

The Cooperation Argument for Fairness in International Trade

ISUS X Panel: Welfare Consequentialism and Global Distribution

Abstract. Many moral complaints about the way in which the contemporary multilateral trading regime distributes benefits and burdens across the members of distinct nations are put in terms of fairness. We are told that the current system of international trade not only impacts adversely on the basic welfare of the global poor, but is also *unfair*, in the sense that it arbitrarily skews distributive outcomes in favor of developed countries. The intuitive force of such complaints appears to pose a problem for welfare consequentialists. Not only do claims of fairness have no intrinsic significance within welfare consequentialism, but fairness can conflict with the pursuit of goals that do have clear welfare consequentialist warrant. Some have suggested that the connection between fairness and socially beneficial cooperation provides a way for welfare consequentialists to incorporate concerns of distributive fairness in international trade into their theory. This paper outlines that argument and considers the challenges that it faces.

Helena de Bres

Department of Philosophy

Wellesley College

hdebres@wellesley.edu

The Cooperation Argument for Fairness in International Trade

1.

Many moral complaints about the way in which the contemporary multilateral trading regime distributes benefits and burdens across the members of distinct nations are put in terms of fairness. We are told that the current system of international trade not only impacts adversely on the basic welfare of the global poor, but is also *unfair*, in the sense that it arbitrarily skews distributive outcomes in favor of developed countries. The ethical debate over the WTO's TRIPS agreement provides a case in point.¹ Some criticisms of TRIPS clearly derive from a concern with what we can call "decency": the attainment by individuals of a minimally decent life, in which, at the very least, health care and food security are assured. In the health sector, TRIPS is charged, in this respect, with delaying the entry of generic medicines into developing country markets, keeping their prices out of the reach of the global poor. In the agricultural sector it is said to threaten the rights of poor farmers to save, exchange and sell seeds, thereby rendering access to vital foodstuffs insecure. But other common criticisms of TRIPS point to a distinct set of moral concerns. Given that the huge majority of the intellectual property rendered eligible for international patent protection under TRIPS derives from industrialized countries, the incorporation of TRIPS into the WTO system will result in a significant transfer of resources from the developing to the developed world.² Many have claimed that the standard arguments given for domestic intellectual property protection are insufficiently compelling in the international case to justify this disparity of impact.³ Moreover, the benefits in increased market access that poor countries were promised in exchange for signing the TRIPS agreement have largely failed to materialize. People who criticize TRIPS on the preceding two grounds are moving beyond decency-related concerns, to question, on grounds of equity or fairness, the way in which TRIPS distributes benefits and burdens across countries.

¹ TRIPS (Trade-related Aspects of Intellectual Property) introduces a single global blueprint for the protection of intellectual property rights, involving the extension of minimal regulations concerning patents, copyrights, geographical indicators, and industrial designs across all member countries of the World Trade Organization.

² Chaudhuri, Goldberg and Jia (2004) suggest that foreign producers would gain something like \$53 million per year if India satisfied international patent protection standards – while the Indian economy would lose \$450 million, \$400 million of which would have gone to consumers.

³ See James 2008(a) for a good discussion.

The intuitive force of such complaints appears to pose a problem for welfare consequentialism: the ethical theory according to which all of morality rests exclusively on a single foundational duty to promote wellbeing.⁴ Welfare consequentialists have a reputation for indifference or hostility to fairness, in all of its forms. The reasons underlying this reputation apply as much to the case of distributive fairness in the multilateral trading regime as to anything else. So it is natural to assume that welfare consequentialists will focus any criticisms of the contemporary system of international trade on that system's implications for decency rather than its fairness; and this, in fact, is what they do.⁵

The core commitments of welfare consequentialism have considerable appeal. But the idea that the contemporary multilateral trading regime raises genuine concerns of fairness is also compelling. So it is natural to wonder whether there might not, after all, be a way of consistently holding those two sets of ideas together. The aim of this paper is to describe and assess one such welfare consequentialist argument for fairness in the system of international trade. I begin by outlining the general problem that welfare consequentialists are felt to have with accommodating considerations of fairness, and the range of strategies of response available to them. I then describe the particular argument I am interested in - an argument that draws on the supposed connection between fairness and socially beneficial cooperation - and spend some time discussing its merits and problems. I conclude that the argument under consideration is unlikely to be able to do all of the work it is intended to do, and finish with a brief assessment of the more general prospects of reconciling distributive fairness in

⁴ Slightly more formally: *consequentialism* is the thesis that the moral status of actions, rules, institutions or other entities is exclusively a function of an ordering of the available options in terms of the overall value in which each results; *welfarism* is the thesis that wellbeing (however construed) is the sole item of moral value. Welfare consequentialist theories differ chiefly over their choice of "interpersonal aggregation function": *i.e.* the principle by which the overall value of a state of affairs is determined. Utilitarians opt for sum-ranking, according to which the value of a state of affairs is determined additively, by summing up the value of its components (those being, for utilitarians, the wellbeing of individuals). The result is a theory that requires us to maximize total wellbeing. Other variants require us to maximize average wellbeing, to ensure that as many people as possible attain a certain "sufficient" level of wellbeing, to ensure that the wellbeing of the worst off is given special consideration, and so on. I have tried, for the sake of generality, to structure the discussion below in such a way that it is neutral between each of these variants. The discussion in section 5, however, highlights one respect in which a welfarist's choice of interpersonal aggregation function has distinctive implications for the argument I will consider here.

⁵ Elfstrom, for instance, writes of his book applying utilitarianism to global politics: "In general terms the position of this work is that all have a strong obligation to work to satisfy basic wants wherever they are found, but a much weaker obligation to look after the secondary wants of each and every human being (1990, 15) - and then spends no time at all on the second kind of obligation, which will include concerns of international fairness. Similarly, Goodin (1985) confines his global concerns to "protecting the vulnerable", where the latter are presented as those persons whose basic welfare interests are at risk. Goodin explicitly presents his principle of protecting the vulnerable as a "partial principle of morality" making "no claim to order the entire moral universe" (1985, 117), but it is nonetheless all that he offers.

international trade with welfare consequentialism.

2.

Welfare consequentialism's problem with fairness stems from an apparently fundamental conflict between the structure of the former and the following two core features of the latter.⁶ First, fairness is exclusively concerned with the satisfaction of *claims*. A person (call him Cary) has a claim to a certain good or form of treatment, *x*, if *x* is: i) *owed* to Cary and ii) owed explicitly to Cary *himself*. The implied contrasts here are with *x* as something that it would simply be good to give to Cary and with a reason to give *x* to Cary that derives from a duty of general benefit of which Cary just happens to be the current beneficiary. (One way of capturing the core idea here is to say that fairness ensures that people receive what they are owed, with the emphasis on both the "they" and the "owed"). Second, fairness is not concerned with the satisfaction of claims *per se*, but rather with their *proportionate* satisfaction: with their satisfaction, that is, relative to the satisfaction of other people's claims. For instance, If Cary and Gregory have equal claims to the 1959 Best Actor Oscar, fairness is indifferent between neither getting the Oscar and both getting it. If Cary has a stronger claim to the Oscar than Gregory, on the other hand, fairness requires that Cary receives it, or that neither do.

While it is open to welfare consequentialists to use this language of claims and their proportionate satisfaction, both ideas are awkward within a welfare consequentialist framework. Take the first, first. A welfare consequentialist might say that Cary has a claim to a certain good or form of treatment, *x*, if the act of giving *x* to Cary would produce the best welfare consequences. But this interpretation is undeniably strained. For one, producing the best welfare consequences hardly seems like something that is owed to Cary *himself*. If anything, it is a duty owed to persons at large (a notion that some will doubt makes sense). For another, welfare *per se* does not seem the sort of basis on which claims might arise. While the fact that giving *x* to Cary would affect his or anyone else's welfare in a valuable direction might be a reason to give *x* to Cary, it seems odd to say that Cary is thereby *owed* *x* (whereas saying so would not be odd if desert or entitlement, rather than welfare, were the basis for his claim). The fact that fairness is not directly concerned with

⁶ In singling out these two features, I follow Broome (1991), 195.

the satisfaction of claims, but rather with their proportionate satisfaction, also appears problematic for a welfare consequentialist. The point of requiring satisfaction of claims in proportion to their strength rests on the idea that conflicting claims to a single good will sometimes arise. But note that on the welfare consequentialist interpretation just proposed, there will only ever be *one* claim to *x* (i.e. the claim of whoever's having *x* will produce the best welfare consequences). One might remedy this by suggesting that each person has a claim to every good or form of treatment *x*, and that the proportionate satisfaction of each person's claim to *x* requires that the welfare of each count equally in the process of aggregation that determines who should get *x*. But, aside from the simple implausibility of the idea that everyone has a claim to everything, this suggestion renders the core fairness idea of proportionate satisfaction redundant. The point of requiring satisfaction of claims in proportion to their strength rests not only on the idea that conflicting claims may sometimes arise to a given good (as above), but also on the idea that the strength of those claims will at least sometimes vary. On the interpretation just given, however, they never will.

