “the worst of the worst” (Nili 2011: 8). In the worst cases, it is clear that the conditions are not met and it becomes much easier by Wenar’s framework to condemn the governments. However, more concerning are the countries in which Wenar’s conditions are met but the people of the country are still being exploited. Nili (2011: 9) explores a connection of the “harm to citizens” argument, positing that it could be the case that citizens are aware of the trade and are able to dissent, but do not do so for fear of the consequences and harm of trade being curtailed. Thus, Wenar’s framework still allows for cases in which the government is stealing resources from people, which means it is not sufficient in its aim of preventing governments from selling resources that are stolen due to a lack of consent.

Wenar’s proposal aims to tackle the violation of property rights of the citizens of resource-cursed countries through a combination of bans on the purchase of these resources and a system of taxation on those who violate the ban with proceeds being held in trust for these citizens. His plan is beneficial in that it incentivises the toppling of authoritarian governments, tracks stolen goods and ensures their (monetary equivalent) reimbursement to the people they were stolen from. However, I have argued that protecting citizens’ rights in this way could lead to worse situations for them, both in terms of fewer trickle down benefits from resource sales, as well as potentially increasing hardships imposed by their governments through public spending cuts. Furthermore, his conditions are not strong enough to ensure consent to the sales. Therefore, Wenar’s proposal, although perhaps encouraging from an idealistic perspective due to its grounding in property rights, cannot be justified when implemented in the real world.
References


The Case for a Total Economic Embargo of Failed States

Qiang Hua*

* Philosophy, Logic and Scientific Method, London School of Economics, Houghton Street, London, WC2A 2AE, UK. Email: q.hua1@lse.ac.uk

Abstract

Leif Wenar (2008, 2016) argues that autocratic governments have no property rights over their natural resources and the international community should therefore stop trading natural resources with autocracies. I shall argue that Wenar’s ban should only target failed states and be expanded to cover all economic assets. First, I will outline Wenar’s principles of property rights, which result from ‘popular sovereignty’ over natural resources and the self-interests of consumers. Second, I will analyse Wenar’s proposal for an import ban on natural resources from autocratic states and a taxation plan called the ‘Clean Hands Fund’. I will argue that Wenar’s proposal is unfeasible and inconsistent. Finally, I will outline my proposal for imposing a total economic embargo on failed states as an alternative to Wenar’s position.

Keywords: Autocracies; Oil; Embargo; Terrorism; Conflicts.

Leif Wenar (2008, 2016) argues that autocratic governments have no property rights over their natural resources and the international community should therefore stop trading natural resources with autocracies. I shall argue that Wenar’s ban should only target failed states and be expanded to cover all economic assets. First, I will outline Wenar’s principles of property rights, which result from ‘popular sovereignty’ over natural resources and the self-interests of consumers. Second, I will analyse Wenar’s proposal for an import ban on natural resources from autocratic states and a taxation plan called the ‘Clean Hands Fund’. I will argue that Wenar’s proposal is unfeasible and inconsistent. Finally, I will outline my proposal for imposing a total economic embargo on failed states as an alternative to Wenar’s position.

1. Wenar’s Principles

The first principle underlying Wenar’s position is a principle of property rights in the form of popular sovereignty. His principle states that the resources of a country belong to its people (Wenar 2008: 2). Autocracies grossly violate this principle as their people have no influence over the sale of the country’s resources and seldom enjoy the benefits of the transactions (Wenar 2008: 4). Thus, natural resources sold by autocracies are stolen resources even if
they are sold via a legally sound entity such as an oil company. In the same manner that an individual will be held liable to the original owner of a stolen good if they knowingly purchased that good from a legal store, we cannot legally recognise stolen resource simply by the legitimacy or illegitimacy of the corporations involved (Wenar 2008: 16).

Wenar’s second principle is the principle of self-interest. Wenar argues that any benefits consumers derive from the import of autocratic natural resources will be outweighed by negative 'spillovers' (Wenar 2008: 30). Resource windfalls have empowered autocracies such as Russia and Iran, enabling them to undermine the interests of countries purchasing their resources. In addition, autocratic resource-exporters are extremely unstable (Wenar 2008: 3). Of the enormous wealth that natural resources generate, only a small portion of that wealth is sufficient to fund the security forces necessary for safeguarding resource harvest and transport (Wenar 2008: 4). Thus, the ruling elite has a minimal need to tax the population and provide public services, leaving a large portion of the windfall available for private consumption. This generates cycles of conflicts as the magnitude of the resource wealth incentivises others to seize the resources by force (Wenar 2008: 7). International terrorist groups like ISIS can exploit instability and injustice to gain territories and recruits. The instability generated harms consumer interests, and it is therefore not in the self-interest of consumers to support trade which promotes such instability.

