

The Limits of Commodification Arguments: Framing, Motivation Crowding, and Shared Valuations

Natalie Gold

University of Oxford, Faculty of Philosophy,, Radcliffe Observatory Quarter 555, Woodstock Road,
Oxford, OX2 6GG; natalie.gold@rocketmail.com; ORCID ID orcid.org/0000-0003-0706-1618

Abstract. I connect commodification arguments to an empirical literature, present a mechanism by which commodification may occur, and show how this may restrict the range of goods and services that are subject to commodification, therefore having implications for the use of commodification arguments in political theory. Commodification arguments assert that some people's trading a good or service can debase it for third parties. They consist of a normative premise, a theory of value, and an empirical premise, a mechanism whereby some people's market exchange affects how goods can be valued by others. Hence their soundness depends on the existence of a suitable candidate mechanism for the empirical premise. The "motivation crowding effect" has been cited as the empirical base of commodification. I show why the main explanations of motivation crowding—signaling and over-justification—do not provide mechanisms that could underpin the empirical premise. In doing this, I reveal some requirements on any candidate mechanism. I present a third explanation of motivation crowding, based on the crowding out of frames, and show how it fulfills the requirements. With a mechanism in hand, I explore the type of goods and services to which commodification arguments are applicable. I establish that the mechanism would only enable markets to break down "shared valuations", which is a subset of the valuations that proponents of commodification arguments are concerned with. Further, it can only break down relatively fragile shared understandings and therefore, I suggest, it cannot support a commodification argument regarding the sale of sexual services.

Key words: Elizabeth Anderson, blood, commodification, domino theory, framing, motivation crowding, prostitution, Margaret Radin, shared valuation, Richard Titmuss

Many people have the intuition that there are some goods and services that should not be bought and sold. But there is little agreement on what and why. The *what* and *why* questions are intertwined.

When thinking about why some things should not be for sale, we can contrast essentialist arguments, which hold that it *constitutes* a harm to buy and sell some things, with arguments against trade because of its bad *effects* (Satz, 2010). Essentialist arguments capture the idea that market exchange degrades or *corrupts* some goods and practices (Sandel, 2012). To take the example of prostitution, an essentialist argument against prostitution is that it is inherently degrading to the prostitute, a non-essentialist argument is that prostitution reinforces gender inequality. It is clear that non-essentialist arguments about *what* should not be for sale rely on the existence of corresponding mechanisms that show *why* sales have bad effects.

However, even essentialist arguments may rely on the existence of mechanisms. Some corruption arguments go further than the idea that trade is corrupting for the buyer or the seller, asserting that that some people's trading a good or service can debase it for third parties (Anderson, 1993; Radin, 1987, 1989). To take the example of prostitution again, the idea is that some people's buying and selling sex degrades what is valuable about sex for everyone. The overall shape of the argument is that if a good or service is treated like a commodity by some people, then it takes on the characteristics of a commodity for everyone, so let us call these *commodification arguments*.

Commodification arguments combine elements of the essentialist argument, about the inherent negative effect of sales, with an argument about its effects on third parties. Another way of putting this is that they consist of a normative premise and an empirical premise (Radin, 1989, p.173). The normative premise is a theory of value according to which it is important that the non-market version of some goods and services is available. The empirical premise is that, once some people treat a good or service as a commodity, then it is debased for everyone. Some people's behavior affects other people's valuations. The soundness of commodification arguments depends on the truth of the empirical premise, on the existence of a mechanism that would cause market exchange to drive out

moral and civic values for third parties.

A mechanism that supports commodification as an answer to the *why* question will also have implications for the answer to the *what* question. We buy and sell many goods and services, and it is not plausible that the value of all these things is corrupted by the market. If we can establish a mechanism by which commodification might occur, then we can assess what goods and services may be vulnerable to that mechanism and, hence, to commodification. I will present a mechanism by which commodification may occur, an exercise in moral psychology that answers the *why* question, and then show that the mechanism has implications for *what* goods and services may be commodified and therefore has ramifications for political theory.