The preceding points, of course, reflect three quite familiar criticisms of welfare consequentialism: its focus on the collective rather than the individual; its neglect of all moral concerns other than welfare; and its monistic rather than pluralistic character. A fourth frequently criticized feature of welfare consequentialism - its focus on ends rather than means - worsens matters further for consequentialists out to capture our intuitive ideas of fairness. It is key to our concept of fairness that the latter imposes stringent constraints on the legitimate pursuit of the general welfare. Welfare consequentialists, however, cannot countenance such welfare-independent constraints without abandoning their theory. If an act (rule, institution, decision procedure) produces the type of welfare consequences favored by the theory, that act (rule, institution, decision procedure) must be right, whether fair or unfair. The result is a dilemma: either the dictates of fairness conflict with those of welfare consequentialism, in which case they must be rejected, or else they coincide with those of welfare consequentialism, in which case they are redundant.

One welfare consequentialist response to all of this is simply to reject concerns of fairness as misguided. Because intuitions about the importance of fair treatment run deep and strong, however, this is a large bullet to bite. An alternative, more palatable, response attempts to dissolve the conflict between welfare consequentialism and fairness by

showing it to be apparent rather than real. This attempt is a variant of a more general strategy on the part of welfare consequentialists that we can term “assimilationism”. Assimilationism consists in the attempt to convincingly incorporate core features of common-sense morality within an uncompromised welfare consequentialism.⁷ The assimilationist project is a delicate operation, in that it must satisfy two criteria that are in some tension. For one, the attempt at assimilation must provide a sufficiently *robust* case for the targeted feature of common-sense morality. The case must show why the feature in question is morally acceptable, and justify steady and stable compliance with its dictates. For another, however, the attempt at assimilation must not give a stronger case for the feature in question than sustained rational reflection itself supports. It is a core feature of our thought about morality that common-sense morality can itself be wrong. This means that if the case given by the assimilationist over-reaches – brooking no violations of its favored principle, even under exceptional circumstances – it fails by its own lights. What the assimilationist seeks is the happy medium between fully rejecting common-sense morality and fully absolving it of any criticism.

Some (including the bullet-biters mentioned just previously) will not see the point of this project. Assimilationist welfare consequentialists are suspected of wasting their time concocting disingenuous and inevitably unsuccessful arguments aimed at converting the doubtful to their *prima facie* unappealing theory, or, worse, harboring their own hidden non-welfare-consequentialist sympathies, which they are aching to assuage. I am more sympathetic to the project, though as yet agnostic on its prospects of success. Assimilationism is motivated by the idea that an ability to accommodate the core features of common-sense morality (suitably filtered by sustained reflection) is a virtue in a moral theory. Although the general methodological question of how moral intuitions ought to figure into moral theory lies far beyond the scope of this paper, let me just say that I do not find this idea outrageous. Assimilationism seems to me to be a perfectly respectable project, and, insofar as it holds out the *potential* for having one’s cake and eating it too, I am all for it – in the case of principles of fairness, as in others.

⁷ Given that welfare consequentialists are not always explicit on their meta-philosophical commitments, it can be unclear where to place them on the assimilationist-non-assimilationist divide. I take Shelly Kagan to be a prime example of a non-assimilationist bullet-biter. Hooker’s rule utilitarian account in *Ideal Code, Real World* has an assimilationist character, insofar as it uses alignment with common-sense morality as a decider when selecting between pair of otherwise equally acceptable moral codes.

Assimilationist welfare consequentialists such as myself have three broad strategies available to us. One option is to include fairness within the foundational values recognized by one's theory. This strategy comes in two variants. On the first, fairness is claimed to be, alongside welfare, a direct contributor to the value of a state of affairs. We are no longer required simply to produce the theory's favored welfare consequences, that is, but to produce these *and* fairness. This strategy represents a sacrifice in the theoretical simplicity that is one of welfare consequentialism's key attractions. It also has the counter-intuitive implication that, other things being equal, we should not only be fair whenever the occasion arises, but also actively *create* opportunities to be fair (e.g. baking birthday cakes ripe for division) in order to add to the world's fairness count. Someone who treats fairness as the appropriate object of production in this way would appear to have fundamentally misunderstood its point. A similar problem arises in relation to the second variant of the strategy at issue, according to which fairness should be understood as an irreducible constituent of welfare, so that my being fairly treated is, in itself, a component of my life's going well for me. While there are certainly many and ingenious ways to increase a person's welfare, a concerted campaign of acting fairly toward her – *repeatedly* - does not seem to be one of them.

A second option for assimilationists is to appeal to people's preferences for fair treatment (either of themselves or others). Kaplow and Shavell refer, in this fashion, to the "possibility that individuals have a taste for a notion of fairness, just as they might have a taste for art, nature or fine wine" (2002, 21). One common objection to this line of thought is that it is unacceptably deflationary: calling our commitment to fairness a mere "preference" or a "taste" fails to capture its moral importance. This response is unlikely to move many welfare consequentialists. After all, on perhaps the most prevalent version of welfarism, preference-satisfaction is the *only* thing of moral value, so that saying that fairness satisfies preferences is actually paying a significant moral compliment to it. A stronger objection to this strategy is that treating fairness as a simple first-order preference will not generate a sufficiently robust welfare consequentialist argument in favor of fairness. We have preferences for fairness, certainly, but we also have preferences for a great many other things, and there is no systematic reason to think that our preferences for fairness take any special priority.

It seems that what welfare consequentialists need, in order to justify the substantial and constraining role that fairness plays in our moral thought, is some reason to think that

compliance with principles of fairness has valuable welfare consequences beyond people's direct attachment to those principles themselves. The third strategy I will mention, an instance of the general approach known as "indirect consequentialism", attempts to provide this reason. The hallmark of indirect consequentialism is the claim that, although a rigid moral rule (such as a duty to speak the truth) may be superficially in tension with welfare consequentialism, the end result of adopting such a rule, at least over the long term, will be more productive of the welfare consequences valued by the theory than the alternative in which those rules are rejected. This is said to be so for essentially practical reasons – standardly, limits on information, coordination, or altruism, and our capacity for self-deception or error – that limit our ability to discern and carry out optimal courses of action on a case-by-case basis. The third assimilationist strategy is to argue that principles of fairness are among these indirectly justified rules. On this approach, we are to understand the proportionate satisfaction of Cary and Gregory's claims to the 1959 Oscar as a kind of social practice, the internal dynamics of which may be difficult to state in welfare consequentialist terms, but the effects of which over time are entirely consonant with welfare consequentialist priorities. Although indirect consequentialism has its own problems (some of which will come up below), it is the most promising of the options available to welfare consequentialists in this area, and the argument I consider in this paper is an instance of it.

3.

That argument, in its most simple and compressed form, goes like this:

1. *The fairness-cooperation connection.* The multilateral trading regime must satisfy norms of distributive fairness if cooperation within that regime is to be sustained over the long term.
 2. *The cooperation-welfare connection.* Sustained cooperation within the multilateral trading regime will produce better welfare consequences than the non-cooperative alternative.
- Intermediate conclusion.* A multilateral trading regime that satisfies norms of distributive fairness will produce better welfare consequences over the long term than a regime that violates those norms. (1,2)
3. *Welfare consequentialism.* We ought to produce the best welfare consequences

available to us.

Conclusion. Welfare consequentialism requires that the multilateral trading regime be distributively fair. (1,2,3)

Call this “the cooperation argument”. To clarify its terms: “distributive fairness” refers to fairness in the distribution of benefits and burdens (generally, resources) across persons. The specific content of the norms in question can be left open for the moment. So too, in the interest of generality, can the question of which welfare consequences are better than others, given that this depends on which version of welfare consequentialism one endorses. In general, “cooperation” refers to coordinated action between two or more actors in pursuit of some common aim assumed to be of mutual benefit. Here, the actors at issue are participants within the “multilateral trading regime”, and the common aim is to promote liberalization in the cross-border market exchange of goods, capital and (more recently) services. The regime in question is centered in the World Trade Organization (WTO), and incorporates i) a set of rules, policies, agreements and standing expectations regarding tariffs, subsidies, quotas and other types of trade-related action, ii) an ongoing schedule of structured negotiations over the development and elaboration of those rules (etc.), and iii) a body for resolving any related disputes and encouraging compliance. The “non-cooperative alternative” in this case would include such actions as withdrawal of membership in the WTO, or sustained refusal to abide by its rules and guiding principles.

