# Moral Trade*

## Toby Ord

If people have different resources, tastes, or needs, they may be able to exchange goods or services such that they each feel they have been made better off. This is trade. If people have different moral views, then there is another type of trade that is possible: they can exchange goods or services such that both parties feel that the world is a better place or that their moral obligations are better satisfied. We can call this *moral trade*. I introduce the idea of moral trade and explore several important theoretical and practical implications.

INTRODUCTION

Our world contains a great diversity of moral views. There are utilitarians, Kantians, and virtue ethicists. There are Hindus, Buddhists, Christians, Muslims, and Jews. There are some who believe there is an unalienable moral right to have an abortion, others who hold it to be the moral equivalent of murder; some who believe eating meat is the moral equivalent of murder, others who hold it to be completely innocuous; some who believe there is a moral obligation to share society's wealth, others who hold that each person has a right to keep all the fruits of their labor.

People with different moral views often find themselves in antagonistic relationships. They might find cause for a local alliance if their views agree on a salient issue, such as a utilitarian and a Hindu joining forces to promote vegetarianism even though they support this practice

for very different reasons.[1] However, if their views differ on a salient issue, then the other side is often regarded as the enemy and battle lines are drawn.

In this essay, I want to show a different way forward. I want to show how people with conflicting moral views can cooperate, creating an outcome that is better according to each of their moral theories. Consider the following vignette:

> *Animals and the Poor:* Victoria and Paul are close friends. Victoria is a vegetarian because she is aware of how much animals suffer in the modern meat industry and wants to reduce this suffering as much as possible. Paul is deeply concerned about global poverty and donates 10 percent of his income to organizations that can effectively improve people's lives in poor countries. He enjoys many vegetarian dishes, but still eats meat most of the time since he finds it a bit more convenient, and, as a contractualist, he doesn't think that animals matter from a moral point of view. Victoria can see that Paul is prepared to make sacrifices to help others, but it frustrates her that he isn't prepared to be a vegetarian—especially since it would only be a small sacrifice for him. He has similar views about her lack of serious financial support for the world's poorest people. One day they decide to make an arrangement. She will start to donate 1 percent of her income to fight global poverty if he agrees to become a vegetarian. Either of them can end the deal at any time, releasing them both from any commitment. They could check up on the other person, but as they trust each other they decide it is sufficient to ask each other how they are keeping up once each year.

Both Victoria and Paul see this agreement as a good deal. Paul doesn't think that animals matter morally and wouldn't be prepared to forgo eating meat just to help them. However, with this arrangement his vegetarianism causes several hundred dollars per year to go to fighting poverty, which he thinks is well worth the small sacrifice. Victoria values poverty reduction in other countries and used to donate about $100 per year to poverty-related charities. But she cares much less about it than she cares about animal welfare and wouldn't be prepared to give 1 percent of her income each year just for this cause. Through their arrangement, however, she is thrilled at being able to roughly double her impact on preventing animal suffering at only a modest financial cost.

Victoria and Paul have just engaged in what I shall call *moral trade.* Different self-regarding tastes allow for gains from trade. For example, if

---

1. John Rawls calls this sort of instrumental agreement a "modus vivendi," which he contrasts with an "overlapping consensus," where the different parties agree to collaborate "for the right reasons." See John Rawls, "The Idea of an Overlapping Consensus," *Oxford Journal of Legal Studies* 7 (1987): 1–25.

two children would prefer to eat each other's packed lunches, they can trade them and both benefit. Similarly, if people have different moral views, they can potentially realize gains from moral trade, making them both think that their moral view is better satisfied if they make a trade than if they don't. For example, consequentialists may think that the trade produces a better outcome, while deontologists may think that it better fulfils their duties or that it is a supererogatory option. Not all moral theories will be able to gain from moral trade, but I shall argue later that many can.

For the moment, I'll set aside questions about whether a moral trade could be bad from the perspective of people who weren't involved in the trade and whether it could be bad from the perspective of the correct, or best justified, moral theory (if any). These issues, and how one might deal with them, will be discussed in the final section. Until then, we shall just consider the gains from moral trade from the perspective of each of the moral views that the trading parties endorse.

In the above example, Victoria and Paul had fairly compatible moral views. Paul valued poverty reduction and was neutral about animal suffering. Victoria valued both, but valued reducing animal suffering substantially more. Moral trade is even possible on an issue when the two parties have diametrically opposed views, so long as there is some area where they are not diametrically opposed. Consider this second vignette:

> *Gun Rights and Gun Control:* Rebecca and Christopher are coworkers at a technology start-up. Rebecca is a staunch defender of the right to bear arms and fiercely opposes gun control. Christopher doesn't believe that there is such a right and is deeply concerned by the harm that guns can do to society. A new bill on gun control is being debated in Congress and has caused many heated arguments at the water cooler. In one of these conversations, it is revealed that Rebecca is planning to donate $1,000 to a gun rights charity to fund their fight against this bill, while Christopher is planning to donate $1,000 to a gun control charity which is supporting the bill. An onlooker suggests that their combined behavior is pointless: they are just fighting a zero-sum battle. She points out that it would clearly be better if they both donated the money to Oxfam instead—a charity that Rebecca and Christopher both support. While Rebecca and Christopher would each rather their own $1,000 went to fighting the battle for the gun control bill, they accept that the onlooker is right. They each think that Oxfam is a good charity, and would prefer that $2,000 is donated to this organization than that $1,000 is given to help the bill and $1,000 to hinder it. So they decide to make the deal.

Situations like this can be quite common. As well as gun rights and gun control, there are other pairs of almost diametrically opposed charities,

such as pro-life and pro-choice, or the Republican and Democratic political parties. Donations to each of these pairs range from hundreds of millions of dollars per annum to billions of dollars per annum, so there are significant gains to be made.