2. Wenar’s Solution

Wenar uses the Freedom House classification to determine whether a regime can “legitimately sell resources from that country” (Wenar 2008: 25). Any regime classified as 'unfree' will be subject to an import ban. These regimes do not meet the minimal standards of accountability as their people have no basic rights concerning the use of their resources, such as the right to information and the right to protest (Wenar 2008: 20). Wenar further proposes to tax goods sold by countries that defy the ban. The tax will extract the exact value of the stolen resources imported by that country. Revenues generated will be used to create a fund known as ‘Clean Hands Trust’ which will be transferred to the resource-exporters if they transition to democracy (Wenar 2008: 28). Wenar believes the fund and the ban would, in conjunction, incentivise the populace to replace their unaccountable governments (Wenar 2008: 29-31).

Wenar’s solution suffers from two problems. First, the ban and taxation plan are unfeasible as they require cooperation from the majority of leading economies. Without cooperation countries defying the ban could simply divert their exports to an alternative market. Such cooperation is unlikely as most countries have no incentive to impose a ban on all autocracies following Wenar’s plan. Autocracies control more than 60% of the world’s oil reserves (OPEC 2017). Banning their supply would cause a surge in oil prices. This would lead to severe global inflation, especially in middle and lower-middle income countries where oil constitutes 15-20% of merchandise imports (World Bank 2017). By contrast, negative ‘spillovers’ of autocratic natural resources have limited global impacts. For example, Russian Revanchism mainly harmed NATO and post-Soviet countries (Mearsheimer 2014: 84-89) while the Iranian nuclear program primarily targets American allies in the Middle East (Knepper 2008: 456-457). Therefore, countries that suffer no ‘spillovers’ from Iran and Russia would have no incentive to impose Wenar’s ban which targets all autocracies. Similarly, although the instability and injustice in many autocracies have created terrorism, it does not justify Wenar’s ban which targets many autocracies that have not contributed to international terrorism. Data from 159 countries suggests that autocracies contribute as equivalently to international terrorism as democracies (Gaibulloev et. al. 2017: 491).
Only energy-independent countries that suffer significantly from negative ‘spillovers’ from autocracies are likely candidates to benefit from and adopt Wenar’s proposal. However, the only countries which seem to satisfy this classification are countries such as the United States and Norway, due to their extensive carbon deposits and exposure to geopolitical threats and terrorisms generated by instabilities in autocratic resource-exporters. However, these countries are isolated internationally as they only consist of around 20% of the world’s total merchandise imports (WTO 2018). This means that the taxation plan would be ineffective as countries defying the ban can divert trade.

Secondly, Wenar’s ban is inconsistent with his principles. The ban only targets natural resources ignoring other economic assets such as private corporations which can be stolen and traded externally. Consumers could be equally complicit if the property rights of these stolen assets were recognised. Wenar may argue that economic assets that are not resources are often private properties that warrant no sanctions. However, analysis of 165 countries suggests that autocracies require large foreign reserves so that luxury goods can be imported to maintain the loyalty of the elites (Vadlamannati and Soysa 2018: 4). Thus, if the sanctions only target natural resources, autocracies would be incentivised to appropriate the private export sectors to substitute for the foreign reserves lost due to the natural resource ban, in order to maintain regime stability. For example, the Islamic Revolutionary Guard Corps controlled 68% of the Iranian export sector following the American oil sanctions (Wehrey et. al. 2009: 65). Such economic dominance was obtained by appropriating private property through anti-competitive behaviours such as government-provisioned market monopoly and collusion with government to acquire juicy privatisation deals (Alfoneh 2010).

Moreover, the ban may destabilise resource-exporters and therefore contradicts the principle of self-interest. Autocratic resource-exporters are highly reliant on natural resources to maintain stability. Oil industries generate 80% of revenue for Arab autocracies and over 70% of total employment (IMF 2016: 13). Additionally, 20% of the population of Arab autocracies are between the ages of 15-24 (United Nations Economic and Social Commission for Western Asia 2007). By strangling the oil industry, the ban will consequently cause fiscal crises in those countries and create mass youth unemployment. Such a situation would likely generate civil wars. According to Urdal (2004: 16), data from all political entities since 1945 indicates that the combination of a youthful population and high unemployment has a robust positive relationship with armed conflicts. Thus, ‘spillovers’ are exacerbated when unjust yet stable regimes are destabilised, allowing extremist groups to exploit fresh suffering for sympathy and recruits.