Although I have yet to see a philosopher explicitly addressing this argument, gestures in its direction are fairly common amongst social scientists working on the economics and political economy of international trade. Premise 1, for instance, is suggested in the following passages:

“Unless there is consensus about the fairness of the rules and procedures in the trading system, countries will not abide by them indefinitely. Fairness is therefore an element in the existence and functioning of the system...there may be circumstances in which a country may emerge worse off from a round of negotiations, yet to find it has no choice but to accept the worsened status. Nevertheless, it remains true that, if ideas of fairness are seriously and persistently violated, the cooperation on which the system rests will be threatened and the system thereby undermined.” (Brown and Stern 2006)

“International [trade] agreements require the agreement not just of the rich countries, but

also of the poor, and if the citizens of democratic poorer countries believe that they are being unfairly treated, they will refuse to agree or comply.” (Stiglitz and Charlton 2005, 74)

“The lesson of the recent heroic exercise to gain congressional approval of the Uruguay Round Agreement, including the provisions establishing the WTO, is that any rules-based approach to international trade is unlikely to be durable unless, in the end, it is able to command significant public legitimacy” [which in the context of the paper, implies distributive fairness]. (Howse and Trebilcock 1996)

Endorsements of premise 1 in a more general form (*i.e.* applied to all international regimes, including the trading system) are also present, as in the following two passages:

“The pursuit of a shared conception of fairness is the necessary starting point for devising any lasting allocational rules: rules which are likely to command respect and pull towards voluntary compliance.” (Franck 1995, 13)

“The gains of efficiency that derive from inter-state cooperation...[do] not suffice [to explain the motivation of states to form international organizations and abide by their rules] ...decisions arrived at in international forums must also be seen to be fair by their stakeholders. The key to the acceptability of international agreements lies in this perception...” (Narlikar 2006)⁸

Endorsements of premise 2, by these and other authors, are ubiquitous. Premise 3 is less often explicitly stated, although it is taken as a given by most economists. Brown and Stern are the most explicit, of those cited above, about the indirect consequentialist nature of their argument:

It should be understood that we do not advance [fairness] as a high moral principle. It is an

⁸ As these more expansive examples suggest, the cooperation argument is clearly of the sort that might generalize, so that fairness might be said to promote socially beneficial cooperation within all international regimes (or, indeed, within any form of relationship). I do not take on this broader form of argument here, due to its formidable scope. It seems plausible enough that if the cooperation argument works for trade, it will cross over, suitably adjusted, to other international regimes, *e.g.* those concerning human rights, nuclear non-proliferation, or the environment. (For an endorsement in the case of the later, see Franck: “Obviously, if [the] costs of climate stabilization were to be incurred they would have to be distributed in a manner generally perceived as fair by the international community, otherwise the necessary co-operation would not be forthcoming” (1995, 389); see also the similar argument in Athanasious and Baer, 2002). But since there is no principled reason to believe that in advance, this is a case that needs to be made regime by regime (or relationship by relationship).

instrumental criterion to be valued for its consequences, namely that it facilitates the reaching of inter-governmental agreements that protect and enhance the mutually advantageous trading system.⁹ (2006)

4.

One of the problems with indirect consequentialism, here as in general, is that its arguments come across as too good to be true. The apparent conflict between welfare consequentialism and fairness is something of a major public relations problem for welfare consequentialists, and it would be very good for them if that conflict could be dissolved. But, as in the case of claims by the similarly downtrodden that respecting basic labor standards is always good for business, or that liver is not just nutritious but *also* tastes great, wishing something does not make it so. That said, wishing something does not *not* make it so, either. So the preferable route is to ignore the aura of convenience and focus instead on the content of the argument. For present purposes, I will take premise 3 as given, and concentrate instead first on premise 1 (the “fairness-cooperation connection”), and then more briefly on premise 2 (the “cooperation-welfare connection”).

The general idea expressed in premise 1 - that distributive fairness is necessary for sustained cooperation – is intuitively appealing. Both casual empiricism and a now extensive body of research in behavioral economics and social psychology suggest that people who perceive themselves or others to be getting a less than fair share of the gains of a cooperative enterprise are prone to defect or become hostile, even when doing so comes at a personal cost.¹⁰ This behavior is partly explicable in terms of the direct emotional-motivational force that moral norms in general can exert once internalized. If I am strongly committed to principles of fairness, I will generally feel indignant, enraged or insulted when I see myself or others being treated in ways that

⁹ Similarly see Franck’s more general (and on the face of it quite implausible) claim that “the answer to [the] question [‘What shall we do about sharing and conserving in order to maximize human well-being?’] or rather the starting point in the search for an answer, is misleadingly simple: ‘we must do that which is fair’.”

¹⁰ Okun (1981) offers a detailed account of the demands of customers and employees for fair treatment and of the role of perceived unfairness in triggering a search for alternative suppliers. Fehr and Gächter (2000) present the results of more recent experiments suggesting that people are willing to hold back their contributions to a public good when others fail to pay their fair share, even when doing so goes against their self-interest. See also Kahneman, Knetsch and Thaler (1986), and, more generally, the two collections Mellers and Baron (1993) and Messick and Cook (1983).

conflict with those principles, and this may trigger a desire to withdraw membership, withhold future contributions, or punish those perceived to be responsible. This everyday feature of human moral psychology supports a causal connection between fairness and cooperation (or between unfairness and defection) in cooperative environments where norms of fairness are strongly internalized by participants. However, we can also point to additional factors that might establish a positive causal connection between fairness and cooperation even in cases where participants lack a direct personal attachment to fair treatment. The best way to introduce these factors is by means of the following puzzle.

In some cases, cooperation arises and is sustained naturally and easily. In other cases, however - particularly in instances where it is expected to extend over a lengthy period, and to impact substantially on participants' interests - cooperation is difficult both to get going to begin with, and to maintain once commenced. This can seem odd, given that cooperation, by definition, is assumed to be in every participant's interest. We can explain the puzzle by reference to the following two facts. First, even if agreement is reached on the desirability of cooperation itself, there remains plenty of scope for disagreement over how the surplus value generated by that cooperation is to be distributed across participants. We can assume that each party would prefer, other things being equal, to receive more rather than less of the gains from cooperation, and that each will characteristically attempt to hold out until a favorable bargain is struck. If an acceptable solution to this "division problem" cannot be found, arrival at a mutually advantageous arrangement will be impossible (Coleman 1992). The threat to cooperation here posed is increased by the fact that, in the case of cooperative ventures that extend over a considerable period, the division problem may need to be solved not just once (*i.e.* at the outset of the venture) but repeatedly, as the venture evolves over time.

The second reason why cooperation can be difficult to initiate and sustain derives from uncertainty on the part of potential cooperators concerning the exact consequences of participation. Cooperation essentially involves agreement by participants to impose certain constraints on themselves in order to further the common goal. Typically, cooperators forgo various opportunities for personal advantage that they would otherwise have chosen to exploit. Often these constraints are only in the interest of their adopters on the conditions that a) fellow cooperators adopt similar constraints, and b) adopters'

circumstances do not significantly change. Adopting such constraints is therefore *risky*, because it leaves cooperators vulnerable to the failure of these conditions: either to exploitation by other parties in the scheme or to unforeseen developments that render previously acceptable constraints disadvantageous. The risk at issue derives from the fact that potential cooperators cannot know in advance whether their partners can be trusted to do their part, or how their own position will change in the future, and thus applies even to those who have the power to press for favorable terms at cooperation's outset. In the worst case, exploitation or changed circumstances may result in some cooperators becoming worse-off than they would have been in the absence of cooperation. If the risk of this outcome is felt to be too high, cooperation will stall.

