

Bowdoin College

## Bowdoin Digital Commons

---

Philosophy Faculty Publications

Faculty Scholarship and Creative Work

---

8-1-2017

### Solving which trilemma? The many interpretations of equality, Pareto, and freedom of occupational choice

Kristi A. Olson  
*Bowdoin College*

Follow this and additional works at: <https://digitalcommons.bowdoin.edu/philosophy-faculty-publications>

---

#### Recommended Citation

Olson, Kristi A., "Solving which trilemma? The many interpretations of equality, Pareto, and freedom of occupational choice" (2017). *Philosophy Faculty Publications*. 2.  
<https://digitalcommons.bowdoin.edu/philosophy-faculty-publications/2>

This Article is brought to you for free and open access by the Faculty Scholarship and Creative Work at Bowdoin Digital Commons. It has been accepted for inclusion in Philosophy Faculty Publications by an authorized administrator of Bowdoin Digital Commons. For more information, please contact [mdoyle@bowdoin.edu](mailto:mdoyle@bowdoin.edu), [a.sauer@bowdoin.edu](mailto:a.sauer@bowdoin.edu).

# Solving which trilemma? The many interpretations of equality, Pareto, and freedom of occupational choice

Politics, Philosophy & Economics  
2017, Vol. 16(3) 282–307  
© The Author(s) 2017  
Reprints and permissions:  
sagepub.co.uk/journalsPermissions.nav  
DOI: 10.1177/1470594X17704401  
journals.sagepub.com/home/ppe  


**Kristi A Olson**  
*Bowdoin College, USA*

## Abstract

According to the trilemma claim, we cannot have all three of equality, Pareto, and freedom of occupational choice. In response to the trilemma, John Rawls famously sacrificed equality by introducing incentives. In contrast, GA Cohen and others argued that we can, in fact, have all three provided that individuals are properly motivated by an egalitarian ethos. The incentives debate, then, concerns the plausibility of the ethos solution versus the plausibility of the incentives solution. Considerable ink has been spilled on both sides of the debate. Yet, in this essay, I argue that we cannot have this debate until we clarify the terms. Once we clarify the terms, however, we might discover that there is no debate to be had. This is because, depending on how equality, Pareto, and freedom of occupational choice are interpreted, there might not be a trilemma in the first place. Specifically, I use a small but crucial distinction in how equality, the egalitarian ethos, and Pareto are assessed – what I call the internal/external distinction – to disentangle the various paths each solution – the ethos or incentives – could take. I conclude that both solutions have gained illicit plausibility by virtue of not keeping the distinction straight.

## Keywords

trilemma, ethos, GA Cohen, equality, Pareto, freedom of occupational choice, incentives

---

## Corresponding author:

Kristi A Olson, Department of Philosophy, Bowdoin College, 8400 College Station, Brunswick, ME 04011, USA.

Email: [kolson@bowdoin.edu](mailto:kolson@bowdoin.edu)

## Introduction

Imagine that we live in an egalitarian society in which everyone receives the same income for a full day's work.<sup>1</sup> In this egalitarian society, Abigail runs a small catering business for the same income as everyone else. Nonetheless, Abigail has the rare gift of impeccable business sense. If she became a business consultant, businesses would become more profitable. And, since our egalitarian society would distribute these gains equally, everyone would benefit. There is, however, one catch: Abigail prefers her current career running a catering business. Indeed, she would need at least £10,000 extra pay as a business consultant in order to be indifferent between the two jobs. The additional £10,000 pay is insignificant compared to the total gains: even if we paid Abigail the extra money, everyone would still be better-off. Nonetheless, paying Abigail the extra money would sacrifice equality in our otherwise egalitarian society.<sup>2</sup>

The scenario above illustrates what has come to be known as the trilemma, which says that we cannot have all three of equality, Pareto, and freedom of occupational choice.<sup>3</sup> As GA Cohen (2008: 185), among others, has noted, we seem to be faced with the following three options: we could insist on equal pay and allow Abigail to pursue the catering career she prefers, we could insist on equal pay and force Abigail to become a business consultant, or we could pay Abigail just enough additional income to entice her to become a business consultant. Yet none of these options secures all three of our desiderata: the first sacrifices Pareto, the second violates freedom of occupational choice, and the third forfeits equality (Cohen, 2008: 185).

In response to the trilemma, John Rawls (1999) famously endorsed the last option. Rawls's difference principle would permit Abigail enough additional income to entice her to become a business consultant. Yet, according to some of Rawls's critics – most notably, GA Cohen – if individuals were as committed to equality as Rawls stipulates they are, then they would not need such incentives.<sup>4</sup> If Abigail were to act in accordance with an 'egalitarian ethos' – a set of moral principles governing occupational choices – she would freely choose to become a business consultant at equal pay. And thus, we can in fact have all three of equality, freedom of occupational choice, and Pareto – or so proponents of the ethos solution claim.

To be sure, proponents of the ethos are not making an empirical claim about what will actually take place; they are not arguing that individuals will in fact be so motivated. Like Rawlsians, then, proponents of the ethos might embrace incentives as a policy matter, especially if individuals would otherwise be very badly off. Nonetheless, the ethos solution rescues the *ideal* of equality. The reason the trilemma originally appeared to be so threatening, according to (Cohen, 2008: 185), was because it implied that equality could be endorsed only at the cost of freedom or Pareto. The ethos solution saves us from this Sophie's choice because, 'if we were egalitarians, we should be sacrificing neither' (Cohen, 2008: 196).

Considerable ink has been spilled on both sides of the incentives debate. Yet, as I will argue in this essay, we cannot have a meaningful debate until we clarify the terms.<sup>5</sup> And, once we clarify the terms, we might discover that there is no debate to be had. This is because, depending on how the terms are clarified, there might not be a trilemma in the first place. Specifically, in this essay, I use a small but crucial distinction in how equality,

the egalitarian ethos, and Pareto are assessed to disentangle the various paths each solution – the ethos or incentives – could take. I do not, and indeed cannot within the scope of one essay, follow each path down all of its twists and turns. Nonetheless, the very existence of these many paths suffices to establish why the debate cannot (yet) take place.

The distinction I make in how equality, the egalitarian ethos, and Pareto are to be assessed is not unknown.<sup>6</sup> Yet, the debate typically proceeds without making these distinctions – to its detriment. In particular, *both* the ethos and the incentives solution gain illicit plausibility by virtue of ignoring the distinction. In Part I, I lay out GA Cohen's ethos solution as a way of entering the debate. My general argument, however, is not tied to Cohen's particular version of the ethos or, for that matter, to Rawls's particular version of incentives. Indeed, my essay will address the Pareto argument for incentives, which, as we will see, is importantly different from Rawls's argument for the difference principle. In Part II, I identify the various paths the ethos solution could take and demonstrate the different assumptions and problems for each. Part III does the same for the incentives solution.

The upshot of Part II is that the ethos solution is exceedingly fragile. Only under very particular specifications of the terms will the ethos be a necessary, adequate, and plausible response to the trilemma. Yet, it would be hasty to conclude from this that equality is once more on the chopping block; the upshot of Part III is that the incentives solution is also fragile. Indeed, it is hard to see how the Pareto argument for incentives can be endorsed without fundamentally changing the nature of the debate. The main conclusion of my article, then, is that we should reject both sides of the debate, at least as they are currently presented.

Before we begin, let me explain why, instead of describing Abigail and similarly situated individuals as the 'talented' who help the 'worst off', I will refer to 'agents' who help 'beneficiaries'. I do so for two reasons. First, we should not prejudice the incentives debate by presupposing that Abigail and those similarly situated are already better-off (for example, by virtue of being more talented). Nor should we presuppose that some individuals are already worse off. To the extent individuals are not equal at the starting point, those inequalities require a distinct justification that I do not consider here.<sup>7</sup> Second, the assumption that individuals in Abigail's position are necessarily more talented is simply not true. Instead, the only necessary differences between agents and beneficiaries are that (i) agents do not prefer the option that maximally benefits their society and (ii) the extent to which an agent disprefers the option that maximally benefits her society – as measured by the additional income necessary to make her indifferent between the two options – is less than the benefit to her society (that is, the benefit to society exceeds the cost of an incentive). Thus, it is entirely possible for the most talented individual in our society to be a beneficiary and for the least talented individual in our society to be an agent.

One final clarification is in order: The Pareto requirement Cohen (2008: 184) endorses is the 'weak' one that prohibits maintaining one state of affairs if there is another state of affairs in which everybody would be better-off. This requirement is weak in the sense that it is violated only in the rare circumstance in which *everybody* could be made better-off. A strong Pareto requirement, in contrast, is violated whenever *at least one* individual could be made better-off without making anyone worse off.

Following Cohen's lead, I will also assume the weak Pareto requirement. Since weak Pareto involves circumstances in which *everyone* would be made better-off, it is the hardest version to reject.

The arguments I make in this essay primarily concern a distinction in the way equality, the egalitarian ethos, and Pareto is assessed. Before we can see why this distinction matters, however, we first need a shared understanding of the trilemma and the ethos solution. That, then, is the task of Part I.