Though the numbers in the vignette were carefully balanced to make the situation clearer, the ability to trade when supporting opposing causes is quite robust. To begin with, it works even if the opposing charities aren't equally efficient at using donations to achieve their aims. For example, it would work if the gun control charity were twice as effective but Christopher was only planning to give half as much money as Rebecca. It would also work if they were planning to help by very different amounts. For example if the charities were equally effective, but Christopher was planning to donate $10,000, then they could just swap off $1,000 of the donations, leaving Christopher donating $9,000 to the gun control charity. It could even work if they had different empirical beliefs about the effectiveness of the charities. For example, the trade would still make sense if they each believed that their own supported charity was somewhat more effective than the other—so long as they thought that this gap was sufficiently small compared to the value of donating to their compromise charity.[2]

In reality, apparently opposed charities will produce some additional benefits or costs that both parties agree on, such as a more educated public, making the existing regulation more nuanced, or creating a more partisan political culture. However, even if these other effects are net positive, the perceived value of them is likely to be only a small fraction of the perceived value of the donation and in many cases there will be a compromise charity that is judged by both to outweigh these side benefits of the opposed charities.

VARIETIES OF MORAL TRADE

These examples have focused on a direct clash between two different moral views, but similar examples arise when a moral view is contrasted with prudence. For example, Victoria—the vegetarian in the first vignette—could find people and simply pay them to become vegetarians. For some people who are already on the fence about vegetarianism, or who wouldn't

---

2. For example, suppose they each thought their campaigners were twice as effective as those of the other side (so that $1,000 going to both charities was worth the same as $500 to their charity). Then they are comparing something equivalent to $500 to their charity with the option of $2,000 going to Oxfam. If they think that money going Oxfam is more than a quarter as good as money going to their lobbyists, they should agree to trade. One reason people might systematically estimate the effectiveness of their campaign to be higher than their opponents think, is if they think that their arguments are stronger and hence more likely to gain traction. (I would like to thank an anonymous referee for this point.)

find a vegetarian diet particularly difficult, there would be a sum of money that Victoria could offer such that they would be prudentially better off making the trade and Victoria would consider the improvement to the world to be worth the financial sacrifice.[3]

There can also be cases where each party's gains from trade involve self-regarding and moral components. Indeed the first vignette with Victoria and Paul was already like this, since they were both making internal trade-offs between prudence and morality. The moral trade allowed them a new way to gain a lot of what they thought was good (animal welfare, poverty reduction) for a relatively small prudential cost (donating 1 percent, giving up meat), leading to an outcome that they each judged to be better, all things considered.

We could summarize these cases by what each person sees as the change in the moral value and prudential value from making the trade, compared to the status quo. For Victoria and Paul, it is moral gain plus small prudential cost versus moral gain plus small prudential cost. For Victoria just paying someone to become a vegetarian, it is moral gain plus small prudential cost versus pure prudential gain. For Rebecca and Christopher, it is simply moral gain versus moral gain. Since the mixed cases are relevantly similar to the pure cases and it would be useful to discuss them together, let us define *moral trade* so as to include both:

> *Moral Trade:* Trade that is made possible by differences in the parties' moral views.

I shall use the term *regular trade* to refer to trade that is not moral trade. I shall use the term *pure moral trade* for moral trade where both parties view the result as morally superior and *mixed moral trade* for other cases.

Note that there are trades where each person thinks the outcome is morally superior which don't count as moral trade on this definition. For example, suppose the directors of two charities with different resource needs trade a car for a website redesign. Even if they have different moral views, this is a trade that was made possible by the different resource needs—it didn't require a difference in their moral views. Cases of this sort are much more similar to instances of regular trade than to the cluster of examples discussed here, and so the definition of moral trade was constructed to exclude them.

Note also that as well as substantive differences in moral views, I include differences due to different indexicals. For example, if Mary and

---

3. Donations to charities with paid employees could be seen as examples of this kind of moral-for-prudential trade. A donor effectively pays an employee of the charity to perform some good act that the employee would not have otherwise done. This can be more efficient than the donor doing that good act directly. Such situations could be seen as examples of specialization or professionalization via (mixed) moral trade.

Susan both hold an agent-relative view which gives a special place to the agent's family, then for Mary this will be a special place for Mary's family and for Susan it will be for Susan's family. This could give rise to opportunities for moral trade. Similarly, a moral theory might give an agent a strict duty to tell the truth but not require any action concerning whether others tell the truth. If two people held this theory, it would give them different duties, which again could give rise to interesting opportunities for moral trade.

An interesting edge case for the above definition is the possibility of intrapersonal moral trade. Many people are guided by prudential considerations and moral considerations in a rather ad hoc mixture. It is possible that they will notice that a certain pair of behaviors is dominated by another pair of behaviors, being both prudentially and morally better. For example, consider the case of Alex, who would dearly like to fly from Boston to Paris for a cultural holiday, but is deeply concerned about carbon emissions and climate change.[4] He decides that he can't justify the prudential benefits at such a moral cost. However, then a friend suggests that he could fly there and offset twice that amount of carbon at a relatively small cost. He likes this idea, since it allows him to finally see Paris despite a small additional cost (enhancing his overall prudential interests), while simultaneously helping fight climate change overall (better fulfilling his moral view). Flying with doubly offset carbon emissions is therefore an improvement over staying in Boston on his two different value systems and constitutes a kind of internal or intrapersonal moral trade. I shall not formally include intrapersonal trade in the definition, but some of what I say below would apply to these cases as well.

## SOME ECONOMIC THEORY OF MORAL TRADE

What can we say about the theory of moral trade? First, it is worth pointing out that moral trade is already encompassed within the standard definition of trade in economics. When economists speak of a person's preferences, they include both the person's prudential preferences and their moral preferences[5] in an undifferentiated set of preferences. So when it comes to trade, moral trade is mixed in with regular trade. For economists, moral trade is thus not an additional kind of trade, but a subcategory.

4. Julia Driver provides a detailed discussion of a case like this in "Moral Bookkeeping and Carbon Offsets," in *Consequentialism and Environmental Ethics*, ed. Avram Hiller, Ramona Ilea, and Leonard Kahn (London: Routledge, 2014), 164–73.