As several theorists have pointed out, norms of fairness are well suited to respond to each of the foregoing obstacles.¹¹ Where the division problem is concerned, generally accepted principles of fairness can serve as a "focal point" in negotiations amongst potential cooperators over the division of cooperative gains.¹² The use of fairness as a default solution reduces or eliminates the kind of time-consuming and costly bargaining over distributive matters that can, at best, cut into the surplus value of cooperation, and, at worst, result in complete stalemate.¹³ Where the problem of uncertainty is concerned, a shared commitment to fairness can operate as an insurance mechanism, assuring all potential cooperators, first, that sacrifices made on their part will not be exploited by others, and second, that, in the event that their circumstances shift unfavorably, they will nonetheless retain some downside protection.¹⁴ The latter implication follows from the

¹¹ See, e.g., Wertheimer (1996), Young (1999), Messick (1993).

¹² The United Nations scale of assessments, which is based on the fairness-related principle of capacity to pay, has come to play something like this role in international relations. Originally intended to apply to the financing of the U.N. as a whole, it was later applied directly to the WHO and other U.N. specialized agencies, as well as to the financing of the Montreal Protocol on ozone protection (see Barrett, 2007, 114). U.N. assessments are based loosely on a formula that reflects both population and income per capita, subject to a ceiling and a floor.

¹³ One might accept that fairness provides a solution to the division problem, but question its unique suitability for the role. Might not other default norms function just as well for these purposes? One virtue possessed by fairness in particular is its frequent (though not essential) connection to equality. The latter, as Messick (1993) argues, has its own heuristic virtues, chiefly in the form of simplicity and symmetry (*i.e.* the absence of a need to identify the particular characteristics of the parties involved).

¹⁴ See the economist William Baumol: "A rule of fairness is a sort of insurance arrangement that selfish people accept to make sure they are not mistreated, and pay for it by providing assurance to others that they, too, will not be mistreated" (1986, 17, fn 7), and the philosopher (and welfare consequentialist) Richard Arneson: "The perception that others are trustworthy in the sense that they are likely to reciprocate benefits received is a significant influence on people's willingness to cooperate with others especially in ways that involve acceptance of costs that benefit others now in the hope that later those whom one has helped will be willing to

fact that fairness generally ensures that, however matters turn out, every participant will gain at least some portion of the benefits of cooperation.

I take it that it is arguments such as these that underlie the surface appeal of premise 1. Some combination, that is, of 1) an intrinsic attachment to fairness, 2) the utility of fairness as a focal point in cooperative bargaining, and 3) the appeal of fairness as an insurance mechanism, is felt to establish a positive causal connection between fairness and cooperation in general, and that generalization is then applied to the case of the contemporary multilateral trading regime.¹⁵ Such an inference is, however, insecure. Given that the arguments rehearsed above draw on a combination of anecdotal evidence, small-scale laboratory experiments, and (not to denigrate it) armchair theorizing, the usual caveats apply about extending their conclusions to larger-scale settings, and to the real world. How applicable are the above considerations, in fact, to the specific case of distributive fairness in the highly complex system of contemporary international trade?

Here advocates of premise 1 face the following problem. The defenders of the cooperation argument cited earlier apparently believe both that the current multilateral trading regime is unfair, and that increasing recognition of its unfairness has contributed to a reduction in popular support for trade liberalization. If premise 1 is correct, what we are seeing in this dissatisfaction are the ripples before a storm. While the international trading system remains roughly stable for the moment, eventually the unfairness built into it will undermine the interstate cooperation necessary to sustain it. The truth of this prediction is, of course, unknowable until we reach whichever date “eventually” is intended to

accept costs when doing so efficiently secures pay-back to the initial helper.” (2007). See also Goodin (1992).

¹⁵ Young (1999) proposes both arguments 2) and 3) in relation to the role of fairness in the formation of international regimes in general:

Institutional bargaining can and often does become a protracted and costly affair...And there is no guarantee that the resources expended on regime formation will yield any clear-cut benefits at all...there is often good reason to pay attention to considerations of fairness or equity as a means of avoiding protracted disagreements or even stalemate.... (250)

Institutional bargaining differs from ordinary, everyday bargaining in ways that make it difficult for participants to anticipate with an precision how the outcomes will affect their interests over time...Many regimes are just too complex or multifaceted...In other cases, regimes turn out to harbor unforeseen loopholes...Beyond this, there are cases in which institutional arrangements themselves change or evolve in a manner that is unpredictable at the outset...These conditions give participants in institutional bargaining a distinct incentive to devise arrangements that conform to well-defined or widely accepted criteria of fairness. (252-3)

refer to. But one would hope for at least some significant supporting evidence for it, drawn either from the present state of multilateral trade or from broadly comparable political-economic regimes of the past or present. And such evidence is weak. The problems are twofold. First, it is far from obvious that the current regime is in the grip of an impending fundamental crisis. The dramatic collapse, in July 2008, of the latest set of Doha round negotiations certainly suggests discontent with some elements of the current system.¹⁶ But it is at this point unclear both whether that collapse spells discontent with multilateral trade *per se*, rather than with some particular aspects of it¹⁷, and whether the current stall is a short- or a long-term phenomenon. The second problem is that, even if we are seeing the early signs of the regime's complete disintegration, it is unclear that unfairness is its cause. Negotiators who participated in the recent talks were inclined to lay the blame on the regime's record in securing decent lives for the global poor, or to a perception on the part of India and China that multilateral trade negotiations were no longer in their national interest.¹⁸ Neither of these factors need have much to do with unfairness, *per se*. Nor do the lessons of history or comparative politics offer much assistance. Many unfair regimes have flourished for decades or centuries, and it is again hardly obvious, in the case of those that do ultimately collapse, that unfairness is the root cause thereof. Faced with these facts, an unkind observer might be tempted to attribute the appeal of premise 1 more to a faith-based conviction in the general unsustainability of unjust institutions than to the deliverances of cool-headed empirical analysis.

How might a defender of the cooperation argument respond to this charge? Let us take the most problematic scenario (for the argument under consideration), and assume that the contemporary multilateral trading regime is in fact in reasonable health. There are two ways to reconcile this fact with premise 1 as it stands, each of which is problematic. The first is to dig in one's heels and claim that, although cooperation within the current regime is not yet in fundamental crisis, this state of affairs is not to last. This assertion might be paired with the claim that there would have been still greater

¹⁶ See "After 7 Years, Talks Collapse on World Trade", *New York Times*, July 30, 2008. Available at: http://www.nytimes.com/2008/07/30/business/worldbusiness/30trade.html?_r=1&hp&oref=slogin

¹⁷ The direct stumbling block in the recent set of talks was the important but hardly comprehensive issue of the WTO's "escape clause" (Article XIX of the GATT), which permits countries to impose temporary trade barriers in order to protect industries that suffer "serious injury" or the threat of it due to "increased quantities" of imports.

¹⁸ See fn 15.

cooperation than we see in the current regime had that regime been fairer. (Similarly, market enthusiasts sometimes imply, China's economy would have recently been growing at *twenty* percent per year, were it not stifled by communism). *Perhaps* this is true. But it is a move unlikely to convince skeptics.

A second attempt at reconciliation would be to claim that the current multilateral trading regime is in fact fair. Perhaps the most common variant of this latter claim comes in two steps.¹⁹ First we are told that all that fairness requires of the multilateral trading regime is that it deliver a pareto improvement over the pre-regime baseline (*i.e.* a state of affairs in which some participants are better off and none are worse off as a result of the existence of the regime). Next it is claimed that the multilateral trading regime satisfies the requirement in question. Both of these claims are highly controversial. To take the first, satisfaction of the pareto-superiority criterion does not in general guarantee anything recognizably like distributive fairness. Instead it can sanction "mutually beneficial exploitation": a situation in which, although all gain from cooperation, the gains nonetheless go disproportionately to a sub-set of participants, for no reason other than the latter's greater bargaining potential. Given that such situations appear to many of us to be paradigm cases of unfairness, the salvaging of premise 1 by this second route will represent small comfort for welfare consequentialists wishing to integrate something close to the everyday understanding of fairness into their theory.²⁰ Second, it is not clear

¹⁹ Srinivasan (2005) provides a good example.

²⁰ It is possible to dispute the unfairness of mutually beneficial exploitation, as defined above, in the following way. (I draw here on Wertheimer 1996). We can understand a party's bargaining potential as that party's "threat advantage": *i.e.* his/her relative willingness not to cooperate if his/her proposal over the division of the cooperative gains is not accepted. A party's threat advantage, in turn, is naturally understood as a function of that party's welfare gain from the pre-cooperative baseline. In short, parties with greater bargaining potential are those who are better able to "hold out", and those who are better able to hold out are those to whom the gains from cooperation are less valuable in welfare terms. An implication of this is that those who are "exploited", by virtue of gaining relatively less objective resources from cooperation, are those for whom the attainment of those objective resources can be expected to matter the most, in welfare terms. This being so, if the exploited had *not* been exploited - if they had, for instance, gained an amount of resources equal to that gained by those with the greater threat advantage - they would have done better, in welfare terms, than the latter. And it can be argued that *that*, rather than "exploitation", would have been the unfair outcome.