## I

A. According to what GA Cohen (2008: 189) calls the trilemma claim, it is impossible to have all three of equality, freedom of occupational choice, and Pareto, at least in the absence of a 'wildly improbable fluke'. Cohen specifies the conditions under which each of the three desiderata is satisfied as follows: In order to satisfy the equality desideratum, nobody should be '*substantially* better off than others are with respect to *both* income and job satisfaction' (Cohen, 2008: 184). In order to satisfy the freedom of occupational choice desideratum, nobody should be 'coerced into particular jobs' (Cohen, 2008: 184). And, in order to satisfy the Pareto desideratum, we should not maintain one state of affairs if there is another state of affairs in which everybody would be better-off (Cohen, 2008: 184).

Before turning to Cohen's ethos solution, let me first preempt a possible source of confusion. The trilemma, as Cohen presents it, is not based on the idea that inequalities in pay are necessary because, in the absence of such inequalities, individuals would not know how to allocate their labor efficiently.<sup>8</sup> After all, if inequalities in pay were necessary merely for signaling purposes, this would be consistent with allowing the market to operate and then taxing the incomes down to equality (Cohen, 2008: 122) referring to Carens (1981). See also Wilkinson (2000). Individuals would know where their labor is most needed; that would be determined by the pretax incomes. Equality, however, would be preserved in after-tax incomes. Thus, we can have the signaling function of the market without forgoing equality.

Instead, the trilemma is thought to arise because, in the absence of inequality-generating incentives, the individual's preferred occupation is unlikely to coincide with the occupation that maximally benefits her community. The problem in the example with which I started the article is not that Abigail does not know that the community would be better-off if she switched to consulting. Rather, the problem is that, at equal levels of pay, she disprefers the option that maximally benefits her community. It is this fact that gives rise to the trilemma.<sup>9</sup>

Of course, as Cohen acknowledges, it is possible that everyone could just happen to prefer the option that maximally benefits their community. And, if that did happen, then the three desiderata could be simultaneously satisfied. Nonetheless, although possible, such a scenario is 'wildly improbable' (Cohen, 2008: 189). The task Cohen undertakes, then, is to show that we can satisfy all three desiderata, even in the absence of a wildly improbable fluke.

B. According to the ethos solution, the trilemma disappears once we add the idea that individuals might be motivated by some sort of egalitarian commitment to pursue the occupational position that maximally benefits society, even when doing so requires a career sacrifice on their part. Cohen (2008: 189–191) uses the following analogy to

illustrate. Suppose we want an adequate supply of blood, without paying individuals for donating blood but also without forcing them to donate. The problem, according to Cohen, is structurally similar to the original trilemma. It seems as though we must sacrifice at least one of the three desiderata: if we do not pay individuals for donating blood and we do not force them to donate blood, then we cannot secure an adequate supply. Cohen's solution is to say that we can, in fact, have all three, provided that individuals are motivated to give blood 'through some combination of principled commitment and fellow feeling' (2008: 189). And, as Cohen (2008: 189) notes, at least in certain times and places, a sufficient number of individuals have been motivated to give blood without being forced or paid to do so.

According to Cohen, a roughly analogous solution applies to the original trilemma. If individuals are similarly motivated to choose the occupation that maximally benefits society, then we can, in fact, have all three of equality, freedom of occupational choice, and Pareto (Cohen, 2008: 189). Since individuals are not paid more to choose those occupations, equality is maintained. Since individuals are not forced to choose those occupations, freedom of occupational choice is preserved. And, since there is no other state of affairs in which everybody would be better-off, the Pareto criterion is satisfied. Thus, according to the ethos solution, we can simultaneously satisfy all three – provided that individuals are properly motivated.

Of course, the ethos solution does not *guarantee* that we will secure all three. The relevant individuals, after all, might not be properly motivated. The trilemma claim, however, stated that it was *impossible*, apart from a wildly improbable fluke in which each individual just happens to prefer the option that maximally benefits her community, to have all three. Yet, if the ethos solution works, that claim is false. And this rescues the ideal of equality.

C. In Part II, I will evaluate whether or not the ethos solution succeeds as a response to the trilemma.<sup>10</sup> In order to do that, however, we first need to understand the 'better-off' relationship. In portraying the three desiderata, Cohen (2008: 184) invokes the phrase better-off twice: First, the equality desideratum asks whether one individual is better-off than other individuals. Second, the Pareto desideratum asks whether there exists a state of affairs in which all individuals would be better-off. Yet, despite the fact that both desiderata ask whether an individual is/would be better-off, it is not clear that the metric remains the same.

With respect to equality, Cohen is (in one respect) explicit: 'equality is satisfied if no one is *substantially* better off than others with respect to *both* job satisfaction and income' (2008: 184).<sup>11</sup> In contrast, with respect to Pareto, Cohen tells us only that the Pareto desideratum prohibits maintaining a state of affairs if there is another state of affairs in which everybody would be better-off; he does not specify better-off *in what respect*. Nonetheless, given the example Cohen uses – in which a gardener could make the rest of her community better-off by becoming a doctor – it seems as though he had a different metric in mind. After all, the reason individuals would be better-off if the doctor-gardener became a doctor is presumably not because this would improve their income or job satisfaction (although it might do that as well), but rather because it would improve their health.<sup>12</sup> Cohen's Pareto metric, then, seems to include (among

other things) health. His equality metric, in contrast, is limited to job satisfaction and income.

My purpose here, however, is not to decide what each metric should or should not include. My claim is only that the metric for each desideratum needs to be specified. And, if the metric for Pareto is different from the metric for equality, we might reasonably demand an explanation for the difference. (Why, after all, would egalitarians care if someone is better-off in one state of affairs than in another state of affairs but not care if someone is better-off in that way than other individuals?<sup>13</sup>) Thus, before we can evaluate the trilemma or the solutions to it, more needs to be said. Nonetheless, for our purposes, I will assume that the same metric is used for both and will restrict my focus to just one question: how is the metric to be assessed?

At the most general level, there are two approaches: what I am calling the internal standard and the external standard.<sup>14</sup> The internal standard, as I am defining it here, defers *entirely* to the individual's first-personal assessment. On such a standard, we do not need to determine what criteria to include in our metric or how much weight to give different criteria because such considerations are left entirely to each individual. Whatever the individual includes is, by definition, relevant. The external standard, in contrast, does not defer entirely to the individual's first-personal assessment (although the individual's first-personal assessment could be one among other considerations).

The internal assessment comes in two versions, each with its own assumptions and problems. The first version assumes that interpersonal comparisons of utility are possible and sets each individual's wages such that everyone has the same utility, regardless of the job they perform. Under this version of the internal assessment, in our initial starting point – in which Abigail runs a catering business but could make everybody better-off by becoming a business consultant – everyone, by hypothesis, has the same utility. If Abigail prefers running a catering business to becoming a business consultant, then (assuming preferences correspond to utility<sup>15</sup>) equality requires that she receive extra compensation as a business consultant. Otherwise, Abigail's utility as a business consultant would be lower than everyone else's, in violation of equality.

The second version of the internal assessment, in contrast, selects a common reference bundle (for example, a basic income for no work) and then adjusts each individual's wages such that each individual is indifferent between her job–wage bundle and the reference bundle.<sup>16</sup> To illustrate: in the initial scenario, we adjust Abigail's compensation as a caterer such that she is indifferent between catering and the reference bundle. If Abigail prefers running a catering business to becoming a business consultant, then equality requires that she receive extra compensation as a business consultant. Otherwise, Abigail would not be indifferent between being a business consultant and the reference bundle, in violation of equality.

By appealing to a common reference bundle, the second version of the internal assessment avoids the controversial assumption that interpersonal comparisons of utility are possible. Yet, in doing so, it introduces a new problem: the choice of reference bundle needs to be justified. To see why, compare two possible reference bundles: no work/low pay and hard work/high pay. Individuals who find unemployment and low pay more attractive will do better with the former reference bundle; in contrast, individuals

who find hard work and high pay more attractive will do better with the latter reference bundle. Thus, if the internal assessment is to be plausible, the choice of reference bundle would need to be justified.<sup>17</sup> For now, however, I will set these concerns aside. What matters for our purposes is that, under both versions of the internal assessment, Abigail is indifferent among her equal options.

The external assessment, in contrast, does not defer entirely to Abigail's first-personal assessment. There are a number of reasons why such deference might seem inappropriate. For example, an egalitarian who thinks that an individual's actual preferences are a poor guide to what is best for her might have a paternalistic reason for endorsing an external assessment that instead tracks the individual's fully informed and rational preferences. Yet, this reason for endorsing an external assessment seems at odds with the incentives debate. To see this: Suppose that an individual's options are set such that they are equal according to her fully informed and rational preferences. Nonetheless, she prefers her current option to the equal option that maximally benefits her society. The reason we now need an incentive or the ethos is to entice the individual to do what is in fact aligned with her interests (if only she could appreciate that fact). This, however, fundamentally changes the common understanding of the debate. The argument for the ethos, after all, is not: 'Abigail, you should become a business consultant at no extra pay because it maximally benefits your community, *and, in addition, it is consistent with the interests of your fully informed and rational self*'. Rather, the reason an incentive or an ethos is thought to be necessary is because the agent's interests (including her fully informed and rational ones) *are not aligned* with the option that maximally benefits her community. Thus, basing equality on the individual's fully informed and rational preferences changes the incentives debate in a way that most of the parties involved in the debate would, I assume, reject.