5. This raises a metaethical question. I have attempted to neutrally refer to 'moral views' rather than 'moral beliefs' or 'moral preferences'. Does it matter for a theory of moral trade whether moral disagreements are more like disagreements about beliefs or

This does not mean that it is well known or extensively studied. Apart from contexts when they are carefully outlining definitions, economists pay relatively little attention to the fact that their definition of preferences includes moral preferences. While it is technically covered by the existing formalisms, a paper on moral trade would be considered surprising in economics.

One way to look at moral trade from the perspective of economics would be to note that while we have extensive trade based on variation in prudential preferences, there is strikingly little trade based on variation in moral preferences. Moreover, there are almost no mechanisms for facilitating such trade. A common assumption in economic modeling is a "complete market" where everything is available for trade,[6] but in reality the market for moral trade is very incomplete.

Economic approaches could be used to transform the very simple examples of moral trade that I give (which we could call *moral barter*) into more sophisticated systems to realize greater gains. For example, we could introduce moral trade equivalents of *currency*, *markets*, *bargaining*, and *professionalization*. I shall touch on each of these, but a full treatment is beyond the scope of this essay.

We can also combine techniques from economics with techniques from normative ethics to get more clarity on the nature of moral trade. When two parties are in a situation where moral trade could take place,[7] there will be a set of options available to them. These include all combinations of them each acting in independent ways as well as all the ways they could act in concert with each other. Each of these options can be assessed according to each of the two parties' moral views. We could associate each option with a pair of numbers $(c_1, c_2)$ representing its choiceworthiness according to the first view and according to the second view.[8] We could then represent these options as points on a graph, where the vertical axis represents the choiceworthiness on the first party's moral view and the horizontal axis represents the choiceworthiness on the second party's moral view (see fig. 1).

---

about desires? It seems to me that it does not. After all, just as different tastes can give us mutual reason to perform regular trade, so different beliefs can give us reason to perform regular trade. For example, a large amount of stock market trade only occurs because people have different beliefs about the future valuation of the stock.

6. See, for instance, Mark D. Flood, "An Introduction to Complete Markets," *Federal Reserve Bank of St. Louis Review* 73 (1991): 32–57.

7. This can be scaled up to any number of parties in a trade, but the diagrams would become much more complex.

8. Not all moral theories will have choiceworthiness that behaves exactly like a number: some theories may allow for incomparability, and for some theories the differences between numbers may not be meaningful. Figures 1 and 2 only assume comparability (and their point could be made less elegantly even without comparability). Figures 3 and 4
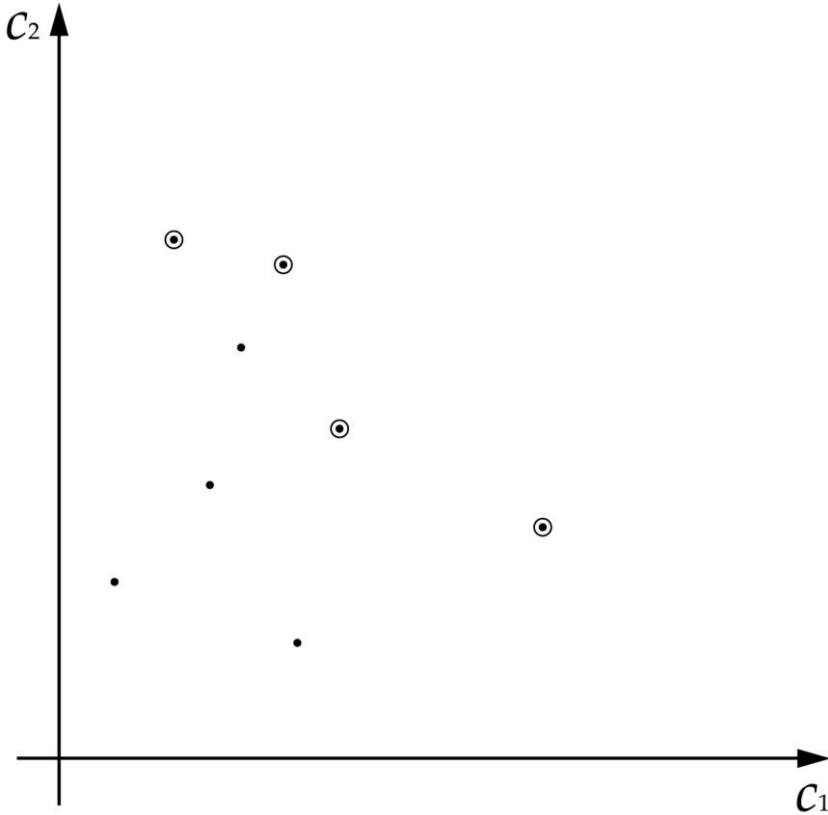
Fig. 1.—The available options displayed according to the choiceworthiness each party assigns them. The Pareto optimal options are circled.

     I have circled four of the options, because they have a special status. They are *Pareto optimal*, meaning that for each of them there is no other option which is better according to one view and at least as good according to the other. In contrast, the uncircled options are *Pareto dominated*. For each of them there is at least one option that dominates them: being better according to one view and at least as good according to the other. From the perspective of a neutral onlooker, we would hope that they could agree to choose one of these circled, Pareto optimal, options.

     It may look like this diagram builds in an assumption that the parties' moral views are orthogonal to each other, but it does not. To the extent to which the views are aligned, the points representing the options will be

will assume that the differences between the numbers are meaningful: that the theory is *cardinal*. None of this analysis assumes that the zero point or the scale is meaningful.

correlated, being scattered closer to the positive diagonal. Opposed moral views will create a negative correlation.