The force of this conclusion depends on our construing the welfare gains from cooperation as the appropriate metric for assessing the fairness of cooperation. Although the general point has already been made in the text, is worth reemphasizing that such a construal implies what for many of us would be a radical reinterpretation of what fairness in many situations involves. Take the multilateral trading regime. Due to the decreasing marginal utility of resources, developing countries can be expected to generate greater welfare than can developed countries from a given amount of trade-generated objective resources. If (to take a common view) what fairness requires is some approximation of equality of gains, the result of treating welfare as the metric of fairness will therefore be to sanction a system that, in terms of objective resources, radically favors affluent countries (a system, some would suggest, like ours).

that the contemporary system of international trade has managed even the dubious moral achievement of mutually beneficial exploitation. Whether or not all states have gained on balance from cooperation in the multilateral trading system, it is quite clear that not all individuals have (take, for instance, those who have lost their livelihoods as a result of their country's opening its markets to foreign producers).

In light of these difficulties, defenders of the cooperation argument might contemplate amending premise 1, while retaining its general thrust. In this connection, note that premise 1, as formulated above, suggests that 1) *actual* fairness (rather than the perception of fairness), and, for that matter, 2) *full* fairness rather than partial fairness, is 3) *necessary for* (rather than simply promotes) cooperation within 4) *the system* of multilateral trade (rather than, say, particular agreements within that system). All of this together amounts to a very strong contention. Perhaps adopting one or more of the parenthesized alternatives would result in a stronger claim.

Let me comment on each of the four possibilities individually, starting with the suggestion that although fairness might not be necessary for sustained cooperation in international trade, it nonetheless helps (amendment 3 above). The careful eye will have noticed that the defenders of the cooperation argument cited earlier themselves alternate between these two suggestions. Although Franck, for instance, claims at one point that “the pursuit of a shared conception of fairness is the *necessary* starting point for devising any lasting allocational rules [...] likely to command respect and pull towards voluntary compliance”, only 5 pages earlier he had made the weaker claim that distributive fairness merely “encourages” the latter (1995 13; 8).²¹ The weaker claim is certainly easier to swallow (although the difficulties of proving it empirically remain considerable). But one can see why some defenders of the cooperation argument might find themselves drawn to the stronger claim, in spite of this. Assume that one's motivation (even if inchoate) in making the cooperation argument is to generate its conclusion. (I

Treating welfare as the metric of fairness may sound like a natural move for the welfare consequentialist. In fact, however, it is far from obvious that a welfarist should make it. If principles of fairness are understood in indirect consequentialist terms, the question of what metric they should employ is itself to be determined by welfare consequentialist considerations. We are to weigh up, that is, the long-run welfare consequences of compliance with principles of fairness that calculate cooperative gains in terms of welfare, and the long-run welfare consequences of compliance with principles of fairness that calculate those gains in terms of objective resources instead. My view is that the latter is the superior option, and hence that the line of argument explored in this footnote is unsuccessful. Explaining why would take us far into the (welfare vs. resources vs. capabilities vs...) currency of justice debates, a topic beyond the scope of this paper.

²¹ I have likewise been a little free and easy with the distinction, for purposes of exposition.

take it that there is nothing wrong with this, provided that one goes in willing to admit defeat). The weaker the causal connection between fairness and cooperation, the weaker that conclusion will have to be. What we will end up with, if we opt for the weaker claim, is something like “welfare consequentialism provides some *prima facie* reason in support of distributive fairness in the multilateral trading regime”. And, to those welfare consequentialists whom I earlier termed “assimilationists”, this will be unsatisfying. Assimilationists do not seek a fully airtight case for compliance with fairness. Common-sense morality itself accepts that sometimes fairness must be sacrificed in light of other competing considerations. Nonetheless, an assimilationist case for fairness ought to be *sturdy*: violations of fairness are supposed to be rare, regrettable, a last resort. A cooperation argument of the form presently under consideration does not, I submit, fit this bill. Something similar applies to the suggestion that premise 1 appeal to partial rather than full fairness. A claim such as:

1b. At least *most* aspects of the multilateral trading regime must be at least *mostly* fair to at least *most* of its participants if cooperation within that regime is to be sustained over the long term.

is easier to believe than premise 1 as it stands, but will not generate the sort of robust and general condemnation of unfairness that is a core feature of common-sense morality.²²

What about adjusting premise 1 so that it is not actual fairness but rather the *perception* of fairness that is claimed to be necessary for sustained cooperation with the regime - *i.e.* as follows?

1c. The multilateral trading regime must be *perceived* (by the majority of its participants) to satisfy norms of distributive fairness if cooperation within that regime is to be sustained over the long term.

Again, at least some of the theorists cited earlier seem to have premise 1c. rather than 1 in mind. And it is arguable that it is in fact the widespread perception rather than the fact of fairness that is doing the real work in at least two of the general arguments for the connection between fairness and cooperation given earlier. People committed to fairness are prone to defect when they perceive themselves or others to be unfairly

²² Such a claim might also continue to run up against the awkward fact of the contemporary regime’s health, since it is debatable whether the condition that it specifies is currently satisfied.

treated, whether or not they *are* being so treated, and the utility of fairness as a focal point in cooperative bargaining clearly depends on a shared perception of what fairness requires.

Should an assimilationist welfare consequentialist adopt 1c? We can begin our answer by recalling that, although assimilationists are, by definition, out to capture the core features of common-sense morality, they also want to retain a reasonable degree of critical purchase on it (see p.6 *supra*). A project of accommodating the perception of fairness rather than its actuality (via a version of the cooperation argument featuring 1c) is problematic from this perspective, given that, although it satisfies the first of these desiderata, it struggles with the second. The result of adopting 1c will be the following amended conclusion to the argument:

Conclusion (c). Welfare consequentialism requires that the multilateral trading regime be perceived (by the majority of its participants) to be distributively fair.

This conclusion does not seem likely to satisfy the assimilationist, given that perceptions of fairness, after all, are not guaranteed to track the moral facts about fairness. In the specific case we are considering, the possibility of disjunction derives from two sources. The first is people's systematic tendency to evaluate fairness in their own favor.²³ The second, more interesting, source is the fact that everyday perceptions of fairness in international trade may attach to an entirely different *subject* than at least some assimilationist welfare consequentialists themselves have in mind. Explaining this second source of disjunction between the perception and actuality of fairness will take us on a brief detour into the fourth possible amendment of premise 1 listed above.

Since Rawls (1971), what the majority of Anglo-American political philosophers have had in mind when talking about distributive justice is *structural* rather than *transactional* fairness. This distinction aims to capture the difference between the way in which an institutional structure as a whole distributes benefits and burdens and the way in which particular actions taken within that structure do. The moral interest of this distinction for those who rely upon it is the possibility that the two types of fairness may come apart: a structurally fair system might be rife with transactional unfairness, and a system in which all transactions are fair might be structurally unfair. If we take this

²³ See Baron (1998) for a survey of the relevant experimental evidence in social psychology.

structural approach to the multilateral trading regime, what we will have in mind in speaking of distributive fairness is the way in which the net gains from participating in that regime are distributed across all countries or individuals, over a substantial time period. We will not have in mind the fairness of, say, a particular trade deal between Australia and the U.S., that of U.S. trade policy in general, or that of the price you paid for your coffee this morning. To adopt this structural focus is not to deny that things of the latter sort, along with many other trade-related matters, might be unfair.²⁴ It is to believe, however, both that it is sensible to speak of structural fairness in international trade, and that the latter has a certain priority, in both method and substance, over other forms of fairness in the area.