Nonetheless, there are other versions of the external assessment that do seem consistent with the incentives debate. Two, in particular, are worth mentioning. First, egalitarians of a certain sort might refrain from deferring to individuals' preferences because they believe individuals should be held responsible for their choices. Second, egalitarians of a different sort might refrain from deferring to individuals' preferences on the basis that their preferences, and hence the internal assessment, cannot be publicly verified. These reasons are by no means exhaustive. Nonetheless, they suffice to capture a distinction in the reasons for endorsing an external assessment that will later become salient.

Before turning to our analysis, let me mention another, perhaps more intuitive way of capturing the distinction: internal assessments are subjective, while external assessments are objective. Although this captures the general idea and has the benefit of being a familiar distinction, I avoid this terminology because it conceals an important ambiguity. In particular, the subjective/objective distinction is commonly used to track two distinct variables that can come apart. To illustrate, imagine we ask each individual how much additional pay would make her indifferent between the basic income/no work reference bundle and being a business consultant. We then take the median individual's response and apply it to Abigail. We might reasonably balk at calling such an assessment objective. After all, we are just aggregating – or, rather, selecting from – subjective assessments; we have not appealed to any objective facts about the job. Yet, on the other hand,



we might also balk at calling such an assessment subjective. In particular, it is not subjective in the sense that matters here; it is not *Abigail's* assessment. The internal/external distinction, in contrast, avoids this confusion. An assessment is internal if and only if each individual determines her own well-being. All other assessments are external.

D. The target of Cohen's criticism, John Rawls, endorsed an external standard.<sup>18</sup> Thus, one might have expected Cohen to do so as well. Yet, Cohen rejects Rawls's claims that the distribuendum must be 'objectively measurable' and 'open to public view' (2008: 200, 369). And, in doing so, Cohen leaves open the possibility that Pareto and equality are to be internally assessed.

Indeed, with respect to Pareto, Cohen seems to have had the internal standard in mind. I return to that issue in Part III. With respect to equality, however, Cohen's stance is less clear. Elsewhere Cohen (1989) endorses an equalisandum with externally assessed components. Yet, when discussing the trilemma, Cohen sometimes seems to have had the internal standard in mind. To begin with, unlike Rawls, Cohen does not take it to be a strike against the inclusion of job satisfaction that it is not (fully) open to public view.<sup>19</sup> Instead, Cohen explicitly acknowledges that how satisfying a job is will vary across persons, yet denies that this is a reason to exclude job satisfaction from the metric.<sup>20</sup> And, although Cohen readily acknowledges that his equalisandum fails Rawls's stringent publicity test, he claims that it satisfies a weaker form of publicity: 'one's intimates at least often know how repugnant or rewarding one's job is for one' (2008: 369–370).

Even when Cohen refers to seemingly objective criteria, he still has the individual's first-personal assessment of the criteria in mind. For example, although Cohen mentions long hours of work as a criterion of burdensomeness, he also explicitly acknowledges that how burdensome long hours of work are will vary across persons.<sup>21</sup> And it is not merely the case that he thinks burdens are agent relative (an agent-relative standard could, after all, still be externally assessed); he also refers to the individual's own assessment: if 'A hates job j', then 'A should receive extra compensation' (Cohen, 2008: 202). And 'people's job-and-income packages should, other things equal, be comparable in welfare terms' (Cohen, 2008: 370).<sup>22</sup> Thus, Cohen seems to defer to the individual's first-personal assessment, consistent with the internal assessment of equality.

Of course, the fact that Cohen gives some deference to the individual's first-personal assessment does not rule out the external standard. The external standard, recall, is defined as not deferring *entirely* to the individual's first-personal assessment. And, given the views Cohen (1989) expresses elsewhere, there is good reason to think Cohen did not intend the internal standard here. Moreover, Cohen's language is often ambiguous. For example, Cohen describes his view as capable of explaining such judgments as 'It's true that he's not very well paid, but that seems not so unfair when you consider *how enjoyable his work is*' (2008: 200, my emphasis). (Contrast this with the more clearly internal standard: 'It's true that he's not very well paid, but that seems not so unfair when you consider *how much he enjoys his work*'.)

My suspicion is that Cohen left the question of how equality is to be assessed unresolved because he thought he did not need to give an answer. This is consistent

with the fact that Cohen treats his argument as generalizing across different equalisanda.<sup>23</sup> And, in doing so, Cohen is not alone: Other proponents of the ethos, such as TM Wilkinson (2000: 49, 68), also leave this issue unresolved. The task of Part II, then, is to show why the different assessments matter. And, for this task, Cohen's account once again provides a useful starting point.

## //

A. Cohen (2008: 184) endorses what he calls a 'weakish' equality requirement: 'the relevant equality might be only that no one is *substantially* better off than others are with respect to *both* income and job satisfaction'. Presumably, however, Cohen did not intend for his criterion to imply that equality is satisfied whenever an individual is not better-off with respect to *both* income and job satisfaction. To illustrate the implausibility: if Bart and Brenda perform the same job with precisely the same level of job satisfaction, but Bart is paid £200,000, while Brenda is paid only £20,000, the equality criterion is apparently satisfied. It is true that Bart is (substantially) better-off with respect to income, but he is not (substantially) better-off with respect to *both* income and job satisfaction. Yet, surely no self-respecting egalitarian would deem their situations equal.

Consider, then, two other interpretations of Cohen's equality requirement. First, we could evaluate income and job satisfaction holistically, such that two individuals might be equal even if one has more income so long as the other has more job satisfaction.<sup>24</sup> Second, we could interpret the criterion as establishing merely a necessary condition for equality: if an individual is substantially better-off in both respects, then equality is violated. So interpreted, the criterion does not specify sufficient conditions for equality.<sup>25</sup> Although the analysis is slightly different, both interpretations reach similar conclusions. Thus, for ease of exposition, I will assume the first interpretation (with the differences noted in endnotes).<sup>26</sup>

In the remainder of Part II, I will show first, that, if we endorse the internal assessment of equality, then whether or not the ethos is necessary depends on our interpretation of freedom of occupational choice. Second, if we endorse the external assessment of equality, then, even if the ethos is necessary (which depends, again, on our interpretation of freedom of occupational choice), the ethos is either too weak to avoid the trilemma or else too demanding to be plausible. Yet, third, if we endorse a mixed assessment of equality – drawing on both the internal and external assessment – then the ethos is necessary, adequate, and arguably plausible, but the solution now depends on an idiosyncratic assessment. Each of the three scenarios is addressed separately below.

B. Assume first that proponents of the ethos solution endorse the internal assessment of equality. The problem – if it can be called a problem – is that, if we have internally assessed equality, then the ethos is not necessary because there is no trilemma. To illustrate this, consider Cohen's own example (2008: 184–185) in which an individual is choosing between £20,000 as a gardener and £20,000 as a doctor. The doctor-gardener would maximally benefit her community by becoming a doctor. Yet, at these levels of pay, she prefers to be a gardener: although her job satisfaction as a doctor for £20,000 is greater than that of others, her job satisfaction as a gardener for £20,000 is *much* greater than that of others.<sup>27</sup>

Cohen treats both these options – at £20,000 pay – as satisfying the equality desideratum. Yet, if equality is internally assessed and we take both income and job satisfaction into account (trading them off of one another to reach a holistic assessment) then allowing the doctor-gardener to keep £20,000 as a gardener *violates* equality: at £20,000 as a gardener, she is better-off, internally assessed, than everybody else. If the doctor-gardener were instead to choose from internally assessed equal options, then we would not need the egalitarian ethos.<sup>28</sup> To illustrate: suppose that, if she were offered £15,000 as a gardener, her job–wage bundle, internally assessed, would be equal to the internally assessed bundles of others and, at £20,000 as a doctor, her job–wage bundle, internally assessed, would be equal to the internally assessed bundles of others.<sup>29</sup> Under these circumstances, the doctor-gardener would be indifferent between £15,000 as a gardener and £20,000 as a doctor. (Since her bundle as either the doctor or the gardener is equal, internally assessed, to the bundles of others – or to a common reference bundle – it follows that her two bundles must also be equal, internally assessed, to each other and, if her bundles are equal, internally assessed, to each other, then it follows that the doctor-gardener would be indifferent between them.) Yet, if the doctor-gardener is indifferent between her two options, we do not need Cohen’s egalitarian ethos to solve the trilemma. To the contrary, all that is required of the doctor-gardener is a willingness to use the benefit to society as a tiebreaker. Call this the non-misanthropic ethos. Whereas Cohen’s egalitarian ethos requires the individual to take the job that maximally benefits society even when she disprefers it, the non-misanthropic ethos requires only that an individual take the job that maximally benefits society when she is otherwise indifferent.