In cases of moral trade, there is an additional important element. One of the options represents a default or status quo: what would happen if they do not trade.[9] In figure 2 I have marked this option with a cross through it. Two of the Pareto optimal options do not dominate this default. This means that they are worse than the default according to one of the parties, so if the trade is being freely made, that party would refuse to consider that option. This leaves us with the options that lie in the indicated quadrant: above and to the right of the default, of which a neutral party would hope that they choose one of the two circled ones.

The question of exactly which option would get chosen depends upon how the parties negotiate. In economics, this is covered by bargaining theory.[10] Developing this here would take us too far afield, but we can at least note that the first party (measured by the vertical axis) would prefer the topmost option and so on down to the default, while the second party would prefer the rightmost option and so on down to the default. The distance by which an option is above the default represents how much "moral surplus" there would be from the first party's point of view if they agreed on it (and similarly for the distance to the right for the second party). Depending on how the parties negotiate it is possible that a considerable portion of the potential total moral surplus might be lost.

In some cases, there might be no options that are above and to the right of the default (see fig. 3 where I have removed all such options). If so, it might look like no trade is possible. For example, suppose that two people are working out how to spend their Sunday afternoon and were by default going to work on their individual projects, related to their different moral views. They realize that if they worked together on the same project, they could achieve significantly more gains. The options where they team up like this are labeled *A* and *B*. However, they each think that the other person's project is not as morally valuable as their own, such that joining the other team would be slightly worse than continuing to work alone on their own project.

9. The analysis of moral trade up to this point is similar to that of moral uncertainty (the theory of how to make decisions when uncertain about moral considerations), with options being compared in virtue of their choice-worthiness according to different moral theories and the ideal option to choose being restricted to the Pareto optimal options. However moral uncertainty does not have an analogue of the default option. For an introduction, see Ted Lockhart, *Moral Uncertainty and Its Consequences* (Oxford: Oxford University Press, 2000).

10. See, e.g., Abhinay Muthoo, *Bargaining Theory with Applications* (New York: Cambridge University Press, 1999).
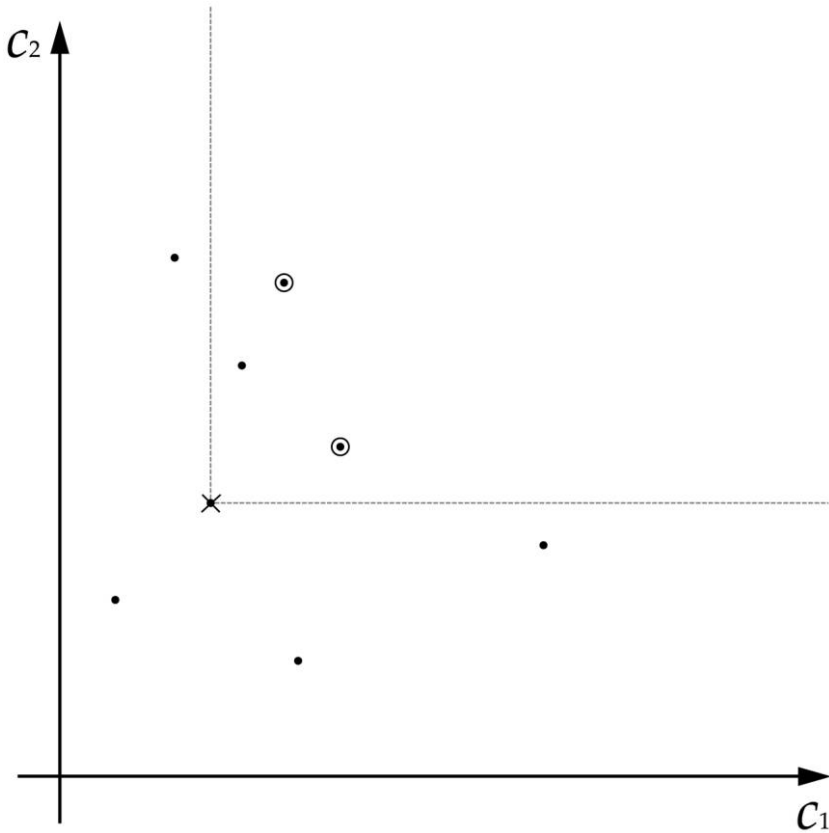
Fig. 2.—Adding a restriction that options must be Pareto superior to the default option

If they know that this situation will be repeated (say, each week) then they could agree to alternate the project they team up on. This shows the value of considering a wide array of options at once when trading to increase the gains of moral trade. However, suppose that this case is a one-off opportunity. What then?

One way out would be to suggest a gamble, for example, to flip a coin and to team up on the winner's project. Alternatively, they could use a randomized method with unequal probabilities. Such randomized methods will typically be judged to have values that span an arc from $A$ to $B$.[11] For example, a 99.99 percent chance of working on $A$'s project would typically be considered by both parties to be almost the same value

---

11. If both parties are risk neutral, it will form a straight line segment.
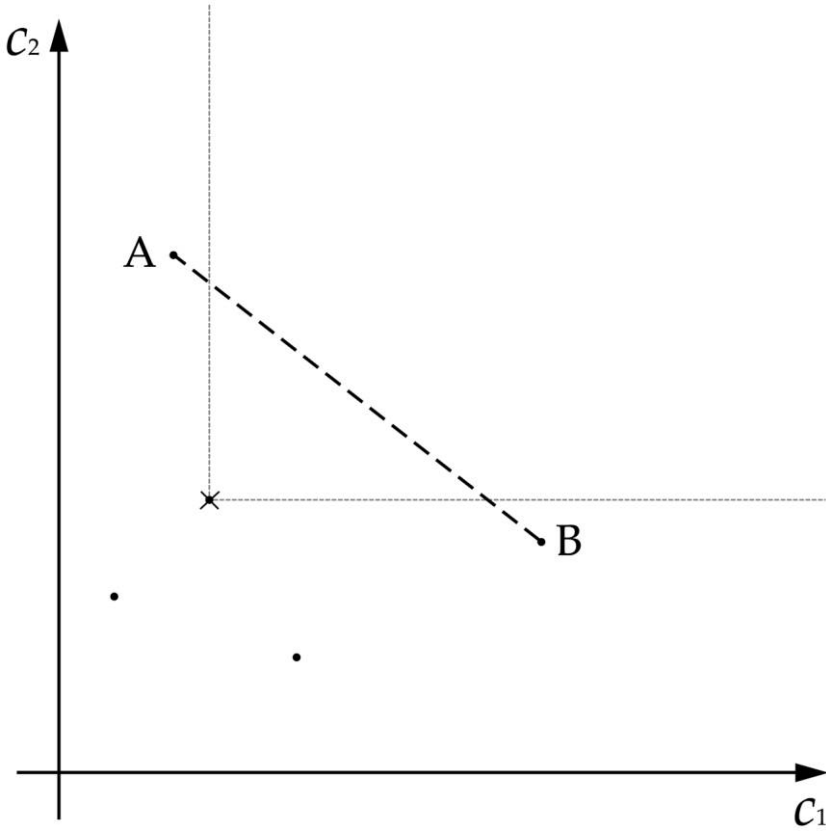
Fig. 3.—The choiceworthiness of new options representing weighted lotteries between options *A* and *B*.