Now, although, as mentioned, commitment to the structural approach to distributive justice is now the majority line in analytic political philosophy, common-sense moral thought on fairness in international trade does not naturally operate in this structural mode. The most common determinants of perceived distributive unfairness in international trade are arguably beliefs about the terms or consequences of particular trade deals, rather than those of the multilateral trading regime as a whole. In particular, when speaking of fairness in international trade, both trade negotiators and the broader public often seem to be relying on the following transactional principle of reciprocity:

Fair Return. Concessions in market access made by one country to another ought to be returned by the latter in the form of concessions in market access conferring a similar benefit.²⁵

This principle is frowned upon by economists, due to its underlying mercantilist assumption that every granting of market access is to be considered a “concession” or “burden”. Trade theory suggests, to the contrary, that the economic case for free trade is unilateral: the result of country A’s reducing its trade barriers to country B is of benefit to A, whatever B does with its own. More to the present point, philosophers of global

²⁴ Risse (2007), for instance, argues that the choice to engage in international trade with oppressive governments may be criticizable on grounds of unfairness to the oppressed, and that a country that does so trade may owe its own population compensation for losses suffered as a consequence of that trade, again on grounds of fairness.

²⁵ Brown and Stern report that “In bilateral trade negotiations, negotiators have often directly compared, product by product, the size of the tariff cuts and the volume of trade involved as indicative of equivalence”, and that, in the multilateral case, “in recent negotiations among the developed countries a rough sense of equivalence could be perceived to have guided them even if there was no close accounting of the gains in market access.” (2006)

justice of the structural persuasion are likewise wary of *Fair Return*, due to its transactional nature. To cite James, the criterion of reciprocity “focuses too much on the trade deal on the bargaining table, instead of on what ongoing negotiations can and ought to be about: whether the system of trade is structurally fair” (2008(b)). The underlying concern here is that some violations of *Fair Return* may in fact be required by structural fairness, if the ultimate aim is to rectify a morally more fundamental imbalance in the multilateral trading regime as a whole.

We can refrain, for present purposes, from a direct assessment of the attractiveness of *Fair Return* as a transactional principle. The present relevance of the foregoing discussion is its underscoring of the possibility that popular views about international trade constitute a case of the divergence of perceptions of unfairness from the sort of unfairness with which we should actually be concerned, and therefore a case where assimilationists should not be interested in accommodating the former. If this were so, 1c would not deliver the conclusion that the assimilationist welfare consequentialist seeks.

Whether or not that conclusion is in fact obtainable via 1c depends on the answer to two questions. First, *is* it in fact the case that popular perceptions of unfairness in international trade relate primarily to particular trade deals, rather than to the distributive effects of the multilateral trading regime as a whole? It might be argued that, although people are concerned mainly with a principle such as *Fair Return* in the day-to-day assessment of the fairness of trade, over longer time-frames – for instance when assessing an extended negotiation round - they are also concerned with a structural principle of non-discrimination something like the following:

Fair Benefit and Burden. The benefits and burdens created by the multilateral trading regime ought to be distributed equally amongst participants, unless special justification can be given for unequal treatment.

If popular perceptions of fairness do incorporate this structural element, a cooperation argument incorporating 1c would have more chance of satisfying both structural and transactional theorists of justice, although the problem of personal bias distorting fairness perception would remain.

Second, is a structural approach to distributive fairness in international trade, after

all, appropriate? Assuming now that the people are indeed focused primarily on transactional fairness, perhaps the people have it right. Maybe it does not make sense, as a quite general matter, to speak of the fairness of a structure as a distinct subject of assessment from the fairness of the sum of the transactions taken within it (Nozick 1974). Or, even if the notion does make sense in some cases, perhaps it does not in the case of international trade, or that of global politics in general. If, as some philosophers have argued, the latter cases lack the kind of dense, coercive background institutions that render talk of structural fairness appropriate in domestic politics, it may make perfect sense to apply international principles of distributive fairness exclusively at the level of particular economic transactions.²⁶ It is not my intention to adjudicate this complex debate here. My aim in raising it is to make the point that the attractiveness of a cooperation argument incorporating 1c will depend in part on what sort of global justice theorist one is: i.e. on one's broader moral conception of both the nature of global politics and of distributive justice. Those assimilationist welfare consequentialists who do think that principles of structural fairness apply to international politics and ought to be the focus of distributive assessment there may find such a strategy disappointing. So too, of course, might those who deny the latter claims, given the quite general possibility that perceptions of fairness are mistaken.

Let us take stock. I began by suggesting that premise 1 of the cooperation argument is not currently supported by the facts. The multilateral trading regime is plausibly unfair, yet it is far from obvious that cooperation within it is genuinely faltering. The ensuing discussion suggested that welfare consequentialists attempting to amend premise 1 so as to better fit those facts run up against a dilemma. While such amendments may make the first premise of the cooperation argument more plausible, they will do so at the cost of either significantly weakening its conclusion (in the case of amendments 2 and 3), or of very likely curtailing its appeal to a select, and not very common, type of justice theorist (amendments 1 and 4). We have seen that a cooperation argument incorporating 1c might succeed, but only if one of the following two conditions are satisfied: i) people in general care about structural fairness in international trade, or ii) concerns of structural fairness in international trade are misguided - and if perceptions of fairness can be expected, more generally, to reliably

²⁶ See Nagel (2005).

track the facts of fairness.

It may be worth considering, before moving on, how we should reconcile this somewhat dismal situation with what had looked earlier like a set of fairly compelling arguments for a strong, positive and quite general connection between fairness and cooperation. There are two possible diagnoses of what has gone wrong. One is to claim that, on reflection, those arguments are weak across the board, perhaps by virtue of stumbling over a conflation of perceived fairness with actual fairness, facilitation with necessity, *etc.*, along the lines considered above. Another is to claim that international trade exhibits special features that make cooperation within it function as an exception to what is in some other circumstances a strong case. We can point to two possible sources of exception. The first is scale. Perhaps fairness is necessary for sustained cooperation amongst small groups of people, but not so in the case of very large groups, such as that of the multilateral trading regime. Zajac reports, of the social psychological evidence, that “[i]n bargaining experiments with a small group of persons (six or fewer) in a personal setting, fairness considerations typically loom large...At the other extreme – market-like experiments in an impersonal setting especially involving large numbers of persons (twenty or more) - fairness considerations tend *not* to influence outcomes” (1995, 109). However, it is difficult to know what to make of this evidence. Given that, in principle, the fairness-as-a-focal-point and fairness-as-insurance arguments should hold as much in large-scale as in small-scale situations, what we may be running up against here are the evidential limits of laboratory testing rather than the limits of fairness in fostering cooperation. Another possible source of exception is the extreme imbalance of power that exists between participants in the system of international trade. The force of the fairness-as-insurance argument, in particular, plausibly depends on each party to cooperation being genuinely vulnerable to the others. If the more powerful have a secure expectation of not being at risk of exploitation by the less powerful at any point in the future, the insurance motive for acting fairly loses its force. Given that the U.S. and E.U. do not seem especially vulnerable to Bangladesh or Ethiopia, perhaps this condition fails to be satisfied in the case of international trade. Again, however, the force of this diagnosis is unclear. As Young suggests, “the broader the functional scope of a regime and the longer it is expected to remain in place” (1999, 254), the greater the utility of fairness as insurance. If we think of the multilateral trading regime, as we should, not as a one-shot bargaining situation, but as

an ongoing structure of negotiations and practices expected to last over a significant period of time, trade relations look less like a situation of unmitigated domination by some parties over others than like a situation of mutual but asymmetric dependence.²⁷ So the answer to our question remains unclear.

5.

Let us assume for the moment that the most promising version of premise 1 is true, and turn, more briefly, to premise 2. The claim that:

Sustained cooperation within the multilateral trading regime will have better welfare consequences than the non-cooperative alternative

can be understood as the conjunction of the following two claims:

2.1. Sustained cooperation within the multilateral trading regime is necessary for²⁸ the full or nearly full liberalization of international trade.

2.2. The full or nearly full liberalization of international trade will produce better welfare consequences than the current system.

These theses are both contentious. Defenders of 2.2 draw on the substantial body of trade theory in economics, which purports to demonstrate that free trade between countries promotes economic growth, by virtue of allowing each country to fully exploit its comparative advantage in production. Gains from free trade include the benefits of specialization (economies of scale), access to larger markets, foreign ideas and technology, and the availability of a broad range of cheaper or otherwise inaccessible goods. These economic benefits are said to translate into increased national welfare - at least up to a

²⁷ This response need not rely only on the rather abstract possibility that the bargaining power of developing countries will increase in future decades. As James notes, developing countries already have several negotiating tools at their disposal:

Although lesser dependence allows some countries to erect trade barriers without immediately threatening their own access to the markets of more dependent parties, this power is limited by their own dependence and vulnerability to trade protection...More dependent and relatively less powerful parties...can threaten partial or issue-specific retaliation without wholly ceasing trade; they can increase the cost of retaliation by banding together in coalitions; they can legitimize the specified retaliation, or induce capitulation, by seeking rulings in local and foreign courts; they can rely on moral argument with the support of international society; and, if nothing else, they can impose the costs of an inconvenient trade war (which may have implications for other issue areas, such as security).