Thus, if we have internally assessed equality, then the ethos as traditionally conceived is not necessary. The much less demanding non-misanthropic ethos will do. And this is an entirely general claim: If we have internally assessed equality, then the individual will be indifferent between her options. And, if she is indifferent between her options, then (unless she is a misanthrope) she will freely choose the option that is best for her society.<sup>30</sup>

C. In the previous section, we saw that if we have internally assessed equality, then we do not need the egalitarian ethos. What I now want to suggest, however, is that we might need the egalitarian ethos in order to achieve internally assessed equality in the first place. Specifically, if reducing the doctor-gardener’s income as a gardener violates freedom of occupational choice, then we cannot secure internally assessed equality without the ethos. If the analysis below is correct, then we now have a dilemma (as well as a trilemma): even if we ignore Pareto, we cannot ensure both equality and freedom of occupational choice. To illustrate, if the doctor-gardener is free to choose from among her original unequal options (and, by hypothesis, her freedom of occupational choice prevents us from changing her options such that they are equal), she will choose to garden for £20,000, violating internally assessed equality.

There are at least two distinct objections to reducing the doctor-gardener’s income as a gardener. The first objection focuses on what Cohen calls the ‘invasion of the [individual’s] inner economy’ (2008: 221–222). Specifically, the state must somehow ascertain each individual’s preferences and capabilities and then use that information to structure the individual’s options (2008: 222). I am skeptical that this counts as a violation of *freedom of occupational choice*, although it might be a violation of a fourth

desideratum, for example, privacy. But even if we accept that it counts as a violation of freedom of occupational choice, for reasons discussed in Part II.D below, proponents of the ethos are unlikely to rest their case on the privacy violation.<sup>31</sup>

The second objection focuses on alterations to the doctor-gardener's options; that is what violates the individual's freedom of occupational choice. Yet, presumably it is not the case that *any* alteration of an individual's options violates freedom of occupational choice. That, after all, would seem to rule out even the ordinary income tax, since the tax alters the individual's options (and, moreover, may change the individual's preference ordering if, for example, there is a certain level of income she needs or desires). And surely we do not mean to imply that individuals have a claim to the options they would have in a tax-free state (even assuming we could determine what those options would be).

Perhaps, then, what is objectionable is not merely that the individual's options are altered, but rather, that they are deliberately altered in such a way to manipulate the individual into choosing a particular job. The state, however, might not have this intention; perhaps their intention is merely to make all of the doctor-gardener's options equal, internally assessed (and not to force her into a particular job). It is merely a fortunate side effect of internally assessed equal options that the doctor-gardener no longer has any incentive to avoid the option that maximally benefits her society. Moreover, it presumably would not be less of a violation if the individual's options were altered randomly.

My own view is that the reduction in the doctor-gardener's income as a gardener does violate freedom of occupational choice. Elsewhere I argue that we should distinguish a tax on the talent rent portion of an individual's income (that is, the extra pay attributable to the fact that the talents requisite for that job are scarce) from a tax on the nontalent rent portion of an individual's income. It is plausible that the latter but not the former is a *prima facie* violation of freedom of occupational choice. This would explain why certain forms of taxation impermissibly infringe on freedom while other forms of taxation do not (Olson, 2010a). If that argument is right, then assuming gardening skills are not scarce – and thus the income associated with gardening does not include talent rent – reducing the doctor-gardener's income as a gardener is at least a *prima facie* violation of freedom of occupational choice. And thus perhaps we need the ethos in order to achieve internally assessed equality.

It is beyond the scope of this article to determine when freedom of occupational choice is violated. My claim here is merely that whether we need the ethos depends on what exactly freedom of occupational choice requires, and we need to determine what exactly freedom of occupational choice requires in order to assess the cost of sacrificing it in order to solve the trilemma. Reducing the doctor-gardener's income as a gardener, after all, falls far short of 'Stalinist forcing' (Cohen 2008: 219). Thus, we cannot use our abhorrence of Stalinist forcing to justify freedom of occupational choice as a desideratum if the protections guaranteed by freedom of occupational choice are in fact much more extensive. The terms of the debate, then, need to be made clear, before we can decide whether a sacrifice is needed and, if so, which desideratum to sacrifice.

My tentative conclusions are these: If equality is internally assessed and if we can make each individual's options equal consistent with freedom of occupational choice,

then the ethos is not needed because there is no trilemma. On the other hand, however, if we cannot make each individual's options equal consistent with freedom of occupational choice, then the ethos is necessary not only to solve the trilemma but also to solve a *dilemma*. Significantly, however, the more expansive freedom of occupational choice becomes, the more willing I am to reject it: Although I would not sacrifice freedom of occupational choice if that meant accepting Stalinist forcing, I might be willing to tolerate the minor infringements necessary to make each individual's options equal. The decision of which desideratum, if any, to reject, then, crucially depends on their specification.

D. Turn now to externally assessed equality. Before we can determine whether the ethos is necessary to solve the trilemma if equality is externally assessed, we first need to clarify the equality desideratum (as well as freedom of occupational choice). Specifically, does equality require the doctor-gardener to be equal, externally assessed, to others in each of her options – including those she does not take – or does equality require only that she be equal to others in the option she actually takes?<sup>32</sup> To illustrate why this matters, assume that £20,000 as a gardener is equal, externally assessed, to £20,000 as a doctor but the doctor-gardener is indifferent between gardening for £15,000 and doctoring for £20,000. If we were to offer the doctor-gardener, say, £14,000 as a gardener and £20,000 as a doctor, she would choose to doctor for £20,000. By hypothesis, £14,000 as a gardener is inconsistent with externally assessed equality (as well as internally assessed equality). Yet, if the equality desideratum applies only to the option she actually takes, then externally assessed equality is preserved since, faced with these options, she would choose to doctor for £20,000. Thus, by reducing her income in her other options, we can induce the doctor-gardener to choose doctoring for £20,000 while arguably maintaining equality, freedom of occupational choice, and Pareto. If so, the ethos is not required.

Proponents of the ethos solution would presumably reject this strategy as violating freedom of occupational choice. Yet, unlike the previous example involving internally assessed equality, the necessary invasion of the individual's inner economy would be much less. So long as we can determine what is best for the community, we can simply dramatically reduce the doctor-gardener's income in all of her other options; we do not need to determine her indifference point. Thus, if this violates freedom of occupational choice – and it seems to me that it does – the violation presumably rests in the alteration of her options and not in the invasion of privacy. All of this, however, just reinforces my main claim: both freedom of occupational choice and equality need additional specification before we can know if the trilemma even exists or which desideratum to sacrifice.

If, as I assume to be the case, it is impermissible to reduce the doctor-gardener's income as a gardener to a level of pay below externally assessed equality because doing so would violate either equality or freedom of occupational choice, then the ethos is necessary. Before we can determine whether the ethos is also adequate and plausible, however, we need to know how the ethos itself is to be assessed. Is the individual required to accept the level of pay that would make her equal, internally assessed (the internally assessed ethos) or is she required to accept the level of pay that would make her equal, externally assessed (the externally assessed ethos)? Although I will argue that

neither version succeeds as a solution to the trilemma, the reasons they fail are quite different: the internally assessed ethos is too weak to escape the trilemma, while the externally assessed ethos is too demanding to be plausible.<sup>33</sup>

The following example illustrates both possibilities. Suppose that everyone in society except Alicia is very happy with their externally assessed equal job–wage bundle. Alicia is the one exception. Although Alicia is equal, externally assessed, to everybody else, she is significantly worse off, internally assessed. Nonetheless, suppose that, even though Alicia is already the worst off individual, internally assessed, she could benefit the rest of her community by taking a different job, albeit a job that would make her even worse off, internally assessed.

If the egalitarian ethos is internally assessed, then the ethos would allow Alicia to accept extra compensation. After all, the internally assessed ethos only requires that Alicia not accept more than the level of pay that makes her equal, internally assessed. Yet, the extra compensation violates externally assessed equality. As a result, even perfect compliance with the internally assessed ethos will not secure externally assessed equality. The internally assessed ethos, then, is too weak to escape the trilemma if equality is externally assessed.

On the other hand, however, if the ethos is externally assessed, the ethos will require Alicia to take the job without extra compensation but, as a result, the ethos is now (in at least some cases) implausibly demanding. To illustrate just how demanding and implausible the externally assessed ethos could be, note that all of the following could be true: (i) everyone in society except Alicia is extremely well off, internally assessed; (ii) Alicia, on the other hand, is quite badly off, internally assessed, through no fault or choice of her own; (iii) the externally assessed benefit the rest of society would receive from Alicia making the sacrifice is quite small; (iv) the rest of her society might not benefit at all – and might even be worse off – internally assessed<sup>34</sup>; (v) Alicia has significantly fewer job options than everybody else; and (vi) if Alicia accepts the job she disprefers without extra compensation, she would be *much* worse off, internally assessed, than everyone else.