as working on *A*'s project for certain, and as this chance is lowered bit by bit, the value will move slowly across the graph until it becomes very close to the value of *B*. There will often be some range of probabilities such that randomizing which project they team up on using that probability will dominate the default option of working separately. They could then trade to one of those randomized options, flip the coin (or run the lottery), and head off to the winning project together.

Another way out would be via side payments, ideally in some more or less continuous medium. For example, the person whose project they team up on could give the other person a sum of money. In order to keep this as an example of pure moral trade, let's suppose that they do this by reallocating some of their charitable donations for the year to a charity of the other person's choice. This is represented in figure 4. The
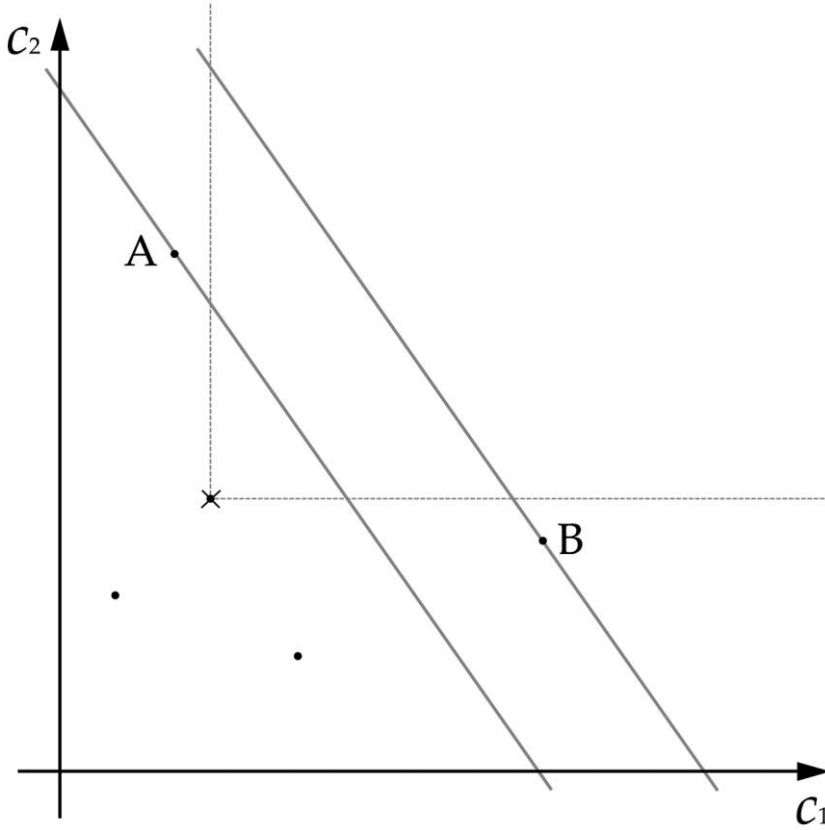
F_{IG}. 4.—The choiceworthiness of new options representing starting with *A* or *B* and making a side payment to the other party.

lines that stretch out from *A* represent the results of teaming up on *A* combined with a side payment of a certain size. As with the randomized method, there is reason to expect small differences in side payments to produce options that are close together in value and so produce a continuous line. The slope of this line would be set by the party's relative moral valuations of the two charities that the donations are being switched between. If the payments are relatively small, then these lines should be relatively straight. If the lines have the shape indicated in figure 4, then the dominant options would involve teaming up on the second person's project and that person switching some portion of their donations to the first person's preferred charity. In this particular setup, the options created by the side-payment approach dominate the options created by the ran-

domized approach and this would always happen unless the lines are quite curved.

## CONSEQUENTIALISM, DEONTOLOGY, AND VIRTUE ETHICS

Let us now consider how moral trade fits in with the major branches of normative ethics. To some extent, it would seem to be a natural fit with consequentialism. Since most forms of consequentialism are based around agent-neutral value, consequentialists typically do not care who achieves a valuable outcome, so long as it is achieved. This means that a consequentialist should see a lot of potential value in moral trades, regardless of whether the other party is a consequentialist or not.

In contrast, many deontological theories involve agent-relative value and side constraints. The most obvious effect of side constraints is that the trade cannot involve the agent doing something she judges to violate a side constraint. However, this should still allow many opportunities for trade: for instance, none of the examples above involve the breaking of side constraints.

There are also more subtle effects. A deontologist who is a vegetarian for moral reasons concerning animals may think that it is impermissible for them to eat meat, and that this is much more morally important than convincing others to not eat meat. They might therefore see relatively less reason to engage in moral trade. However, it seems unlikely that they could see it as without value. For example, if they are a vegetarian because they think it is immoral to inflict suffering on animals, then presumably they are at least somewhat concerned that animals do not suffer and could see at least some value in convincing other people to stop causing them suffering.