²⁸ Again, one could substitute "promotes" for "is necessary for", with consequences similar to those mentioned earlier in relation to premise 1.

certain level of GDP, and provided that certain domestic distributive constraints are satisfied.²⁹ We can add to this economic argument a version of the *doux-commerce* thesis according to which expanded trade relations reduce the likelihood of war. Critics of 2.2 fall into two camps. Some accept the points just mentioned, but claim that trade liberalization is beneficial only under certain conditions that the more ardent free-traders are apt to ignore. To be beneficial, it is suggested, liberalization must be both carefully sequenced and appropriately sensitive to values – such as domestic distributive justice, national security, culture, collective self-determination, and social and environmental standards - not directly related to economic growth. A frequent charge from this camp is that, fueled by a mystic faith in the iron laws of economics, trade liberalization has become for some theorists and policy-makers an end in itself, to the detriment of a broader and morally more fundamental development agenda. Some critics of this kind may be said to endorse 2.2, since, once their concerns regarding process and complementarities are satisfied, their faith in the welfarist desirability of trade liberalization may remain unshaken. Others come closer to a rejection of 2.2. They claim that, however trade liberalization is implemented, factors unaccounted for in the neoclassical trade model (e.g. poverty, inequality, incomplete capital and risk markets), or the unavoidable distortions of WTO governance in a very unequal world, will inevitably make further trade expansion a destructive or dangerous goal.

An obvious fact, but one of which it is easy to lose sight when following the details of these debates, is that they incorporate disputes over moral as well as empirical matters. This is because, fully specified, 2.2. will include a thesis about which welfare consequences are “better” than others. As a result, even if we were sure of all of the relevant facts (and counterfactuals) in this area, we might continue to disagree over the desirability of the results of trade liberalization, precisely because we evaluate those results differently. 2.2, then, is a point in the cooperation argument where the welfarist’s choice of interpersonal aggregation function makes a difference to his/her assessment of the argument (see fn 4). I submit that there is at least some reason to believe that utilitarians are *more* likely to support greater trade liberalization than prioritarian or egalitarian

²⁹ Held (2004) notes that “abolishing all trade barriers could boost global income by \$2.8 trillion and lift 320 million people out of poverty by 2015. In principle, this could cut global poverty by a quarter”. Oxfam claims that “If Africa increased its share of world exports by just one per cent, it would generate \$70bn – a sum that dwarfs the \$14.6bn provided through debt relief and aid combined” (2002, 50).

welfare consequentialists, given their (ultimate) lack of concern for the sorts of distributive considerations that are valued by the latter, and that contribute to much contemporary criticism of free trade.

As it turns out, the same Ricardian insight of comparative advantage that is taken to support 2.2 is used by some to question 2.1. As noted earlier, if economic theory is to be believed, the case for reducing trade barriers is entirely unilateral: each country should drop its own, even if others do not. If this is so, cooperation within the multilateral trading regime would seem to be superfluous to the goal of increased trade: self-interest alone should suffice. Defenders of 2.1 respond that, even if cooperation should not be necessary, it turns out that it is. The fact is that countries have only been willing to liberalize their own markets in return for a similar commitment on the part of their trading partners.³⁰

The partial and fully empirical nature of 2.2 and 2.1, respectively, renders debates over their truth largely outside philosophy's sphere of operations.³¹ Philosophers *are* however, trading to their comparative advantage when considering the implications of the truth or falsity of 2.1 and 2.2 for the rest of the cooperation argument. In that vein, note this. If 2.1 or 2.2 are false, the assimilationist's hopes for the cooperation argument are surely crushed. But, less obviously, the situation might not be materially better if both 2.1 and 2.2 were true. The reason springs from the fact that premise 2, as stated, is ambiguous between the following two readings:

2a. *Some* forms of sustained cooperation within the multilateral trading regime will have better welfare consequences than the non-cooperative alternative

³⁰ On the successes of the regime in promoting liberalization, see Sykes (2004):

Tariffs have declined steadily from an average rate of around 40% when GATT was founded to an average rate of about 4% presently. And although the history of the WTO/GATT system is not free of tension, it has so far avoided the problem of large scale trade wars that plagued the global economy before it...Member states also appear to comply with the rules of the new dispute resolution system. In a clear majority of cases where a measure has been held to violate WTO law, members have abolished it. And although members have declined to abolish a few measures that have been ruled illegal, they have accepted the measured retaliation authorized by the system as an alternative.

³¹ In fact, it is not clear whether social science can settle them either, given the weakness of evidence for and against each. The fact that "to date, not one successful developing country has pursued a purely free market approach to development" (Driskill 2007, 17) makes it difficult to either prove or disprove the merits of that strategy; similarly, it is very difficult to know what degree of liberalization we would have reached today in the absence of the GATT/WTO system.

2b. *All forms of sustained cooperation within the multilateral trading regime will have better welfare consequences than the non-cooperative alternative*

2.1 and 2.2, taken together, imply 2a. But they do not generate the much stronger 2b. Given that 2b appears quite implausible once distinguished from 2a, this will be experienced as no great concern by defenders of free trade. Free traders can happily admit that some specific forms of cooperation within what is an otherwise socially beneficial liberalization regime might have regrettable welfare consequences: indeed, some claim this of the TRIPS agreement mentioned in the introduction.³² But a failure to generate 2b is nonetheless problematic for assimilationist defenders of the cooperation argument. 2a, unlike 2b, has a reasonable chance of being true. But to get the strong conclusion that they are after, it is 2b, not 2a, that assimilationists need. This is because, in a cooperation argument incorporating 2a, the value of fairness will sink or swim with the value of the particular form of cooperation under consideration. Unfortunately for the assimilationist, whether or not most forms of cooperation within the regime do turn out to be valuable, the mere fact that the question needs to be asked for the purpose at hand is counter-intuitive. It is not generally thought to be a legitimate excuse for unfair treatment within a cooperative venture that that venture is, anyway, up to no good.

6.

The foregoing discussion has brushed up against a familiar problem faced by indirect consequentialist arguments in general. Although a consistent practice of adhering to a given moral rule may produce the best consequences on balance, it will usually remain the case that, in some particular instances, breaking that rule will be the locally optimal course of action. In such cases, claim critics, the consequentialist should, according to her own lights, break the rule. To do otherwise would be to perform an action that produces less valuable consequences than those available by means of another action - which, according to consequentialism, is wrong. But if the consequentialist is to make her decisions directly on the basis of the consequences of particular acts falling under the scope of a rule in this way, the rule itself will have ceased to play any role in her deliberations. A willingness to violate a rule whenever the consequences so direct is tantamount to not having a rule at all. Thus, consequentialism cannot justify adherence to moral rules. The standard answer to this objection rejects the

³² For a good philosophical discussion, see James 2008(a).

crucial claim that consequentialists ought to break a generally justified rule in cases where the consequences of doing so would be locally optimal. If the rule in question does indeed have a consequentialist justification, it is argued, the negative consequences of such occasional local failures will already have been taken into account in the process of justifying it. The fact that a justification for the rule exists entails that, in the contest between rigid adherence to the rule and a flexible, situation-specific commitment to it, the consequences of rigid adherence have won out.³³ Of course, this response only succeeds if there *does* exist a consequentialist justification for the rule. So the onus is on the indirect consequentialist to show that that justification exists.

What is needed, in order to generate a stable, convincing and genuine welfare consequentialist case for distributive fairness in relation to the multilateral trading regime is reason to think that distributive fairness within that regime is justified *across the board* by welfare consequentialism, in the way just suggested (with the recognition that rigid rules may themselves incorporate a set of legitimate exceptions). We do not want our case to leave us with the vision of welfare consequentialists humming and hawing over particular instances of distributive fairness in international trade, wondering whether or not they are justified. Does the cooperation argument provide a genuine, secure and convincing welfare consequentialist case of this kind? We have seen several reasons to doubt it. There appears to be altogether too much space for slippage both between fairness and cooperation and between cooperation and welfare to generate the robust case needed. The argument appears to work best for those assimilationists who are both content to confine themselves to concerns of transactional rather than structural fairness (p. 22 *supra*), and who are utilitarians (p. 25-6 *supra*): a rather small set of the range.