There has been remarkably little detailed work on mechanism. Michael Sandel (2012) cites experimental evidence of *motivation crowding*, which shows that paying people can have perverse effects on their behavior, as support for the idea that markets drive out non-market norms and values. However, this evidence is about changes in behavior, not valuations. If motivation crowding is to support the empirical premise of commodification arguments, then the change in behavior needs to be connected to changes in valuations. Further, motivation crowding experiments show that paying someone can change *that person's* behavior, whereas commodification arguments are about the effects of market exchange on third parties. There is no empirical literature on motivation crowding in third parties, which is what commodification arguments require.

A lot of weight rests on motivation crowding, so it is appropriate to investigate how successful it is as a mechanism of commodification. I will show that the two main families of explanation for motivation crowding, based on signaling and over-justification mechanisms, cannot support the empirical premise. In doing this, I will reveal some further requirements on any mechanism that purports to underpin the empirical premise. I will then present an alternative explanation for motivation crowding that relates it to *framing*, or how people conceive of their situation, and show how this could support the empirical premise. With a mechanism that could underpin the empirical premise in hand, I assess the extent of the support that it provides to commodification arguments. Granted that commodification arguments are an answer to the *why* question, I explore the implications of the mechanism I present for the answer to the *what* question. The mechanism puts conditions on the type

of goods and services that are subject to commodification, the scope of commodification arguments is limited to goods and services that satisfy these conditions, which may exclude some popular targets.

My aim is to assess the maximum extent of support for the empirical premise of the commodification argument, by presenting a mechanism that could underpin it. I do not claim that it is the only possible mechanism. However, I do not think that there are any other promising candidates for the empirical premise, so I think that proponents of commodification arguments will have to accept something like my mechanism if their arguments are to be sound. If a reader should think the mechanism implausible or incorrect, then so much the worse for commodification arguments.

1. Commodification arguments, the empirical premise, and Titmuss's puzzle

The best developed commodification arguments are due to Elizabeth Anderson (1993) and Margaret Radin (1989, 1987), and their arguments have the same structure: a normative premise containing a theory of value, and an empirical premise about contagion.

Anderson (1993) provides a pluralist theory of value, where there are different modes of valuation that are appropriate to different kinds of goods. She distinguishes two types of goods: *intrinsic goods*, which are the immediate objects of our valuations, and *extrinsic goods*, which are those things that it makes sense to value only because one values some other thing. One subset of extrinsic goods is those that have *use value*, in that they are only valued as a means to some other end.¹ Money, and many of the things that are bought and sold on markets, have use value. Anderson claims that use value is an inferior mode of valuation to valuing goods intrinsically.

Anderson supplements this normative premise with a second, empirical premise: that putting goods up for sale debases their intrinsic value. She does not provide a general mechanism by which commodification occurs, but she takes prostitution as the “classic example” (1993, p.154). Anderson claims not only that commodified sex is degraded and degrading to the prostitute (a general corruption argument), but that there are also “deep connections between the ways women’s sexuality is valued by men in both [the market and personal] spheres” (p.155), so that if women’s sexuality is legally valued as a commodity, then it may be difficult “to establish insulated social spheres where it can be exclusively and fully valued as a genuinely shared and personal good” (p.155).² Therefore, Anderson thinks that

some markets need to be restricted for the sake of freedom and autonomy, in order to promote access to a wide range of significant options and to preserve superior modes of valuation.

Radin, like Anderson, draws on the Kantian notion of people as ends in themselves, but she puts the emphasis in a different place, on the distinction between persons and things. Radin claims that some goods and services are integral to the self, they are so important for our “personhood” and for human flourishing that to understand them as monetizable and completely detachable from the self “is to do violence to our deepest understanding of what it is to be human” (1987, p.1906). Radin does not provide a general theory about which goods and services are integral to personhood, but she gives babies and sexual services as obvious examples.