Of course, proponents of the externally assessed ethos could make Alicia's situation less intolerable by embracing (as Cohen does) a personal prerogative exemption, which allows the individual to give some extra weight to her own self-interest (2008: 11). Nonetheless, it is not clear that the prerogative (either internally or externally assessed) can save the externally assessed ethos from implausibility. After all, proponents of the ethos must still establish that egalitarians would endorse the externally assessed ethos in the first place. But would egalitarians in fact endorse such a demanding ethos? Cohen, for one, seems to have had the internally assessed ethos in mind.<sup>35</sup> Yet, as we already saw, the internally assessed ethos is too weak to secure externally assessed equality. Thus, if equality is externally assessed, Cohen must either commit us to a much more demanding ethos than he recognized or allow that even full compliance with the ethos will not secure all three of equality, freedom of occupational choice, and Pareto.<sup>36</sup>

E. There are two distinct reasons why the externally assessed ethos might seem implausibly demanding in the example above, and the distinction between them is important since one but not both can be circumvented. According to the first reason – raised by Jonathan Quong (2007: 85–87)<sup>37</sup> – the ethos is implausibly demanding because Alicia

must make herself *intrapersonally* worse off, internally assessed. According to the second reason, the ethos is implausibly demanding because Alicia, who is already the *interpersonally* worst off individual, internally assessed, must make herself even worse off, internally assessed.

As we will see, the second reason can be circumvented; the first reason cannot. But does the first reason actually make the ethos implausibly demanding? Cohen seems to acknowledge that the doctor-gardener must make herself intrapersonally worse off, internally assessed.<sup>38</sup> Yet, he was willing to accept that result – and for good reason. After all, even though the doctor-gardener must make herself intrapersonally worse off, internally assessed, she is still better-off, internally assessed, than everybody else. She should thus count her blessings for her past and current good fortune instead of complaining that she is not *as much better-off* interpersonally as she was before.

Alicia's situation, however, is different: unlike the doctor-gardener, she is *interpersonally* worse off, internally assessed and yet is required to make herself even worse off, internally assessed. And this, I claim, does make the externally assessed ethos implausibly demanding. The doctor-gardener avoids this particular problem since Cohen stipulates that the doctor-gardener is better-off than everybody else in either job, internally assessed. Yet, Cohen needs an ethos that can apply to all situations (or he must give up the claim that the ethos solves the trilemma). And, as applied to Alicia, the ethos is implausibly demanding.

It is, however, possible to avoid this particular objection. According to a third possible interpretation of the equality criterion, equality requires, as a necessary condition, that no one is better-off under *both* the internal and external assessment. This dual assessment allows us to distinguish the two examples. The doctor-gardener, because she is equal, externally assessed, and better-off, internally assessed, cannot accept additional compensation without violating equality; the additional compensation would make her better-off with respect to both. Alicia, in contrast, although equal, externally assessed, is worse off, internally assessed. Thus, extra pay to Alicia (up until she is equal, internally assessed) is consistent with the dual-assessed equality requirement.

The internally assessed ethos is now adequate to ensure that no one is better-off in both respects.<sup>39</sup> Although Quong's objection remains – the doctor-gardener must still make herself intrapersonally worse off – this is not implausibly demanding, given that she is still better-off, internally assessed, than everybody else. Indeed, the only reason she must make herself intrapersonally worse off is because she was much better-off than everyone else in the original scenario.

Of course, this solution to the trilemma depends on a highly idiosyncratic assessment of equality that would itself need to be established.<sup>40</sup> This, however, further illustrates my main claim: until the terms have been specified, we cannot have a meaningful debate.

### III

A. Let me now turn to the problem for proponents of incentives. What justifies incentives, according to the Pareto argument, is that they make everybody better-off. But better-off *how*?

Before addressing this question, let me first make a disclaimer. I am interested in the Pareto argument for inequalities, which is importantly different from Rawls's argument for the difference principle. The difference principle, under at least some of its formulations, permits inequalities that make some individuals intrapersonally worse off – externally and internally assessed – so long as the worst off group in this state of affairs is better-off than the worst off group in any other state of affairs. For this reason (and others), my argument should not be construed as applying to the difference principle.

Nonetheless, for purposes of determining how Pareto is to be assessed, it is worth noting that, although his terminology is different, Rawls is very clear that the difference principle is to be externally assessed.<sup>41</sup> Brian Barry, in his reconstruction of the argument in terms of Pareto, follows Rawls's lead.<sup>42</sup> Yet, Cohen's restatement (2008: 89) of Barry's reconstruction seems to slip into the internal assessment: 'it is irrational to insist on equality when it is a Pareto-inferior state of affairs (why would anyone, and, in particular, the worst off, prefer equality to an inequality in which everyone is better off?)'.<sup>43</sup> This characterization of the Pareto argument makes sense *only if* Pareto is internally assessed. After all, it is not irrational to reject an externally assessed improvement that makes one worse off, internally assessed.<sup>44</sup>

My suspicion is that this equivocation, if it is equivocation, arose because, in the implicit story of how incentives work, the beneficiaries are better-off in both respects. Agents, as a result of incentives, choose jobs that are more productive. Beneficiaries, in turn, each receive a small bonus (for example, extra income), but – and here is the crucial assumption – otherwise remain exactly the same. The beneficiaries are better-off externally assessed: after all, everything is the same but for the fact that they now have additional income. Beneficiaries are also better-off internally assessed: the only thing that changes is that everyone is a little bit wealthier and (under plausible assumptions) everyone would welcome that.

The problem with this story, however, is that proponents of incentives cannot assume that everything else remains the same. Consider, for example, an extension of the scenario with which I started this essay. In that example, Abigail was running a catering business but could make everybody better-off by becoming a business consultant. Suppose that, prior to Abigail's career change, Boris was one of Abigail's employees. After Abigail abandons her catering business, Boris must find a new job. Since his new job is just as good, externally assessed, once we factor in the additional income he receives as a result of Abigail's career change, Boris is better-off, externally assessed. Nonetheless, if Boris preferred his job in the catering industry and the additional income from Abigail's career change does not fully compensate for this loss, then paying Abigail an incentive can make Boris better-off, externally assessed, but worse off, internally assessed.

Nor is Boris unique. In Cohen's example of the doctor-gardener, each of the following scenarios is possible: The community is big enough to support only one health-care provider and they prefer a doctor to a nurse practitioner. As a result, the former nurse practitioner must find a new job. The plant nursery that supplied the gardener goes out of business; those workers must also find new employment. The newly made doctor cures the delivery driver's chronic back pain and, as a result, the driver's temporary



replacement must return to her previous job. Even though, by hypothesis, each is better-off, externally assessed, it is entirely possible that, as a result of the (presumably unwelcome) job change, each is worse off, internally assessed.

My suspicion is that such cases (in which an individual is better-off, externally assessed, but worse off, internally assessed) are not merely possible, but likely. Indeed, it would be a remarkable economy in which a change significant enough to make everybody better-off, externally assessed, had no effect at all on the types of jobs (or products and services) that exist or on who holds them.<sup>45</sup> And it would be equally remarkable if each and every one of these changes left no one worse off, internally assessed.

In response to this possibility, proponents of incentives have three options: First, they could ignore the plight of Boris and other similarly situated beneficiaries. After all, Boris is still better-off, externally assessed, regardless of how he fares internally assessed. Second, they could endorse the internal assessment of Pareto and take back Abigail's incentive since not everybody is better-off, internally assessed. Or third, they could endorse the internal assessment of Pareto but, instead of taking back Abigail's incentive, give Boris enough additional compensation to make him better-off, internally assessed (or at least no worse off). I address each of the possibilities below.

B. The plausibility of the first option – maintaining externally assessed Pareto – depends, in part, on why proponents of incentives opted for an external assessment of equality in the first place. As discussed in Part I.C, one possible reason for endorsing an external assessment appeals to the paternalistic idea that a properly designed external assessment is a better guide to well-being than individual preferences. If this is the reason for endorsing an external assessment, it seems plausible to ignore the fact that Boris *considers* himself worse off. After all, the very reason for endorsing this particular external assessment was because individuals are thought to be poor judges of their own well-being. Recall, however, that if an external assessment is endorsed for paternalistic reasons, then the incentives/ethos debate significantly changes.<sup>46</sup> Moreover, the assumption that Abigail would be better-off (according to her fully informed and rational preferences) even without the extra pay undermines the incentives argument by making the ethos more plausible. After all, we are not asking Abigail to sacrifice her own self-interests; we are instead asking her to do what is, in fact, consistent with her (fully informed and rational) self-interest.

Of course, there are also nonpaternalistic reasons for endorsing an external assessment. And, under these other reasons for endorsing an external assessment – say, a concern for publicity – ignoring Boris's plight is less plausible. Once we drop the assumption that we are the best judges of Boris's own interests, we are left in the awkward position of explaining to Boris that the inequality that makes him worse off, internally assessed, is nonetheless being done *for his sake*. Yet, surely we do not expect him to be grateful to us for making him worse off, internally assessed! So understood, Pareto would not be as difficult to reject as both sides of the incentives debate seem to assume. Boris would not be irrational to reject an externally assessed benefit that makes him worse off, internally assessed. And, from Boris's perspective, the internal assessment is the assessment that really matters. Maintaining externally

assessed Pareto, then, either fundamentally changes the debate or makes Pareto a less appealing desideratum.

Let us then consider internally assessed Pareto. The second option, recall, takes back Abigail's incentive, while the third option compensates both Abigail and Boris. The second option strikes me as implausible when the third option is available. After all, the idea behind the Pareto argument for incentives is that inequalities are justified whenever they make everyone better-off. If, in the scenario we are imagining, the only thing standing in our way of making everyone better-off, internally assessed, is Boris's plight, then surely (for those who endorse the Pareto argument for incentives) we should not withdraw the incentive from Abigail but should instead compensate Boris as well. Indeed, taking back the incentive violates our Pareto desideratum: there is another state of affairs in which everybody would be better-off, internally assessed.