This result would not be disappointing for other assimilationist welfare consequentialists if the respective welfare consequences of distributive fairness and unfairness in international trade appeared likely to differ only slightly. If that were the case, assimilationist welfare consequentialists might be forgiven for giving fairness the benefit of the doubt. But, in fact, there are reasons, independent of the cooperation

³³ See Rawls (1955), 281:

It is a mistake to think that if the practice [of promising] is justified on utilitarian grounds then the promisor must have complete liberty to use utilitarian grounds to decide whether or not to keep his promise. The practice forbids this general defense; and it is a purpose of the practice to do this.

argument, to think that distributive fairness in international trade and global welfare may come into significant conflict.

Perhaps the most pressing of these is the possibility that fairness in trade will divert resources from the core welfarist goal of global poverty alleviation. There exists potential for two distinct kinds of conflict between what I earlier called global “decency” and fairness in international trade. The first derives from the fact that decency might sometimes best be promoted by means of “targeted” actions that are incompatible with a fairness-related commitment to universality of treatment. As an example, take the following case mentioned by James:

[S]uppose that, because India and China combined contain much of the world’s population, and much of the world’s poor, the WTO could maximize poverty reduction globally by allowing trade only with these countries. Perhaps each country is to erect high trade barriers against the rest of the developing and less developed world. We would not call this fairness in trade...the poor of Latin America and Sub-Saharan Africa are supposed to have a chance to benefit from access to global markets, even if that chance costs the poor elsewhere. (2006, 716)

The second kind of potential conflict between decency and fairness derives from the fact that decency might sometimes best be promoted by means of unequal distributions of benefits and burdens that fairness, by virtue of its tendency towards equality, rules out in that particular instance. Say that we could modify the multilateral trading regime so that some or all developing countries were given privileges far in excess of the “special and differential treatment” to which they are entitled on grounds of fairness, including no requirement to reciprocate trade concessions made by developed countries, even in cases where such concessions would not be especially burdensome. Such inegalitarianism, while plausibly violating fairness, might under certain conditions be maximally conducive to global decency. A failure to institute a reform of this kind under those conditions would represent a commitment to fairness over decency, and quite possibly over the best welfare consequences overall.

Just how problematic all of this is for assimilationist welfare consequentialists, in the grand scheme of things, will chiefly depend on the answers to two questions. The first is whether or not there exist additional considerations, distinct from those adduced in the

cooperation argument, connecting distributive fairness in international trade to desirable welfare consequences that might counterbalance the foregoing concerns. One option worth exploring here involves pointing to the just-mentioned fact that fairness often has broadly egalitarian implications, and advocating the welfare benefits of international resource egalitarianism. The second question is the extent to which non-welfare consequentialist accounts do any better than welfare consequentialism in providing a stable and compelling case for fairness in the multilateral trading regime. Given that any reasonable moral theory must pay some significant heed to welfare consequences, it may be that many such accounts have similar implications to those just mentioned. Whatever the answers to these broader questions, however, there appears to be only a slim chance of the cooperation argument, by itself, making distributive fairness in international trade safe for welfare consequentialism.

References

Richard Arneson, "Broadly Utilitarian Theories of Exploitation and Multi-National Clinical Research," in Ezekiel Emanuel and Jennifer Hawkins, eds. *Exploitation in Medical Research* (Princeton: Princeton University Press, 2007).

Tom Athanasiou and Paul Baer, *Dead Heat: Global Justice and Global Warming* (Seven Stories Press, 2002)

Jonathan Baron, *Judgment Misguided: Intuition and Error in Public Decision-making* (Oxford: Oxford University Press, 1998)

Scott Barrett, *Why Cooperate? The Incentive to Supply Global Public Goods* (New York: Oxford University Press, 2007)

William Baumol, *Superfairness: Applications and Theory* (Cambridge, MA: MIT Press, 1986)

John Broome, *Weighing Goods: Equality, Uncertainty and Time* (Cambridge, MA: Basil Blackwell, 1991)

Andrew Brown and Robert Stern, "Concepts of Fairness in the Global Trading System", Working Paper Series Number 13, International Policy Center, University of Michigan, Feb. 2006

Shubham Chaudhuri, Pinelopi Goldberg and Panle Jia, "Estimating the Effects of Global Patent Protection in Pharmaceuticals: A Case Study of Quinolones in India" (Washington, DC: World Bank, 2004)

Jules Coleman, *Risks and Wrongs* (Cambridge: Cambridge University Press, 1992)

Robert Driskill, "Deconstructing the Argument for Free Trade" (ms, February 2007)

Gerard Elfstrom, *Ethics for a Shrinking World* (London: Macmillan, 1990)

Ernst Fehr and Simon Gächter, "Cooperation and Punishment in Public Goods Experiments", *The American Economic Review* (Sept. 2000)

Thomas Franck, *Fairness in International Law and Institutions* (New York: Oxford University Press, 1995)

- Robert Goodin, *Protecting the Vulnerable* (Chicago: University of Chicago Press, 1985)
- Robert Goodin, *Motivating Political Morality* (Cambridge, MA: Blackwell, 1992)
- David Held, *Global Covenant: The Social Democratic Alternative to the Washington Consensus* (Polity Press, 2004)
- Robert Howse and Michael Trebilcock, "The Fair Trade-Free Trade Debate: Trade, Labor and the Environment", *International Review of Law and Economics* 16: 61-79 (1996)
- Aaron James, "When International Intellectual "Piracy" is Fair" (2008(a), ms)
- Aaron James, "Skepticism about Fairness in Trade" (2008(b), ms)
- Aaron James, "Distributive Justice without Sovereign Rule: The Case of Trade," *Social Theory and Practice* 31: 4 (2005): 1-27
- Aaron James, "Equality in a Realistic Utopia," *Social Theory and Practice* 32:4 (2006): 699-724
- Louis Kaplow and Stephen Shavell, *Fairness vs. Welfare* (Cambridge, MA: Harvard University Press, 2002)
- Daniel Kahneman, Jack Knetsch & Richard Taylor, "Fairness and the Assumptions of Economics", *The Journal of Business* 59:4, Part 2 (Oct. 1986), S285-S300
- Barbara Mellers and Jonathan Baron (eds.), *Psychological Perspectives on Justice: Theory and Applications* (Cambridge: Cambridge University Press, 1993)
- David Messick, "Equality as a decision heuristic" in Barbara Mellers and Jonathan Baron (eds.), *Psychological Perspectives on Justice: Theory and Applications* (Cambridge: Cambridge University Press, 1993)
- David Messick and Karen Cook (eds.), *Equity Theory: Psychological and Sociological Perspectives* (New York: Praeger, 1983)
- Thomas Nagel, "The Problem of Global Justice", *Philosophy and Public Affairs* 33 (2005): 113-47

Amrita Narlikar, "Fairness in International Trade Negotiations: Developing Countries in the GATT and WTO", *The World Economy* (2006)

Robert Nozick, *Anarchy, State and Utopia* (New York: Basic Books, 1974)

A. Okun, *Prices and quantities: A macroeconomic analysis* (Washington, DC: The Brookings Institution, 1981)

John Rawls, *A Theory of Justice* [1971] (Revised ed.) (Cambridge, MA: Harvard University Press, 1999)

John Rawls, "Two Concepts of Rules", *The Philosophical Review*, Vol. 64 (1955)

Matthias Risse, Fairness in Trade I: Obligations from Trading and the Pauper Labor Argument, *Politics, Philosophy and Economics* 2007

Peter Singer, *One World: The Ethics of Globalization* (New Haven: Yale University Press, 2002)

T.N. Srinivasan, "Doha Round of Multilateral Negotiations and Development", Working Paper No. 252, Stanford Center for International Development (Sept. 2005)

Joseph Stiglitz and Andrew Charlton, *Fair Trade for All: How Trade can Promote Development* (New York: Oxford University Press, 2005)

Alan Sykes, "The Economics of Public International Law", John M. Olin Workshop Law and Economic Working Paper No. 216 (University of Chicago Law School, July 2004)

Alan Wertheimer, *Exploitation* (Princeton: Princeton University Press, 1996)

Oran Young, "Fairness Matters: The Role of Equity in International Regime Formation," in Nicholas Low (ed.), *Global Ethics and the Environment* (Routledge, 1999), 247-263

Edward Zajac, *Political Economy of Fairness* (Cambridge, MA: MIT Press, 1995)