Radin’s (1989) empirical premise is that the commodified and non-commodified versions of some interactions cannot co-exist, so commodifying something can prevent its non-commodified analogue from existing. She says that “Once something has a price, money must be a part of the interaction, and the reason or explanation for the interaction, when that something changes hands. A sale cannot simultaneously be a gift.” (p.175). Radin illustrates the idea with respect to blood. (This is despite the fact that she is most interested in the market for sexual services and is ambivalent about whether her argument applies to blood.) She says that if the price of a pint of blood is \$50, then giving a pint of blood becomes like giving \$50 of money; anything special about giving blood rather than money is destroyed (p.170).

Eric Mack (1989) has been skeptical about the empirical premise. In particular, he targets Radin’s example of how giving a pint of blood becomes like giving \$50 of money. Mack argues that the value that he puts on gifts is not the same as their price and he says that anyone who equates the value of a gift with its price is making a mistake. He alters the example, so that the item being given is a Christmas gift of a sweater to his wife, and considers why he would rather give her a sweater than its cash value or its value in gift vouchers. He concludes that giving a sweater is not the same as giving money. This is plausible, but it is not my focus here. Mack then turns the question around, asking: if non-market relations are so much better than commodified ones, why don’t people continue using non-commodified versions when markets are introduced? By the end of this paper, I will have provided an answer.

Mack may be right to be skeptical about Radin's example. However, that does not resolve the argument about the empirical premise which, after all, is an empirical question about whether, in Anderson's language, selling a good or service can debase its value for others, or, in Radin's language, commodified and non-commodified versions of a transaction can coexist, and, for both authors, whether some people's market transactions (behavior) can affect the ways in which others can value those goods and services. But let me suggest another source of skepticism that arises from considering Mack's example. Like the sweater, the gifts that we give are often commodities that can be bought. Sometimes we can even guess the price of a gift. However, it doesn't devalue Mack's gift to his wife if she knows that he bought it in a shop and that it cost \$50. She may well prefer to receive the sweater from him than a \$50 note. Radin (1987, p. 1912) says that commodification is a "slippery slope" and uses the metaphor of dominos, where pushing over the first domino causes the rest to topple, but it is not credible that any step towards commodification of any good must lead to total disaster. The sweater is partially commodified and most people do not find that to be a problem. To give Anderson and Radin due credit, both recognize the possibility of partial commodification or, as Anderson puts it, "mixed goods". However, the empirical premise is not plausible unless there is a mechanism that can explain why some goods become devalued by partial commodification but not others. Why are gifts of blood commodified but not gifts of sweaters? A suitable mechanism should answer this question. And can the answer tell us anything about what other goods and services are likely to be commodified? We might hope that the mechanism would also give us some guidance about what properties of goods or their social context makes them vulnerable to commodification. In underpinning the *why* question, it should put boundaries on the *what* question.

It is no coincidence that Radin used an example involving paying for blood. One of the earliest commodification arguments is found, at least implicitly, in Richard Titmuss's (1970) *The Gift Relationship*. Titmuss presented a case study of the methods for the provision of blood in several countries, in particular the UK and the US. Both these countries collected blood through voluntary donations but, in the UK, the entire blood supply was obtained from donations, whilst the US at that time also operated commercial blood banks, which paid people for their blood. Titmuss argued that the market for blood in the US discouraged voluntary donation, and that less blood was supplied than would have been the

case in a purely voluntary system. He also conducted a survey of donors' motivations. Titmuss concluded that the result of paying for blood was that, "Both the sense of community and the expression of altruism are being silenced" (p.157).

Although he cited evidence from national data on blood collection, Titmuss was criticized by two Nobel Laureate economists for not providing an explicit mechanism that would cause these effects. Robert Solow (1971, p. 1696) thought that "the story seems far-fetched" and Kenneth Arrow regarded it as "speculative thinking" because of the concentration on empirical evidence rather than theoretical analysis (Arrow, 1972, p.143). Titmuss left a puzzle: why introducing a market alongside a previously existing non-market mechanism would have caused donations to decrease or, more generally, why some people engaging in market exchange affected third parties' behavior. In honour of its origins, let us call this question about the behavioral mechanism *Titmuss's puzzle*.