Endorsing the third option, however, dramatically changes the debate in a different way. To see this, consider the following: If we endorse the third option, we must give Boris (as well as Abigail) extra compensation, thereby creating another inequality, externally assessed. The justification for this inequality, however, is not the incentives justification. After all, there is nothing we are trying to get Boris to do: unlike Abigail, Boris has no choice about whether to change his career. We might instead say that the new inequality is justified because it is necessary to ensure that no one (Boris included) is intrapersonally worse off, internally assessed. But – and here is the important bit – once we have that justification, we can dispense with the incentives justification altogether because whatever justifies Boris's inequality will presumably justify Abigail's inequality as well. The additional pay to Abigail is necessary to ensure that no one (Abigail included) is intrapersonally worse off, internally assessed. Since we need this more general justification – otherwise we cannot justify Boris's inequality – the incentives justification is extraneous.

If this is right, then the debate is not about the permissibility of *incentives*. It is instead a debate about whether it is permissible to give anyone, agents or beneficiaries, extra pay – that is, pay beyond externally assessed equality – to ensure that they are not intrapersonally worse off, internally assessed. This is not (necessarily) a problem for those who endorse extra pay. To the contrary, if opponents of incentives assumed (incorrectly) that agents were already better-off and thus that incentives exacerbate existing inequalities, then the fact that extra pay is given to both agents and beneficiaries might ease their opposition. Nonetheless, the fact that the justification for the extra pay is independent of incentives is a significant change in the so-called incentives debate.

C. The shift from justifying the inequality as an incentive to ensuring that no one – agent or beneficiary – is intrapersonally worse off raises the possibility of new equalisanda that are, strictly speaking, externally assessed, but which nonetheless permit extra compensation consistent with equality. Let me introduce two distinct possibilities.

To illustrate the first: suppose the only reason our imagined egalitarian rejects the internal assessment – and thus does not defer entirely to individuals' preferences – is because she believes individuals should be held responsible for their choices. For example, suppose that, as a result of goofing off during vocational school in his 20s, Billy ends

up in a job he disprefers in his 30s. If he hadn't goofed off, Billy would have been eligible for a job he finds more appealing and, in that job, he would have been equal, internally assessed, to everyone else (or to a common reference bundle). Since Billy is responsible for the fact that he is interpersonally worse off, internally assessed, our imagined egalitarian would balk at paying him more in his current job. Suppose, however, that Billy now falls prey to misfortune that was not his fault: as a result of company-wide layoffs, he loses his job. Although he finds employment in the next best option available to him, he likes this new job even less. Our imagined egalitarian would continue to balk at paying Billy the extra compensation necessary to make him equal, internally assessed, to everyone else (or to a common reference bundle). Nonetheless, since, by hypothesis, Billy is not responsible for losing his job, our imagined egalitarian would compensate Billy for the internally assessed losses associated with his job change. Indeed, unless Billy is paid the difference between his internal assessment of his original job and his internal assessment of his next best job, he would be worse off, internally assessed, through no fault or choice of his own.

For a similar reason, our imagined egalitarian would also pay Abigail the additional amount necessary to ensure that she is not intrapersonally worse off. After all, she is presumably not responsible for the fact that a different career would make everyone else better-off. And the same is true of Boris, her former employee: He too is not responsible for the fact that he must take a new job. Instead of violating equality, the extra income they receive preserves the following equalisandum: Each individual is equal, internally assessed, unless she is responsible for being worse off. And, since the compensation necessary to ensure that the individual is not intrapersonally worse off is compatible with our equalisandum, there is no trilemma (except in those rare cases in which Abigail is responsible for the fact that she is worse off in the job that maximally benefits her community).<sup>47</sup>

The equalisandum described above is subject to two problems. First, its appeal is limited to responsibility-sensitive welfare egalitarians. And, second, as we saw in Part II.C, it is also vulnerable to the objection that, achieving internally assessed equality – whether or not modified for responsibility – might itself violate freedom of occupational choice. In response, let me introduce a second possible equalisandum.

Imagine an egalitarian who rejects internally assessed equality because she thinks interpersonal comparisons of utility are not possible and any choice of reference bundle is necessarily arbitrary.<sup>48</sup> Taking into account those aspects of the job that are widely preferred (or dispreferred), our new egalitarian determines what each job should pay in order to make it equal externally-assessed.<sup>49</sup> Each individual is then allowed to choose from the externally assessed equal options available to her. But – and this is what makes our equalisandum compatible with internally assessed Pareto – if an individual is asked or required to choose an option other than her first choice, she is given the additional compensation necessary to make her indifferent. The new equalisandum, then, is this: each individual receives the internally assessed equivalent of her first choice from among the externally assessed equal options originally available to her.<sup>50</sup>

Someone might object that eliminating the common reference bundle makes the new equalisandum implausible since there is no common reference point – nothing, that is, making everyone equal. Yet, whereas the common reference bundle approach made each

individual extremely vulnerable to the choice of reference bundle, the new equalisandum alleviates that problem. If each individual has, say, five externally assessed equal options from which to choose, she is less vulnerable to the random luck of the draw: she has five chances, instead of just one, of having an option she likes. And, since everyone begins with an externally assessed equal option, there is, in fact, a common tie.

Our skeptic might now object that some individuals have more options than others, and this makes the equalisandum unfair. There are, however, several possible responses. First, recall that each option is externally assessed equal. The unfairness then – if it is unfairness – is nothing like the inequalities that exist in our world in which some individuals have more externally assessed *unequal* options than others. Second, the original distribution (before the additional compensation) is also vulnerable to the same charge. Thus, the objection is not specific to the additional compensation, but rather would presumably affect any attempt to ensure externally assessed equal options consistent with freedom of occupational choice. And, in any case, there are ways we could potentially address the disparity. For example, if it is consistent with freedom of occupational choice, we might tax individuals at a different rate, based on the number of options available to them. Or, as hinted at above, we might restrict each individual to a choice from among, say, five externally assessed equal options. And the state could – and should – invest in ensuring that each individual has a wide range of equally good options.

Moreover, although it might seem natural to assume that individuals with more options are better-off, this need not be true. If I prefer option A, the fact that I also have option B does not necessarily make me any better-off, intrapersonally. Similarly, if you and I both prefer option A, the fact that I also have option B and you do not does not necessarily make me any better-off, interpersonally. Indeed, even if you preferred option B, we could not conclude that I am better-off interpersonally, internally assessed. After all, by hypothesis, interpersonal comparisons of utility are impossible.

In addition, it need not be the case that those individuals who receive compensation are the same individuals who possess more options. To the contrary, as illustrated by the example of Alicia in Part II.D, agents do not necessarily have more options; nor are agents necessarily more talented than beneficiaries. Moreover, even if agents did tend to have more options than beneficiaries, recall that beneficiaries also receive compensation if, like Boris, they are required to change jobs. And, for each agent who receives compensation, there might be several beneficiaries who receive compensation. Thus, the additional compensation might actually alleviate the unfairness, if it is unfairness.

Finally, it also is not an objection that some individuals are paid more than others, despite doing the same work. But see Van Parijs (1993: 317–318). After all, the individuals are paid the same for *the work*. The additional pay compensates for something else: Namely, the fact that one individual, but not the other, had to give up her preferred career.

To be sure, much more would need to be said to establish that such an equalisandum is plausible. Nonetheless, the possibility of such an equalisandum reinforces my main claim: we cannot have the incentives debate until we have specified the terms. If we endorse the equalisandum described above, then there is no trilemma.

## Conclusion

The incentives debate turns on how the three desiderata are defined. First, if equality is externally assessed, then the ethos is not a viable solution. It is either too weak or else implausibly demanding. This is particularly noteworthy given that the debate originally arose in the context of externally assessed equality. Second, if equality is internally assessed and freedom of occupational choice is compatible with achieving internally assessed equality, then we can avoid the trilemma altogether; the ethos is not needed. If, on the other hand, freedom of occupational choice is not compatible with achieving internally assessed equality, the ethos is still needed, but freedom of occupational choice is now less compelling as an inviolable desideratum. Third, if equality is *both* internally and externally assessed (such that equality is violated only when the individual is better-off in both respects), then the ethos might be necessary, adequate, and plausible. Yet, the ethos solution now depends on whether we can justify this rather idiosyncratic assessment of equality.

The ethos, however, is not the only solution at stake. The Pareto argument for inequalities is also vulnerable. If Pareto is externally assessed, then Pareto is now less compelling as an inviolable desideratum. On the other hand, if Pareto is internally assessed, then the debate is no longer about incentives, but rather about ensuring that no one is intrapersonally worse off. And, depending on our equalisandum, it might (once again) be possible to avoid the trilemma altogether.