Wenar (2008: 31) could counter that the suffering is short-term and not comparable to long-term benefits of democratisation. However, autocratic resource-exporters overwhelmingly tend to be personalistic autocracies which concentrate power on a single leader (Wright 2008: 325). Thus, leaders lose tremendous privileges when regime change occurs. As a result, 90% of autocratic leaders from personalistic regimes between 1946-2010 refused to relinquish power peacefully (Geddes et. al. 2014: 328). This gives autocratic forces, such as the military, better chances than democratic groups to seize power through violence (Geddes et. al. 2014: 330). Therefore, Wenar’s plan is likely to produce violent autocratic transitions, rather than democratisations.

3. Modification of the Proposal

I propose a total economic embargo on failed states as an alternative to Wenar’s proposal. I define failed states as countries whose governments have lost their monopoly over violence and are unable to control the majority of their territories and populations (Rotberg 2004: 5).
Within failed states, armed conflicts contesting territories between at least two rival groups take place frequently (Urdal 2004: 7). Thus, if a resource-exporter becomes a failed state, the fighting would hamper resource production and cause prices to surge in the international market (Luciani 2011: 20). All countries purchasing resources will then have an incentive to cooperate in enforcing a ban which aims to end the instability and allow the resumption of production. Moreover, according to Piazza (2009: 78), data from 197 countries between 1973-2003 shows that the citizens of failed states are significantly more likely to participate in international terrorism. This generates further incentives for international cooperation in stabilising failed states, rendering the proposal more feasible. For example, the Kimberley Process that banned ‘blood diamonds’ was instituted after the 911 attack when conflict diamonds were reported to have financed Al Qaeda (Feldman 2003: 854-855).

A total economic embargo would incentivise failed states to attain stability. Armed groups in failed states obtained their military capability through smuggling natural resources and other illicit trades (Conrad et. al. 2019: 613). Thus, the total economic embargo can lead to conflict resolutions by significantly reducing the utility of fighting. Escribà-Folch (2010: 137) used a sample of 87 civil wars to show that total economic embargo shortens conflicts by making conflict termination 30% more likely. Moreover, the sanctions imposed by international organisations help to attain stability by making peaceful conflict resolution 30 times more likely (Escribà-Folch 2010: 141). Once peace is attained, the embargo should be relaxed to include only natural resources so that international trade and investment can resume in the private sector. This would alleviate the suffering of the population by providing employment and capital (Bray 2009: 20).

Wennmann (2007: 440) shows that armed groups can exploit the weak state capacity of the post-conflict states to obtain illicit financing. The availability of funds could be used to purchase weapons to restart conflicts (Wallensteen et. al. 2006: 31). Thus, the resource embargo should only be lifted after the targeted countries embark on UN-sponsored Security Sector Reform (SSR) and Demobilisation, Disarmament and Reintegration (DDR). SSR refers to the construction of the capability and accountability of all coercive instruments in society, including the military and the police force (Jackson 2011: 1806). DDR refers to the process of taking arms away from combatants and facilitating their transition into the national military or civilian life (United Nations 2014: 24). DDR and SSR programs subject countries to constant monitoring and evaluation to ensure they are on track (United Nations 2014: 117). When implemented, DDR prevents renewal of conflict by denying armed groups the resources to conduct violence (Knight 2010: 42-46) while SSR attains state-building by establishing the monopolies of violence (Dudouet 2011: 23). Resource-exporters such as Angola achieved peace and stability after SSR and DDR (Knight 2010: 42-46).