Titmuss's puzzle has a parallel structure to the empirical premise of the commodification argument. Both posit that there will be a spillover effect of markets on non-market participants. Titmuss's puzzle is about a spillover effect of behavior: some people's buying and selling affects others' behavior. The empirical premise is about a spillover of behavior on valuation: some people's buying and selling affects others' valuations. The empirical premise seems more intractable than Titmuss's puzzle. After all, there are many ways in which some people's behavior can affect other people's behavior, starting with simple imitation. But people have thought that Titmuss's puzzle and the empirical premise are connected, and Titmuss's own conclusions about his findings suggest that the change in behavior corresponds to a change in valuations (although Titmuss thought that what matters about blood-donation is that it is one area where we can engage in anonymous altruism, so his argument differs from commodification arguments since he does not claim that there is some intrinsic feature of blood that means it should not be commodified, but rather that there should be some areas of social life where such altruism is possible). I will go on to present a mechanism that can solve Titmuss's puzzle and show how it can also underpin the empirical premise.

2. The motivation crowding effect: an empirical effect in need of a theoretical explanation

Michael Sandel (2012) draws on examples of motivation crowding to show that market values can crowd out nonmarket norms, and many people think that motivation crowding can solve Titmuss's puzzle (see Frey & Jegen, 2001 and references within). The *motivation crowding effect* is a behavioral phenomenon first documented by Edward Deci (1975, 1971), who found that subjects who had been paid to solve puzzles were less likely to return to them later, after payment had been withdrawn, than a control group who had received no reward for their activity in the first period. The paid subjects also reported a lesser interest in the task than the unpaid. These results have been replicated in numerous other experiments, in the laboratory and in the field. (For literature reviews and meta-analysis see Cerasoli, Nicklin, and Ford, 2014, Frey and Jegen, 2001, and Deci, Koestner, and Ryan, 1999.)

Deci hypothesized that the extrinsic motivation of the reward replaced the subjects' intrinsic motivation, where "[o]ne is said to be intrinsically motivated to perform an activity when one receives no apparent reward except the activity itself." (Deci, 1975, p.175) So extrinsic motivation is doing something for an apparent reward and intrinsic motivation is simply defined in contrast to extrinsic motivation, as anything that is not done for a tangible reward. Intrinsic motivation includes many different sorts of reasons for undertaking an activity including, for example, enjoyment of a task, other-regarding reasons, or moral reasons.

Deci's hypothesis is simple and powerful. Simple because it postulates that, since people no longer do the task for free, they are no longer intrinsically motivated. Powerful because of the range of phenomena it can potentially explain. The motivation crowding effect has been used in models to explain why increasing incentives in the workplace will not necessarily increase output (Kreps, 1997), why countries with stricter constitutions would have more tax evasion (Frey, 1997a), and why increasing punishment does not always decrease crime (Akerlof & Dickens, 1982). But a complete theory will need to say more about why people cease to be intrinsically motivated.

Deci's hypothesis captured the imagination. However, to prevent ambiguity, note that I use "motivation crowding" and "motivation crowding effect" to refer to the body of empirical results and not to any specific hypothesis about what causes the effect. (In fact, one might note that Deci himself drew on more detailed theories to explain the effect, from Cognitive Evaluation Theory in his 1975 book to the broader Self-Determination Theory in Deci and Ryan, 2000.) One might see the

mechanism I will present as elaborating on the hypothesis that extrinsic motivation crowds out intrinsic motivation, but if one prefers to think of it as an alternative hypothesis, that will not have any bearing on the argument of this paper.

Deci argued that “Every reward (including feedback) has two aspects, a controlling aspect and an informational aspect” (1975, p.42). Deci thought that the informational and control aspects would work in opposite directions, with payment as a signal of success in the task, which would counterbalance the negative effect of its controlling aspect. But nowadays, both aspects are seen as having a negative effect, and each of them has formed the basis of a family of theories about the operation of motivation crowding: that introducing payment “over justifies” the agent, which causes her to change her preferences and hence behavior, and that payment is a “signal” which changes information and hence behavior.³ In this section, I assess to what extent each of these theories can provide a mechanism that would explain Titmuss’s puzzle and that could underpin the empirical premise. The reader may be able to think of other candidates for a mechanism, but these are the two that have garnered most attention in the empirical literature, so I focus on them.