The upshot is that, until we have nailed down what equality, Pareto, and freedom of occupational choice require, the incentives debate is just a word game. We can render the ethos implausible by adding two words. We can undermine Pareto by specifying a particular assessment. Indeed, we can eliminate the trilemma altogether with the stroke of a pen. Yet, if we are merely playing a word game, the question of whether equality, Pareto, and freedom are compatible is uninteresting: we can make them compatible or incompatible simply by changing the terms.

We are then at an impasse: no one can win a debate in which each side is using different and unspecified terms. Thus, instead of persisting in this hopelessly tangled word game, I propose a truce. Let's set the trilemma debate aside and focus our attention on what really matters: determining what equality, Pareto, and freedom of occupational choice require in the first place.

## Acknowledgements

This article has greatly benefitted from the criticisms and comments of the editors of this journal and two anonymous reviewers. I am also exceptionally grateful to Mark Berger, Brookes Brown, Adam Hosein, Frances Kamm, Jacob Kesinger, Julie Rose, Lucas Stanczyk, Tim Scanlon, and audience members at the University of Chicago's Pozen Family Center for Human Rights. This article is loosely based on a chapter of my dissertation (2010b).

## Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

## Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

## Notes

1. For now I disregard the complicating factor of differential labor burdens.
2. This example is modeled off of a similar one in Cohen (2008: 184–185). Yet, whereas Cohen specifies in his example that the individual has more job satisfaction than others regardless of which option she takes, my example is intended to be neutral; we do not know how her job satisfaction compares to others. I return to this difference in Part II.E.
3. Cohen (2008: 184). The Pareto desideratum I am using here says that we should not maintain one state of affairs if there is another state of affairs in which everybody would be better-off. As I explain below, this is known as weak Pareto.
4. The idea of using an ethos – or moral motivation, more generally – to secure all three of equality, freedom of occupational choice, and efficiency is by no means unique to Cohen. See, for example, Carens (1981), Wilkinson (2000), and Vandembroucke (2001).
5. To be sure, it might be the case that the specific debate between Rawls and Cohen can be resolved by appealing to other Rawlsian commitments. These resolutions, however, are unlikely to satisfy non-Rawlsians. I am concerned here with the ethos and incentives solutions more generally.
6. See, for example, Quong (2007). Quong, however, does not apply the distinction to all three (equality, the ethos, and Pareto).
7. One possibility, of course, is that the current inequality is the result of a previous incentive. Nonetheless, in order to address the simplest case, I will assume a background condition of equality.
8. To the contrary, Cohen (2008: 368) explicitly states that the individual in his doctor-gardener example knows that society would be better-off if she chose to doctor.
9. Of course, an individual who would otherwise happily pursue the option that maximally benefits society for equal levels of pay might nonetheless refuse to do so if she believes that by refusing to do so she can attain a higher level of pay. Cohen, however, sets such cases – what he calls the bluff case or the bad case – aside (2008: 181). Thus, like Cohen, I will also set them aside. My concern in this essay is the misalignment between the individual's occupational preference and the option that maximally benefits her community, not the individual's strategic refusal to pursue the option she actually prefers.
10. As Jonathan Quong (2010: 327–333) has noted, the ethos, if it is to be a solution to the trilemma, does not merely require the individual to accept equality. The ethos also requires the individual to choose the outcome that maximally benefits her society for egalitarian levels of pay. Quong goes on to argue that, by virtue of going beyond the demands of equality, Cohen's egalitarian ethos is no longer a requirement of justice, nor a plausible requirement of morality. For a related discussion on how to interpret the ethos, see Williams (1998: 235–237). Nonetheless, because I want to focus on a different question, I set these issues aside.
11. Since Cohen refers to all aspects of the individual's quality of work experience in terms of job satisfaction, I do the same. Nonetheless, I assume that Cohen did not intend for job satisfaction (narrowly construed) to be the only consideration. For example, an individual might find a particular job very satisfying, but nonetheless prefer another job if, for example, the second

- job, by virtue of having more flexible hours, allowed the individual more time to spend with his family. I assume that Cohen would permit such trade-offs within the egalitarian reckoning since it is consistent with Cohen's motivation for including burdens in our metric of equality in the first place: Since 'both money and quality of work experience matter to people . . . both should therefore enter the egalitarian reckoning' (2008: 200). It would therefore be odd for Cohen to exclude considerations other than job satisfaction (narrowly construed), given that these other considerations also matter to people and might matter a great deal. In any case, however, regardless of what Cohen intended, the internal standard, as I am defining it, includes all of these considerations.
12. As Quong (2010: 317) and others have noted, the example of doctoring might be illicitly influencing the debate, for example, if we have particular intuitions about the importance of health care.
  13. I do not mean to suggest that there is no answer to this question; my claim is merely that the answer should be given.
  14. Quong (2007: 85–88) makes a similar distinction between subjective and objective burdens. Nonetheless, as discussed below (see especially Part II.E) the conclusions Quong draws are different from mine.
  15. Whether or not individuals' preferences necessarily track utility depends on how we define 'preferences' and 'utility'. For an extensive discussion of the various possible interpretations, see Hausman (2012). I will return to a related issue – whether an individual's preferences necessarily track her best interests – in the text.
  16. The second version is compatible with what economists call egalitarian equivalent. See, for example, Pazner and Schmeidler (1978). The egalitarian equivalent guarantees each individual the compensation necessary to makes her indifferent between her current job–wage bundle and some common reference bundle (for example, strict equality). Van Parijs (1993: 315–318) similarly uses a basic income for no work reference bundle.
  17. See, for example, Fleurbaey (2009: 1054). As Fleurbaey notes, the fact that the choice of reference bundle changes the results does not mean that we must reject the idea of a common reference bundle. Rather, it just means that any such choice must be justified. Fleurbaey is optimistic that we can do so.
  18. 'The difference principle tries to establish objective grounds for interpersonal comparisons . . . . These comparisons are made in terms of expectations of primary social goods'. Rawls (1999: 79).
  19. As Cohen (2008: 368–69) acknowledges, 'how much a given person enjoys, or otherwise finds fulfillment in, or, contrariwise, is burdened by, a given job is not publicly (very) scrutable'. Yet, Cohen goes on to argue that this is not a reason to exclude burdens from the equalisandum.
  20. 'Gardening is much more fulfilling . . . for H than most people's jobs are fulfilling for them, whereas gardening is no more fulfilling for G than most people's jobs are for them' (2008: 368).
  21. The burdensomeness of the number of hours worked 'varies importantly in size with the character of the job *and the make-up of the person*' (2008: 99, note 33, my emphasis).
  22. See also (Cohen 2008: 200):

Just as individuals trade the two desiderata [money and quality of work experience] off against each other in their intrapersonal comparison of the work/money packages

available to them, so egalitarians should bring form and amount of work, as well as income, into their conception of equality . . . .

23. Cohen (2008: 94, note 21) ('Different egalitarians believe that different things should be equalized, so they would endorse and condemn different outcomes, but that does not matter here. . . . [T]he argument we shall examine is for inequality no matter what conception of equality is embraced'.)
24. Indeed, Cohen seems (at times) to have exactly that in mind. See note 22 above. Under this interpretation, it is the modifier 'substantially' in the phrase 'no one is substantially better-off than others are' that makes the equality requirement 'weakish'.
25. Under this interpretation, the equality requirement is 'weakish' in two ways: the modifier 'substantially' and the fact that it merely states a rather minimal necessary condition. I am grateful to the editor for pointing out this interpretation.
26. See notes 30 and 36.
27. Indeed, being a gardener is, for her, 'super-wonderful' (2008: 207–208). Whether her income is the same as everyone else's is actually unclear. Cohen specifies that, at £20,000, she is *not better-off* in income than others (2008: 185). This, of course, is compatible with her income being less than that of others. But Cohen also seems to suggest that, at £20,000, her income is the same as others: 'if in deference to equality . . . we freeze salaries at £20,000 . . .' (2008: 185). In any case, however, the assumption that she receives the same income as others is not necessary for my argument. Even if her income is worse than others, she is still, by Cohen's own stipulation, better-off than others on the internal standard (2008: 185).
28. Quong (2007: 85–87) has made a similar argument. Our conclusions, however, differ in at least two respects. First, Quong assumes that Cohen would be discomfited by the result that the very same – or at least similar – supplemental income would be endorsed under internally assessed equality as under Rawlsian incentives. That is, Quong takes the extra income itself to be Cohen's target. In contrast, I assume that Cohen was bothered by *inequalities* and, if assured that the extra pay was consistent with internally assessed equality and if assured that the internal assessment was the right way to assess equality, I assume he would drop his complaint. Second, as I discuss in II.C, I believe that we cannot secure internally assessed equality consistent with freedom of occupational choice. Cohen himself suggested this problem (2008, personal communication). And thus I conclude that the ethos is still necessary, albeit to solve a dilemma (as well as a trilemma).  
 Van Parijs (1993: 315–318) has similarly observed that subjective compensation for labor burdens allows us to satisfy equality and Pareto simultaneously. Like Quong, however, Van Parijs does not address the possibility of a dilemma: Can we make individuals' options subjectively equal without violating freedom of occupational choice? If we cannot, then the fact that subjective compensation is compatible with Pareto does not help us escape the trilemma. I turn to this issue in II.C.
29. In order to take into account the fact that she is better-off, overall, even as a doctor, we could reduce her salary as a doctor as well. Since this does not affect my argument and only complicates the exposition, I will set it aside.
30. The second interpretation of Cohen's equality requirement – in which the equality requirement merely states as a necessary condition that no one should be substantially better-off