It might be thought that the central role of the agent's intentions in many deontological theories would limit the benefits of moral trade. For instance, if a person was only donating to help the poor because this was part of a moral trade and not out of a concern for human welfare, this might lower or even nullify the moral value of that act. However, there are reasons to think that there would still be gains to make from moral trade even in cases of this sort. First, it might be that even if the act of donating with no real concern for the poor lacks value, the other party's act of convincing this person to donate may possess moral value. Second, it would be puzzling if the theory said there was nothing morally superior about the situation in which poor people were helped.[12] And finally, society routinely tries to give people nonmoral reasons to do

---

12. As John Rawls put it: "All ethical doctrines worth our attention take consequences into account in judging rightness. One which did not would be irrational, crazy." John Rawls, *A Theory of Justice* (Cambridge, MA: Harvard University Press, 1971), 30.

moral things, such as threatening them with jail if they murder someone. Assuming the theories see some value in such ways of avoiding murder, they should also see value in trades that give people nonmoral reasons for abstaining from acting in other unethical ways.

Indeed, it is possible that side constraints or agent-relative value could encourage moral trade. For example, someone might think that it is impermissible for them to lie in order to avoid some suffering but that it wouldn't be impermissible to convince someone else to make this lie in order to avoid the suffering. It could also be that having a rigid view that lying and breaking promises is impermissible could significantly help to foster moral trade. For it could let other parties see that the deontologist is truthful about their claims and will follow through with their side of the deal. It could therefore help with the key trust issues involved in making moral trade work in practice.[13]

It may seem that virtue ethicists would see much less reason to engage in moral trade than consequentialists or deontologists. For example, if a vegetarian virtue ethicist is refraining from eating meat because it is the compassionate thing to do, then it is not obvious that it is equally compassionate to convince someone else to refrain from eating meat.

However, theories of virtue ethics often stress the importance of the promotion of virtue in others. For example, in the *Nicomachean Ethics* Aristotle analyzes the sources of virtue and finds habituation to be particularly important: we become just by doing what is just, temperate by doing what is temperate, brave by doing what is brave.[14] By this he means that doing what is required by justice, even if for the wrong reasons, leads us to become just and to go on to do these just acts for the right reasons. He argues that this justifies moral education to inculcate virtue in others, and even for states to use coercion to induce habituation and hence improve virtue in their citizens.[15]

Moral trade should also be able to create habituation and hence improve virtue. While trading with a friend to convince them not to eat meat wouldn't immediately create the virtue of compassion, it should habituate them into acting more in accord with compassion, which may in turn lead them to become more compassionate. It thus seems that virtue ethics, too, provides reasons for making moral trades.

---

13. The trust issue could be particularly difficult for consequentialists given that, if there is no way of enforcing the deal and no potential reputational costs, most forms of consequentialism would say that it is wrong for the consequentialist to go through with their side of the deal. Consequentialism is thus more reliant on these practical mechanisms than the other branches of ethical theory.

14. Aristotle, *Nicomachean Ethics*, ed. Roger Crisp (Cambridge: Cambridge University Press, 2000), bk. 2, chaps. 1–4.

15. Ibid., bk. 10, chap. 9.

TRUST

The major practical obstacle to the gains from moral trade is that of insufficient trust between the two parties. This obstacle comes in two parts. Neither of them are different in principle from regular trade, but in practice they can be much more difficult to overcome.

The first part concerns trust about factual issues, so we shall call it *factual trust*. Suppose we are engaged in regular trade. I am exchanging some of my money for some fruit from your garden. If we were particularly paranoid, or lived in a particularly lawless society, there may be difficulties in setting up the trade. Do you hand over the fruit first, or do I hand over the money? There might also be some issues of trust after the trade. Is that a genuine banknote? Will those apples go bad more quickly than I was led to believe? However, these issues are not too burdensome and trades like this can happen fairly easily, even between people who do not know each other.

In comparison, if Victoria agrees to donate 1 percent of her income in exchange for Paul becoming a vegetarian, how do they know the other one is continuing with the deal? The continuing nature of this deal makes trust a big issue. This is particularly so because, unlike most instances of regular trade, the services are not being delivered in the presence of the person who bought them. Note that one *could* have regular trade with these properties, and one *could* have moral trade which is one-off with services delivered in the presence of both parties. For example, if two parties are trading a vote for one policy with a vote for another, then they only need to be able to trust each other until both votes have been cast. However, in practice it seems likely that this problem of factual trust will be greater for moral trade.

There are various strategies for overcoming this problem in practice. In cases where the parties know each other, simply trusting one another might suffice, since they are unlikely to be interested in deceiving each other.

In other cases, the parties may be able to occasionally check up on each other and call off the deal if either side has not kept up the arrangement. For instance, someone could ask to see the other party's receipts for their donations at the end of the tax year. When this is possible, it turns the situation from a one-shot prisoner's dilemma to an iterated prisoner's dilemma of unknown duration. If the parties think there are sufficiently many rounds left to come, then they have reason to keep to their side of the deal so as to be able to continue the arrangement.

It is also possible to arrange penalties for noncompliance, including legal penalties, just as we have for regular trade. This would be another way of escaping the prisoner's dilemma: having society act as leviathan to lower the payoff for defection, changing the nature of the game. While

one party can't legally force the other to change their diet even if an agreement was entered into, there are other possibilities. The most simple is for each party to put money into escrow, which would be forfeited if they broke the deal. If enough money was placed in escrow this could give both parties reason to keep to the deal even if the chance of deception being discovered was low. This is how contract law works for regular trade: we don't force an actor to make a movie just because they signed a contract to do it, instead they risk financial losses if they breach the contract.

With regular trade, our societies have established an implicit contract when you buy something, which overcomes a lot of the transaction cost of writing contracts. And even at the greengrocer, most of the factual trust comes from custom rather than from criminal law or contract. For example, if the greengrocer refused to give you a dollar of your change, it would be difficult to appeal to the police to get this back. If moral trade were to become sufficiently popular, customs might evolve to support it, and the law might be updated so that moral trades also establish implicit contracts.