One mechanism that has been proposed to cause the motivation crowding effect is based on the control aspect of rewards. Amongst psychologists, the predominant explanation for motivation crowding is the *over-justification* of the agent, where the payment is seen as controlling, and the external intervention therefore undermines feelings of self-determination and autonomy, which causes the agent to relinquish the intrinsic motivation (Deci and Ryan, 2000; Ryan and Deci, 2000).⁴ This is compatible with the classic experiments, where agents do not get to choose whether or not they are paid. In contrast, in the case study presented by Titmuss there were two systems (the market and the voluntary system) running side-by-side and, if the agent has a choice between receiving payment or not, why should feelings of self-determination or autonomy be impaired?

The idea that having two systems will increase choice suggests another way of putting Titmuss’s puzzle. If we grant that there is a variety of types of agent in the population, then we might expect that having multiple mechanisms for collecting blood would lead to a higher supply because different types would be induced to give blood by different collection systems (Frey, 1997a). The fact that Titmuss found that blood collected in the market regime was more likely to be infected with hepatitis is a sign

that payment attracts a different group of people. The puzzle is why payment would have caused the previous group of voluntary donors to withdraw. A basic tenet of microeconomics is that increasing the range of alternatives leaves an agent at least as well off because if she prefers one of the new alternatives she can switch and, if not, it is still open to her to make the same choice as before. This is exactly what Titmuss denies. He claims that, although it allows some men to live as they like, the market “narrows the choices for all men” (Titmuss, 1970, p.310-11). For Titmuss, it is not just that there have been changes in motivation, but also that some options have been eliminated.

Standard examples of motivation crowding are intra-personal, payments to an agent affect *that* agent’s behavior. Titmuss presented a more complicated, inter-personal puzzle—payments to one agent affect *another* agent’s behavior. Frey (1997a) suggests that there will be a “spillover effect” between the two systems, but he does not elaborate. A satisfactory explanation of Titmuss’s puzzle will have to include an explanation of the mysterious spillover effect.

Another family of explanations understands motivation crowding within the signaling literature, where the use of financial incentives changes information about the agent or about the task. (So, although signaling theories purport to explain the motivation crowding effect, they jettison Deci’s hypothesis about the driving out of intrinsic motivations.) One signaling explanation of motivation crowding is based on the idea that someone may engage in an activity in order to give others information about herself. For instance, if there are different “types” of agent in the population, some who are altruists and some who are not, then engaging in a civic activity may be a signal that a person is an altruistic type. Paul Seabright (2002) and Roland Benabou and Jean Tirole (2006) propose game-theoretic models where altruists give blood as a signal of type in anticipation of later games where they want to be able to match up with other altruists. In these models, offering payment for blood donation encourages non-altruists to donate and makes the signal “noisy”. Since donation is done by both altruists and non-altruists it is no longer a reliable indicator that someone is an altruist. This makes it less useful as a signal, so altruists donate less than before payment was introduced. The participation of non-altruists spills over and affects the behavior of altruists. However, as well as thinking that it is a strange sort of altruist who only donates blood as a signal, we might doubt that this mechanism operates in Titmuss’s puzzle. We generally do not know of any particular person whether they have

given blood, and there is no public register we could check, so it is hard to see how agents could use it as a signal of their types.

Alternatively, offering payment may send a signal *to* someone, either about herself or about the task. Benabou and Tirole (2003) have proposed that, in a workplace environment, an employer who offers incentive payments may send a negative signal about the worker's ability or about the disutility of the job. With regard to Titmuss's puzzle, an information asymmetry where the donor does not know some piece of information about herself that the authorities do know is not relevant (and Benabou and Tirole (2006) acknowledge this).⁵ An information asymmetry where payment is a signal of the unpleasantness of the experience might be relevant. It could explain why payment would deter potential first time donors. However, blood is mainly given by repeat donors (Healy, 2000) and, since the payment would have to convey information that the agent does not already have, the theory that payment is a signal of unpleasantness cannot explain why these donors would stop giving. For the same reason, it cannot explain the results of the standard motivation crowding experiments, where people are already engaging in the task in the first period and the withdrawal of payment causes them to lose interest.