- with respect to both job satisfaction and income – proceeds slightly differently but it leads to a similar conclusion. Allowing the doctor-gardener to receive £20,000 as a gardener does not violate this interpretation of the equality requirement, since, after all, the doctor-gardener is not better-off with respect to both income and job satisfaction. And this is why the ethos might seem to be necessary. Notice, however, that, although allowing the doctor-gardener to receive £20,000 as a gardener does not violate the necessary condition, reducing the doctor-gardener's salary below £20,000 as a gardener *also* does not violate the necessary condition. (Other individuals, after all, will not be better-off with respect to both job satisfaction and income.) Thus, if we do need the ethos, it is precisely because we *chose* not to reduce her salary, even though doing so was consistent with the equality constraint. The irony, then, is that, if we need an ethos under the internal assessment of equality, it is because and not in spite of the weakness of Cohen's equality requirement. But, in any case, see II.C.
31. Proponents of the first version of the internal assessment – requiring interpersonal comparisons of utility – are especially unlikely to make this argument. After all, in order to implement the ethos, they too must know the internally assessed well-being of others and this would seem to require a similar invasion of privacy, unless it is only objectionable when the *state* invades our privacy.
  32. Vandembroucke (2001: 17, 84–85) discusses a similar distinction, introducing the term 'flexibility evaluations' when all options are considered, as opposed to 'elementary evaluations' (attributed to Amartya Sen) when only the chosen option is considered. Nonetheless, Vandembroucke's use of the distinction is different from the one I am making here.
  33. Although Quong (2007: 86) does not address the possibility of an internally assessed ethos combined with externally assessed equality, he similarly argues that the externally assessed ethos is implausible. This is the second prong of his two prong attack. (See note 32 for his first prong.) Yet, as I discuss in Part II.E, our objections are importantly different.
  34. I return to this issue in Part III.
  35. According to Cohen's description, the ethos only requires individuals to demand not 'much more than what is required for her to be roughly on a par with others, given the satisfactions, and frustrations, that attach, for her, to her job' (2008: 370).
  36. The second interpretation of the equality component – stating, as a necessary condition, that no one is better-off with respect to both income and job satisfaction – fares similarly. If income and job satisfaction are both externally assessed, then the internally assessed ethos will be too weak to ensure either externally assessed equal income or externally assessed equal job satisfaction. And, assuming again that both income and job satisfaction are externally assessed, the externally assessed ethos remains implausibly demanding.
  37. Specifically, according to Quong, if occupational burdens are measured objectively, then a talented individual might be required to make herself intrapersonally worse off. Quong goes on to argue (2007: 86) that this gives us a 'strong reason to reject Cohen's ethos, but not Rawls's difference principle, which is guaranteed always to make everyone at least as well off as they would be under [strict equality]'. Quong presumably here means the difference principle guarantees that everyone is at least as well off, internally assessed (since, if he meant externally assessed, his claim against Cohen would be false). Yet, as we will see in Part III, it is not necessarily true that Rawls's difference principle guarantees that everyone is at least as well off, internally assessed.

38. The fact that she makes herself worse off is compatible with Pareto. The Pareto requirement stipulates only that there not be some other state of affairs in which everybody could be made better-off. We first need to decide, however, whether Pareto is internally or externally assessed. I return to that question in Part III.
39. Alternatively, we could endorse an ethos that reflects our equalisandum: individuals must ensure that they are not better-off with respect to both the internal and external assessments. This too would be necessary, adequate, and arguably plausible.
40. The mixed assessment is perhaps consistent with Cohen's own equalisandum that (arguably) contains both internally assessed and externally assessed elements. See Cohen (1989: 921) (referring to the 'unlovely heterogeneity of the components of the vector of advantage'.) Nonetheless, Cohen does not explicitly argue that equality is satisfied so long as the individual is not better-off with respect to both the internal and external assessment.
41. Rawls (1999: 79). ('The difference principle tries to establish objective grounds for interpersonal comparisons . . . . These comparisons are made in terms of expectations of primary social goods'.)
42. Although Barry recognizes that other possible interpretations exist, he here equates being 'better-off' with 'having more income' (1989: 229).
43. Despite Barry's careful statement of the Pareto argument, the internal assessment of Pareto is apparently so predominant that Thrasher and Hankins, in their critique of Cohen's solution, do not even acknowledge the *possibility* of externally assessed Pareto. Instead, Thrasher and Hankins (2015: 175) equate a state in which 'everyone can be made better-off' with a state that is 'unanimously preferred'. This assumption follows on the internal assessment of Pareto but not on the external assessment.
44. Cohen's interpretation might stem from Rawls's allusion to a veto: 'those who benefit least have, so to speak, a veto'. Rawls (1999: 131). The motivation to exercise the veto is, after all, closely correlated with the internal assessment. An individual who is worse off, externally assessed, but better-off, internally assessed *might* still exercise the veto, but her reason for doing so would be different from the one traditionally assumed.
45. To give just one example: Technological changes that arguably improve everyone's lot, externally assessed, affect both the types of jobs that exist and who holds them. For example, the development of new technologies, such as the washing machine, has been correlated with the decisions of married women to enter the labor market (Greenwood et al., 2005). Yet, as a result of these new entrants to the market, some individuals presumably had to find other employment. Thus, we cannot assume that all of these changes – some voluntary but some presumably involuntary – make everyone better-off, internally assessed.
46. See Part I.C above. Assume Abigail would be paid as a business consultant a wage that makes her equal according to her fully informed and rational preferences. If she is still unwilling to change careers, the reason she needs an incentive (or the ethos) must be because she is not tracking her own best interests.
47. To illustrate: Suppose that Abigail chose to watch *Iron Chef* instead of *The Apprentice*, even though she knew that, in doing so, she was likely to create a preference that would be at odds with the best interests of her community. In this event, our imagined egalitarian would not pay Abigail the extra compensation.
48. But see Fleurbaey (2009: 1054) ('even if the theory of references is still in its infancy, the idea that the choice of the reference is arbitrary is simply unwarranted').

49. Elsewhere I argue that we should set compensation at the level of pay necessary for the market to clear if everyone were equally talented. That is, hold the number and type of jobs constant and imagine an equally talented society with our same preferences bidding on those jobs. The market clearing wages will reflect everyone's preferences, but not the distribution of talent. Each individual is compensated according to the job she actually does. The level of compensation for each job, however, is determined by the market clearing wages in the imaginary equally talented society. See Olson (2010b: 52–69). My argument, here, however, does not depend on endorsement of this particular view.
50. The equalisandum is similar to the egalitarian equivalent (discussed in note 16) but for the fact that there is no common reference bundle.

## References

- Barry B (1989) *Theories of Justice*. Berkeley: University of California Press.
- Carens J (1981) *Equality, Moral Incentives, and the Market: An Essay in Utopian Politico-Economic Theory*. Chicago: University of Chicago Press.
- Cohen GA (1989) On the currency of egalitarian justice. *Ethics* 99: 906–944.
- Cohen GA (2008) *Rescuing Justice and Equality*. Cambridge: Harvard University Press.
- Fleurbaey M (2009) Beyond GDP: the quest for a measure of social welfare. *Journal of Economic Literature* 47: 1029–1075.
- Greenwood J, Seshadri A, and Yorukoglu M (2005) Engines of liberation. *Review of Economic Studies* 72: 109–133.
- Hausman D (2012) *Preference, Value, Choice, and Welfare*. Cambridge: Cambridge University Press.
- Olson KA (2010a) The endowment tax puzzle. *Philosophy and Public Affairs* 38: 240–271.
- Olson KA (2010b) *Justice, unequal talents, and the market*. PhD Thesis, Harvard University, USA.
- Pazner EA and Schmeidler D (1978) Egalitarian equivalent allocations: a new concept of economic equity. *The Quarterly Journal of Economics* 92: 671–687.
- Quong J (2007) Contractualism, reciprocity, and egalitarian justice. *Politics, Philosophy and Economics* 6: 75–105.
- Quong J (2010) Justice beyond equality: review essay. *Social Theory and Practice* 36: 315–340.
- Rawls J (1999) *A Theory of Justice*, 2nd ed. Cambridge: Harvard University Press.
- Thrasher J and Hankins K (2015) Why justice demands inequality. *Journal of Moral Philosophy* 12: 172–194.
- Vandenbroucke F (2001) *Social Justice and Individual Ethics in an Open Society: Equality, Responsibility, and Incentives*. Berlin: Springer-Verlag.
- Van Parijs P (1993) Rawlsian, christians and patriots: maximin justice and individual ethics. *European Journal of Philosophy* 1: 309–342.
- Wilkinson TJ (2000) *Freedom, Efficiency and Equality*. Basingstoke: Macmillan Press.
- Williams A (1998) Incentives, inequality, and publicity. *Philosophy and Public Affairs* 27: 225–247.

## Author biography

**Kristi A Olson** is an assistant professor of philosophy at Bowdoin College.