The second problem of trust is what one might call *counterfactual trust*. For example, Victoria needs to trust that Paul wouldn't have become a vegetarian without that agreement, and he needs to trust that she wouldn't have increased her donations to Oxfam to 1 or more percent without the agreement. This is particularly hard to police. It is even difficult for each party to determine the truth of *their own* counterfactual. Paul might be pretty sure he was not about to become a vegetarian that year, but who knows about the future?[16]

In theory this problem of counterfactual trust exists even in regular trade. How do I know you wouldn't have given me the fruit from your garden if I didn't pay you? Such problems of counterfactual trust are pretty serious in trade between friends as their interests are often sharply aligned. To some extent being your friend *means* that I start to want you to get the things you desire and vice versa. It is perhaps one of the reasons that trade between friends can be difficult to set up. In contrast, there is very little chance that the greengrocer would have given me my vegetables if I hadn't given him some money. Most regular trade therefore comes with a high degree of counterfactual trust.

Counterfactual trust appears to be a big problem for realizing the full gains from moral trade. Unlike for factual trust, contract law seems to be powerless. Playing out the transaction over a long time period seems to exacerbate the issue, rather than fix it: it is harder for either party to predict what someone would be doing a long time hence were it not for this trade. Counterfactual trust is minimized in situations where the ac-

16. This could be a particular challenge if he was known to be interested in moral trade and if many moral trades concerned vegetarianism.

tions required by the trade take place over a fairly short timescale (such as a year), and where the parties are friends or people of known moral character who can trust each others' sincere judgments about what they would have done otherwise.

PRACTICAL CASES

The most well known practical examples of moral trade concern the exchanging of votes. For example, consider the role of an independent senator or leader of a small single-issue party who holds the balance of power. In such a case, it is a commonplace for them to trade their support in passing several of the government's major bills in exchange for them implementing a new policy favorable to the senator's (or small party's) most strongly held view. The different parties involved in the trade consider their relative bargaining power and how much they think is at stake in the different cases, then strike a moral trade, trying to make the overall outcome as good as possible from their own perspectives. This kind of trade during legislation can allow Pareto superior combinations of policies to be agreed upon.

This exchange of votes has also occurred in the general public. In the 2000 and 2004 US presidential elections, a number of websites appeared to enable people in different states to change their votes in such a way that they each thought the outcome was better than the status quo. For example, suppose a person in a safe state preferred Al Gore to Ralph Nader to George W. Bush, while a person in a swing state preferred Ralph Nader to Al Gore to George W. Bush, and that absent any trade they were going to vote for their first choice candidates. This would produce a vote for Nader and a vote for Gore, but they would be inefficiently distributed: the vote for Gore would happen in a state where it would have very little chance of changing the outcome. If they instead agreed to vote for the other person's first choice, the vote for Gore would have more chance of electing him and the vote for Nader would be equally good at helping him win the 5 percent share needed to qualify for federal funding at the next election.

Websites such as Nader Trader and VotePair helped to facilitate such trades.[17] They acted as simple markets for these moral trades, allowing thousands of people to put up offers and find others with matching offers. There were initially serious questions about the legality of such sites (leading several to close down in fear of prosecution), but the services they offer have since been officially declared to be legal, with the US 9th Circuit Court of Appeals declaring, "Whether or not one agrees with these

17. At the time of writing, neither of these is still live, though VotePair is archived at https://web.archive.org/web/20041220193356/http://votepair.org/.

voters' tactics, such efforts, when conducted honestly and without money changing hands, are at the heart of the liberty safeguarded by the First Amendment."[18]

Prior to the 2012 US presidential election, an even more interesting website appeared. Repledge.com was set up to allow people to swap off their donations to their preferred political parties in much the same was as my example concerning gun rights and gun control.[19] Users would donate a sum of money, list whether they were supporting the Republicans or the Democrats, and choose an alternative charity. At the end of the process most of these donations would be matched off against each other and sent to the alternative charities, with the unmatched surplus for the political party that got the most financial support being sent on to that party.

This would be a clear case of a market for moral trade, helping pairs of parties to realize exchanges that they both think make the world a better place. However, the site was taken down before such trades could be made, presumably due to unresolved legal questions arising from the fact that it involved political donations, which are subject to additional rules and regulations.[20]

Hopefully such moral trade markets will be more successful in the future. To the best of my knowledge, nothing would stop someone from setting up a site like Repledge.com dedicated to cancelling out donations which aren't election contributions. For example, a site could be set up to cancel donations between advocates of gun rights and gun control. Creating such markets would help to remove the moral and economic inefficiencies in the charity sector that are created by the existence of charities with opposed missions.

One could even cater for the possibility of different efficiencies for the charities and different beliefs about these efficiencies by getting people to register an additional number when they make their contribution. Each contribution would say something of the form "I'm willing to give up to $6,000 to this charity but am prepared to cancel each dollar against 0.8 dollars pledged to this opposing charity."[21] Such contribu-

---

18. Porter v. Bowen, No. 06–55517 (9th Cir. 2007, 9359).

19. I would like to thank Owen Schaefer for bringing this to my attention. At the time of writing repledge.com is down but is archived at http://web.archive.org/web/20120625133032 /http://www.repledge.com/.

20. Repledge's founder, Eric Zolt, made a request to the Federal Election Commission of the United States to see if it would issue legal advice on some questions related to election contributions regulation, but the Commission's vote was tied 3–3, so it couldn't offer clear advice on the matter.

21. This is equivalent to stock exchanges where people offer to buy or sell some number of shares at some price and mutually agreeable offers are automatically matched up.

tions could then be automatically matched against each other, leaving unmatched some people who were contributing to the more popular charity or who were trying to drive particularly hard bargains. This flexibility would ultimately be useful, especially when the opposed charities are clearly not equally efficient, but it does significantly complicate things and a first effort would probably do better if kept things simple and only offered one-to-one exchanges.