The most plausible signaling explanation for Titmuss's puzzle is that donors are sending signals *to themselves* (Benabou and Tirole, 2006). On this account, the agent has a set of preferences or values, and a set of second-order preferences about what values she has or what type of person she is. At the time of taking action, the agent knows her preferences, but after taking action, she only remembers the action and not the underlying values. In this situation, the agent may choose actions partly in order to manipulate her self-image at the later time. Applied to Titmuss's puzzle, the idea is that donating blood is a sign that the self is an altruist, so people give in order to self-signal as altruists. However, if an unavoidable payment is introduced, then the signal is muddled—maybe the earlier self was motivated by the payment?—so donation done for self-signaling purposes decreases. This is slightly different to Titmuss's puzzle, since the fact that the payment is unavoidable is an important part of the explanation: if there were two systems, the voluntary system and the market, then one could self-signal by making a the voluntary donation instead of taking payment. In any case, it seems unlikely to be of help with the empirical premise, which requires that some people's actions affect how goods are valued by others.

The account is a model of self-deception (Mijović-Prelec & Prelec, 2010); self-signaling donors are not really motivated by altruism and when they stop donating their values do not change. This reveals an underlying problem with all the signaling explanations: even if they explain Titmuss's puzzle, that explanation will not port to the empirical premise because they do not involve any change in valuations. Hence, even if you find signaling explanations persuasive as a solution to Titmuss's puzzle, they do not provide a foundation for the empirical premise of commodification arguments.

This discussion has revealed that an explanation of motivation crowding needs to account for several different ways that payment affects behavior. There is an effect of withdrawing payment, as in Deci's experiments. There is an effect of the introduction of payment on agents already engaged in an activity, as in Titmuss's puzzle. And there is a spillover effect, whereby payment may affect behavior even when a non-payment system is maintained alongside the market. Further, an explanation of motivation crowding that would suffice to underpin the empirical premise must relate behavior change to value change.

3. The connection between framing and motivations

Lindenberg and Frey (1993) claim that when motivation crowding occurs a "gain frame" crowds out a "normative frame", but this is not explained in any further detail. Framing is often implicitly and sometimes explicitly offered as an explanation of the effect of payments on effort (e.g., Gneezy & Rustichini, 2000; Heyman & Ariely, 2004). But there is no detailed conceptual explanation of the framing effect on first parties, let alone third parties as required by the empirical premise.

We can think of the agent's frame as the set of concepts that she uses to think about her situation (Bacharach, 2003). Framing is notorious because of Tversky and Kahneman's (1981) work on *framing effects*, where two groups of subjects were given the same options, described differently, and the description affected the choices that their subjects made. In Tversky and Kahneman's original experiment they asked subjects to imagine that the US was threatened by a disease that was expected to kill 600 people, and that they had to make a choice between two alternative vaccination programmes. Half the subjects were told the options in terms of how many people would be saved (200 people saved vs. a 1/3 probability that 600 people would be saved), the other half were given the same options

phrased in terms of how many people would die (400 people die vs. a $2/3$ chance that 600 people would die). Since the subjects know that 600 lives are at stake, one of these pieces of information should suffice. However, changing the description of the options from one in terms of “lives saved” to one in terms of “lives lost” changed the modal choice: in the first group 72% of subjects chose the program that would save 200 people (with 400 dying), in the second, 78% of subjects chose the program with a $2/3$ chance that 600 would die (and a $1/3$ chance that 600 would be saved).

In Tversky and Kahneman’s framing effect, there were two policies to choose between. The outcome of a policy choice had two aspects: number of people who would be saved and number of people who would die. Even if people are only explicitly told only one of these consequences, we expect them to be able to work out the other. However, changing the set of concepts in terms of which the decision was presented, from “save” to “die”, affected choices. I will give two examples that connect framing effects and motivations, in order to motivate a treatment of motivation crowding using some of the ideas that have been brought to bear on framing effects.