The modification makes Wenar’s proposal more coherent. Since failed states breed terrorism, the embargo that promotes conflict-resolution would deny terrorists their most fertile recruitment grounds. Moreover, post-conflict societies often experience a ‘peace dividend’ as long depressed consumption and investments resume. Post-conflict countries grow 1.14% faster than the global average (Collier and Hoeffler 2002: 4). This provides enormous opportunity for foreign direct investment which could benefit the economies of the consumer countries. This reliably resolves the spillover effects and serves the self-interest of consumers. The re-institution of the monopoly of violence also means the new government is channelling part of the resource revenues to fund basic public services and expand the security state to all corners of society to prevent relapse into instability. Citizens are then able to enjoy basic public goods such as security (Jackson 2011: 1810). Thus, popular sovereignty is improved as resource revenues are extracted from armed groups to benefit the general population.
Wenar may argue that this means accepting stolen resources from stable autocratic regimes. However, taking aim at a small, more reliable improvement is better than risking zero improvement. Even if democratisation is attained, ‘theft’ can still occur. If we define popular sovereignty over resource ownership as the use of resources to benefit the people (Van der Ploeg 2011: 321), then partisan politics in democracies may contradict this goal. In particular, the electoral politics of newly democratised states created under Wenar’s proposal tend to be dominated by non-credible politicians, making clientelism the electoral norm (Keefer 2007: 819-820). Incumbent politicians will be incentivised to utilise natural resources for the sole purpose of rewarding supporters for re-election. Such behaviour is tantamount to corruption as public power is utilized for private ends. Independent investment funds using the revenues of resource sales could be established to counter clientelism. However, even with constitutional safeguards, such funds are vulnerable to raids when the government changes hands. This encourages incumbent governments receiving resource windfalls to invest in irreversible partisan projects catering exclusively to their supporters (Van der Ploeg 2011: 406). This is hardly respecting popular sovereignty as resources are only used to benefit a segment of the populace.

In conclusion, I have argued that Wenar’s has taken an approach to tackling autocratic resources that does not sufficiently take all relevant effects into consideration. Wenar’s approach ignores the diverse interests of the international community and the complex political structures of autocracies. Thus, Wenar’s solution is ineffective and inconsistent with his objectives, with the inconsistency compounded by his exclusive targeting of natural resources. My proposed modifications seek to correct these shortcomings by aiming to reliably improve popular sovereignty.
References


Van der Ploeg, F. 2011. “Natural resources: curse or blessing?” *Journal of Economic Literature* 49(2): 366-420. URL: https://doi.org/10.1257/jel.49.2.366


Tariffs and Trusts Won’t End Illegitimate Sales of Natural Resources by Repressive Regimes

Lynne Sakr*

* Philosophy, Logic and Scientific Method, London School of Economics, Houghton Street, London, WC2A 2AE, UK. Email: lynnesakr@gmail.com. URL: https://www.linkedin.com/in/lynne-sakr-21622976/

Abstract

In response to Wenar’s (2008) proposal which addresses the customary rule of ‘might is right’, I argue that his approach fails to provide structural solutions to the illegitimate sale of natural resources by repressive governments. I first explain how his minimal conditions are not able to legitimise sales as they are not based on an exhaustive and authoritative criterion that would prevent false positives. I then contend that Wenar’s trust-and-tariff mechanism, the Clean Hands Trust, depends upon unrealistic expectations of the resource-cursed government’s reactions and its citizens’ position.

Keywords: Resource Curse; Repressive Regime; Economic Sanctions; Oil; Wenar

1. Introduction

In response to the customary rule of ‘might is right’ which reigns in global resource markets, Wenar (2008) proposes that countries stop trading with governments that are not legitimately authorised by their national population to sell natural resources. In doing so, countries would stop funding corrupt authoritarian elites in resource-cursed states. The term ‘resource curse’ refers to the adverse effects on a country’s social, political and economic infrastructures resulting from its large endowment of natural resources, particularly where these resources fall in the hands of repressive regimes (Ross 2015: 240). Wenar first establishes the minimal conditions necessary to legitimise sales and an authoritative standard that determines when the minimal conditions do not obtain (Wenar 2008: 19-25). He then develops national and international mechanisms to block any remaining illegitimate resource sales (Wenar 2008: 25-31).

In this paper, I argue that Wenar’s proposal fails to provide structural solutions to the sale of illegitimate resources by repressive states as its minimal conditions are difficult to verify. Furthermore, its main enforcement mechanism is not realistically and efficiently conceived. I first discuss how Wenar’s minimal conditions fail to provide infallible criteria against which the legitimacy of sales can be assessed. I then explain how Wenar’s international tariff mechanism to enforce property rights is based on unrealistic expectations of the situation in
the resource-cursed country. It is important to note that my analysis will focus on assessing Wenar’s minimal conditions and one of his enforcement mechanisms. As such, my conclusion is bound to be limited.

2. The Fallibility of Wenar’s Minimal Conditions

In order to move away from illegitimate sales by resource-cursed countries, Wenar proposes the minimal set of conditions necessary for citizens to legitimise the sale of national resources. I argue that these conditions weaken Wenar’s proposal because they do not include authoritative frameworks. In other words, the challenges posed by verifying the satisfaction of these conditions make it difficult and risky to use the conditions as the bases for the legitimacy of sales. Wenar’s minimal conditions are as follows (Wenar 2008: 20-21):

1. Citizens should have access to “reliable general information” concerning the resources and their trade.
2. The state must offer mechanisms that allow and acknowledge citizens’ dissent to the sales.
3. Authorisation given by citizens must be minimally independent of the government’s will.