If people found success with relatively constrained cases like these, they could try to extend them to more complete markets. For instance, one could have a moral-trade version of eBay or Craigslist, with people offering and requesting a great variety of moral services with associated prices. Ideally there would be some form of currency to make the trading more efficient. This could be regular currency such as dollars or pounds, or it could be some kind of virtual currency, perhaps one that can only be exchanged for moral services.[22] There are many possibilities, so at this stage it is very difficult to predict how a mature market for moral trade would develop and progress.

And of course there are still the ad hoc arrangements between friends and colleagues that we started with, like those of Victoria and Paul, or Rebecca and Christopher. While the overall scale is smaller, the gains from each individual trade may well be large. In the case of Victoria and Paul, they each thought that by trading they were increasing their personal moral impact by a large proportion. These personal arrangements are particularly well placed to avoid the problems of factual and counterfactual trust. Moreover, one doesn't have to wait for someone to set up a larger enterprise but can start realizing the benefits immediately.

THE GAINS FROM MORAL TRADE

The gains from regular trade have been vast. Without some kind of trade, it is unclear that humans could do much more than merely survive. Our technology, cities, and culture, could not exist without the ability to specialize our labor and this has largely been achieved through trade. We each do the work that other people want done, which gives us a salary to spend on things that enrich our own lives (paying others to do the labor whose fruits we seek). This simple process has enabled the rise of civilization and produced great value.

How large are the potential gains from moral trade? It seems doubtful that they could be on that same scale: partly because having one type of trade is perhaps enough to get the main gains, and partly because of the additional challenges of acquiring the counterfactual trust needed

22. While it has some drawbacks, the latter might help avoid anxiety about trading between sacred and nonsacred values.

for moral trade. However, there is at least some reason to think that the potential gains are large.

How could we quantify these gains? The gains from regular trade accrue to billions of different individuals, but in each case the gains look somewhat similar (improving someone's wellbeing), and we have ways of assessing these gains in aggregate. The gains from moral trade would be much more difficult to assess, as they are gains in how much better the world is, or how morally superior things are from the perspectives of each of the moral views of the trading parties. These are much more diverse than the prudential preferences of different people and it is not clear whether there is any sensible way to aggregate them into something like a total or an average. One could imagine a case where all moral theories agreed that the outcome would be 10 percent better with moral trade than without, in which case we would have a kind of definitive answer (though only expressible as a percentage). However, if they disagreed on these percentages there would be no clear way to weight them to perform an average.[23] More problematically, many theories would completely reject the idea of quantifying the moral difference between the situations.

However, we can at least make some intuitive progress. For example, it seems that moral trade could remove significant inefficiencies in the charitable sector, where opposed charities have budgets totaling billions of dollars per annum. It may also allow us to greatly cut down on the number of people performing acts that don't give them much prudential value, but which some people think are morally terrible (such as purchasing meat that was produced with very bad animal welfare practices), as those who hold these moral views could always offer something of moral or prudential value in exchange.

I should also note that nothing guarantees moral trade will make things better rather than worse. It only guarantees that it will be better according to the views of the parties who make the exchange. Just as regular trade could have negative externalities (making things worse for people who weren't involved in the trade) so moral trade could have negative (moral) externalities: making things worse according to the moral views of people not involved in the trade. If this were sufficiently

23. There are two parts to this. First, it is unclear how to weight the different views in terms of their probability or popularity. Second, even if we could do that, the average of different percentage gains isn't always the midpoint, but depends on what bases they are rising from. This would require a way to compare absolute value differences between moral theories. This problem also arises in cases of moral uncertainty, where it is known as the problem of intertheoretical comparison of value. See Edward J. Gracely, "On the Noncomparability of Judgments Made by Different Ethical Theories," *Metaphilosophy* 27 (1996): 327–32; Ted Lockhart, *Moral Uncertainty and Its Consequences* (Oxford: Oxford University Press, 2000).

common it could overcome the benefits, but if the moral trade is sufficiently engaged in by people of all views then I find this very unlikely. Especially considering that each moral trade is likely to have many positive moral externalities (making things better according to the moral views of people not involved in the trade).

According to moral realism, moral trade could also have negative effects if the trades make things worse overall on the objectively correct moral view—even if no one holds this view. This would also seem to be unlikely. While it is possible according to moral realism for no one to have the correct moral view, it seems likely that our moral views at least cluster near it and that systematically improving the world according to our moral views would be more likely to improve it according to the correct view than to make it worse. Indeed if our views on morality are so misguided that we are systematically pointing away from the objective moral requirements, then we would presumably have bigger problems than moral trade, and it would be difficult to blame moral trade for our situation as in that case practical rationality, moral education, and strength of will, would all be making things worse.[24]

Finally, moral trade could incentivize bad behaviors. For example, someone who thinks that animals have no moral value might start eating meat in the hope that others will pay him to stop. This problem of perverse incentives already occurs in other areas: for instance, insurance makes people more lax about risks since someone else will pick up the bill. If we don't have good ways of dealing with such perverse incentives, they could erode a significant portion of the value of moral trade in some areas. However, there is little fear of such perverse incentives causing moral trade to make things worse rather than better. For if making offers of moral trade in some domain were strongly believed to make things worse overall, then people would simply stop making such offers.

CONCLUSIONS

Trade is everywhere. It is one of the foundations of modern prosperity. However it is usually restricted to cases of mutual prudential advantage. There appear to be great potential gains from moral trade, both in terms of local ad hoc arrangements and larger organized markets. There is also a lot of potential for further philosophical and economic analysis of the theoretical underpinnings of moral trade, exploring how much of the theoretical and practical apparatus of regular trade can be directly brought across and doing the new work required when it cannot. I hope that this essay is just the first step in a much longer journey.

24. The situation would be similar to blaming rationality in a world where we are unfortunately situated such that acting rationally on our beliefs leads to woe.