My first motivating example is a real-life policy decision, which seems to take the possibility of motivation crowding seriously and whose structure is analogous to Kahneman and Tversky’s framing effect. The economic analysis of crime suggests that the amount of criminal activity is inversely related to the severity of punishment and the probability of getting caught (Becker, 1968). In contrast, a standard example in the motivation crowding literature is that increasing punishment does not necessarily decrease crime. In Melbourne in 2002, Yarra Trams ran a campaign against fare dodging. As a part of this campaign, they instituted a system of random inspection and on-the-spot fines. But this was not the way they presented it to the public. The slogan for the campaign was “Fare evasion is stealing”.⁶ Instead of emphasizing that the chance of being caught and punished had increased, Yarra Trams appealed to a form of intrinsic motivation, with no official sanctions or incentives. So Yarra Trams behaved as though they were anticipating a motivation crowding effect. The potential motivation crowding effect also looks like a framing effect. Deciding to evade one’s fare is a choice whose outcome has two aspects: having done something dishonest, and incurring a risk of being caught and punished. As in Kahneman and Tversky’s experiment, even if people are only explicitly told one consequence we

expect them to be able to work out the other. But Yarra Trams chose to emphasize only one of the consequences, so it looks like Yarra Trams was taking advantage of a framing effect.

		Player 2	
		C	D
Player 1	C	12,12	0,18
	D	18,0	6,6

Figure 1. Prisoner's dilemma

		Your Gains	Other's Gains
		C	D
Your Choice	C	0	12
	D	6	0

Figure 2. Decomposed prisoner's dilemma

My second motivating example is a framing effect that researchers have hypothesized is caused by a change in motivations. It involves giving subjects different presentations of the same prisoner's dilemma matrix, known as *decomposed games* (Messick & McClintock, 1968; Pruitt, 1967).⁷ The most familiar presentation of games uses a standard payoff matrix. Figure 1 gives an example of the prisoner's dilemma payoff matrix. Each player must choose between *C* and *D* and their payoffs depend upon the combination of their choices. Whatever the other player does, each player is better off choosing *D* (which gets 18 rather than 12 if the other chooses *C*, or 6 rather than 0 if the other chooses *D*). However, if both players choose *D* they each get a payoff of 6, whereas if they had both chosen *C* they would each have got 12. It may be possible to decompose a payoff matrix so that, instead of presenting the payoffs in a matrix form, the players have to choose one of two allocations of payoffs between the players. Figure 2 presents one possible decomposition of the "parent game" of Figure 1; each player knows that the other also faces the same allocation decision. Calculating the payoffs for each player that would result from the four different possible combinations of the

allocation decisions (CC, CD, DC, DD) gives the outcome matrix of the parent game. For instance, in the decomposed version, if Player 1 chooses allocation C and Player 2 chooses allocation D then Player 1 has assigned 0 to herself and 12 to Player 2 whilst Player 2 has assigned 6 to herself and 0 to Player 1. So Player 1 gets 0 from herself and 0 from Player 2, a total of 0. Player 2 gets 12 from his choice and 6 from player one, a total of 18. The outcome is 0 for Player 1 and 18 for Player 2, as can be read off from the parent game. In any decomposed game, it is possible to work out the payoff matrix from the choices in the allocation decision.⁸ The decomposition and the parent game are two different ways of presenting the four possible payoff outcomes.