Wenar also proposes the lowest Freedom House rating of 7 as a minimum threshold to establish that a country definitely fails to meet his conditions (Wenar 2008: 24-5). In order to test the strength of these conditions, the following analysis focuses on their potential satisfaction in countries that are exceed this minimum standard, i.e. those with a Freedom House score above 7. This includes countries like Libya, Nigeria and the Democratic Republic of Congo (Freedom House 2018: 1226), which have all previously been mentioned by Wenar (2016) as states, which illegitimately sell natural resources.

The first and third conditions do not provide sufficient benchmarks to clearly differentiate between countries that satisfy the conditions and countries that do not. The first condition requires that the citizens receive “reliable general information” regarding the resource sales conducted by the government (Wenar 2008: 21). Although a requirement of access to information does seem necessary, Wenar does not include a criterion against which the quality of the information provided could be assessed. We could then imagine a resource-rich country establishing mechanisms to publicly verify the information it provides while still withholding crucial details from the people. In this case, the state would still be providing ‘reliable’ and ‘general’ information. However, one could plausibly object that this country is not properly disclosing of all relevant information. In order to ensure they are voluntarily authorizing the sales, Wenar’s third condition requires that citizens be independent from the government’s will to “some minimal extent” (Wenar 2008: 21). He enumerates the types of dependence that could be exhibited, such as brainwashing (Wenar 2008: 21), but he does not illustrate the point at which dependence turns into minimal independence. A definition of the latter is crucial in order to clarify what is expected of the resource-cursed country. Without proper frameworks for assessing the information provided by the state and the citizens’ level of independence from the state, Wenar’s conditions cannot systematically and infallibly ensure an environment in which the people can knowingly and willingly authorize the states’ sales.

The second condition necessitates that the citizens be able to “peacefully express dissent” (Wenar 2008: 21) to the sale of their national resources “without justified fear of major harms” (Wenar 2016: 228). The government must allow and acknowledge their dissent but, in order to satisfy the condition, it seems as though any kind of dissent is acceptable. In
other words, Wenar does not include whether the dissent must be independent from any coercion, influence or limit from the state. The ability to peacefully dissent is enough for the satisfaction of this condition. A lack of requirements on dissent can be dangerous since Wenar's proposal addresses corrupt and authoritarian governments (Wenar 2013). These repressive state structures are inherently averse to dissent. As such, a vague second condition could incentivise such states to set up ad-hoc formal and informal platforms allowing citizens to peacefully dissent but only to a limited extent under monitored conditions. They would then be able to satisfy the second condition without substantially empowering the voices of their citizens. In response, Wenar could argue that the conjunction of this condition with the third one would prevent such scenarios of dependent dissent. This is true. However, as seen earlier, it is difficult to rightly verify the satisfaction of Wenar's third minimal condition. Thus such scenarios are not completely avoided. Wenar’s second condition would benefit from clearer guidelines as to what kind of dissent is required from the state.

Since the satisfaction of the conditions could plausibly lead to illegitimate sales, it seems as though they are not currently jointly sufficient. Nonetheless, where the minimal conditions fail to meet Wenar’s desired aim, the enforcement mechanisms he develops could potentially strengthen his proposal.

3. An Unrealistic and Inefficient International Mechanism: the Clean Hands Trust

Wenar (2008: 25-31) establishes national and international mechanisms aimed to enforce property rights against illegitimate sales. I focus on the trust-and-tariff mechanism which he coins The Clean Hands Trust (Wenar 2008: 28-31). This Trust is meant to keep the proceeds of antitheft tariffs imposed on non-complying countries that are still purchasing from the thieving states. By strengthening their property rights, he argues that the Trust benefits the citizens of resource-cursed countries in two ways. First, the Trust would weaken their corrupt and repressive states by affecting their trades (Wenar 2008: 30). Secondly, in order to gain access to the funds, the Trust would incentivise citizens to unite and overthrow their government (Wenar 2008: 29). In this section, I argue that Wenar’s mechanism is unrealistic and inefficient in addressing the issues faced by the citizens of resource-cursed states. In addition to the Trust, Wenar (2011) develops the Clean Trade Act to legally penalise agents accepting such illegitimate sales. In my analysis, I only assess the strength of Wenar’s Clean Hands Trust.

In his conception of the Trust, I contend that Wenar misjudges how the threatened state will react to the ban on trades. As mentioned earlier the countries targeted by Wenar’s proposal are most often authoritarian. By taking away a proportion of their income, resource-cursed regimes are expected to push for more repression and control over their populations in order to secure their positions (Carey 2010). Drury and Peksen (2010) argue that the economic strain caused by sanctions is often strategically used by the state’s leadership to strengthen repressive rule to the detriment of the opposition and the citizens. In his study of countries between 1981 and 2000, Peksen (2009) found that economic sanctions strongly contributed to an erosion of human rights in the targeted country. These findings suggest that Wenar’s proposal would work to harm the population rather than empower it, as it incentivises the state to increase its grip on power. To present a recent example, in January 2019 the United States imposed sanctions on resource-cursed Venezuela by banning American firms from trading with the state-owned Petróleos de Venezuela. Instead of weakening president Maduro, who has been able to forcefully latch on to power despite months of protests (Keating 2019; Specia 2019), these sanctions have further exacerbated the Maduro-led...
humanitarian crisis with shortages of food and medicine on the rise (Sullivan 2019; Sachs and Weisbrot 2019: 18).

Wenar’s proposal indeed overlooks the economic repercussions endured by the population of resource-cursed states as a result of decreased resource trades. Although the government does receive most of the profits from its illegitimate sales, large parts of the population in these countries still depend on the resource industries. For example, four million people directly benefit from small-scale resource mining in the Republic of Sudan (Milicevic and Gayi 2015: 3). Blocking the trade of these national resources would also mean cutting off a significant source of the population’s income. This would place a strain on their short-term livelihood, which is disregarded by Wenar in favour of the long-term benefits of respected property rights (Wenar 2008: 31). Where the governing elite may feasibly have contingency plans to partly mitigate the cut in resource income, the population does not have such luxuries, especially not in the short-term. How does Wenar expect weakened populations to revolt against their targeted, yet relatively empowered, governments? Wenar might argue that foreign funds will support the population. Nonetheless, he insists that the Trust will only be made available once a new “minimally decent” government is in place (Wenar 2008: 29). How are the people expected to survive in the meantime? This question is especially prevalent in cases where the incumbent’s power is much stronger and more persistent than expected, as seen in the case of Venezuela. It is therefore fair to conclude that the acts of rebellion expected of the national citizenry will face practical difficulties.

Finally, the creation of a Trust does not endow the people of resource-cursed countries with long-term solutions and protection. Let us assume that the resource-cursed country receives the funds once the opposition forms a new government. The concern would then be in ensuring the funds are actually used in the people’s interest. As the new government would still be operating within a cursed infrastructure, where a lack of transparency and accountability reign, there is no certainty that the funds would be used for their intended purpose. Wenar (2008) does require that the new government be “minimally decent” before receiving the funds. However, as discussed in the previous section, the satisfaction of the minimal conditions is difficult to verify. Thus it seems plausible that a government which is only ‘minimally decent’ on the surface would be able to satisfy the required conditions. It could then gain access to the Trust without the intention of using it entirely in the people’s interest. Once again the state, rather than the people, would be empowered. Even if the funds were used exclusively in the interest of the people, Wenar’s proposal does not prevent the new government from changing course and illegitimately selling resources in the future, despite a potential ban. This point highlights the fact that these aspects of Wenar’s proposal do not structurally tackle the root of the problem because they do not efficiently address the political, social and economic weaknesses of the resource-cursed state. Consequently, the proposal does not equip the people with long-term tools to protect themselves from the Resource Curse. They still face the risk of exploitation by the new government, which brings us back to the initial problem.

4. Conclusion

In assessing parts of Wenar’s proposal for banning trade with resource-cursed countries, I have argued that they fail to provide robust solutions to illegitimate sales of resources by repressive governments. Firstly, I explained why the minimal conditions he proposes are not currently able to legitimise sales, as they do not include exhaustive criteria that can prevent false positives. Providing infallible frameworks for these conditions would ensure that they become jointly sufficient. Secondly I addressed the Clean Hands Trust, an international mechanism Wenar develops to enforce property rights. I argued that this Trust is based on
unrealistic expectations of the resource-cursed government’s reaction and its citizens’ position. I added that, even if the Trust were to function as expected, the proposal does not efficiently address the structural weaknesses within the cursed country.
References