Dean Pruitt (1967) gave one group of subjects the parent matrix shown in Figure 1 and another the decomposed version of the game shown in Figure 2. Each pair of players played a series of twenty games. Although the decomposition and the parent game are two different ways of presenting the same information about end monetary outcomes, Pruitt found that the rate of co-operation differed in each presentation. In the parent game, the “[p]ercentage of cooperation remained at a relatively constant level” whereas, in the decomposed game above, it “rose precipitously and remained at a high level” (Pruitt, 1967, p.24). The decomposition increased cooperation. This effect has also been found in 3-person prisoner’s dilemmas (Komorita, 1987) and 4-person prisoner’s dilemmas (Cookson, 2000). In a further investigation designed to discover why behavior was different in the decomposed games, Pruitt asked subjects to record the thinking behind their decisions (Pruitt, 1970). He discovered that, in accordance with expectations derived from game theoretic reasoning, those who played D in the above games were motivated by the payoff they could get by doing so. In the decomposed game, responses to open-ended questions showed that many subjects viewed alternative C as a way of being “helpful” or “generous”. Pruitt (1970, p.235) postulated that “the games produce differing motives, which in turn produce differing behavior”, a suggestion that has also been made by Colman (1995) and which indicates a connection between framing and motivations.

Kahneman and Tversky (1979) explained their framing effect using prospect theory. According to prospect theory, decision makers are risk averse over gains but risk loving over losses, so it matters whether they perceive outcomes as gains or losses. Whether an outcome is perceived as a gain or a loss may depend on the decision-maker’s reference point, which may be affected by the way a decision

problem is framed. If “save” connotes rescuing someone who would otherwise be dead, then the reference point is 600 people dead and the choice is over different gains; people are risk averse, going for the smaller number of sure dead rather than gambling and trying to save everyone at the risk of losing all 600. If “die” connotes losing a person who is currently alive, then the reference point has 600 people alive and the choice is over losses; people are risk loving, and they take the gamble. However, it is hard to see how prospect theory could explain the framing effects in my motivating examples.

Decision theorists have given explanations of framing effects that relate them to reasons (Dietrich & List, 2016; Weirich, 2010; Gold & List, 2004; Schick, 2003). What the different models have in common is that the reasons that underpin an agent’s choices depend on how they frame or, in the case of Schick (2003), “understand” the decision. Note that acting and choosing for a reason does not have to be understood as involving a conscious reasoning process. A minimal requirement is that the agent is disposed to be responsive to reasons, where these are based on facts that count in favor of a particular decision or action. In this paradigm, framing effects occur when there are reasons in favor of both options and the reason that the agent responds to depends on the way in which the decision is presented or described. In effect, these agents are not weighing all their reasons, but act on the basis of a single reason. If they have an acceptable reason to hand, then they do not search for others. This has psychological plausibility. It is consistent with evidence of “concrete thinking”, whereby decision-makers appear to use only surface information and information that has to be inferred from the display or created by some mental transformation tends to be ignored (Slovic, Fischhoff, & Lichtenstein, 1988; Fischhoff, Slovic, & Lichtenstein, 1978). Concrete thinking may be connected to people’s desire to justify decisions by saying that they chose for a (single) reason, even to the extent of constructing and selecting choice situations such that there is always a dominant reason for choice (Montgomery, 1983). Once a reason for choice has presented itself, people are not motivated to seek out further reasons. Call this “one-reason decision-making”.

In Tversky and Kahneman’s problem, the fact some people will die for sure is a reason not to choose the policy with certain outcomes, while the fact that there is a possible outcome where no-one is saved is a reason not to choose the risky policy. The two different ways of framing the decision make these different reasons salient, which affects people’s choices (Gold & List, 2004). This explanation is in

accordance with the psychological literature on “reason-based choice”. A classic example there is the custody decision, where the question of which parent *should* get custody elicits the same answer as the question of which parent *should not* get custody: the questions elicit a search for positive and negative attributes respectively, which would be reasons for giving or not giving custody, and one parent has both more positive and more negative attributes (Shafir, Simonson, & Tversky, 1993).⁹ The idea that the presentation of the decision affects the reason that people act on can explain the framing effects in my motivating examples. Reasons are connected to motivations. We can think of the reason for which an agent acts as her motivating reason, so framing can affect an agent’s motivating reason.

In the Yarra Trams example, the fact that fare evasion is dishonest is a reason to buy a ticket. However, there is another way of looking at the decision, the one that is assumed in the economic analysis of crime, namely that not buying a ticket is a gamble with a large probability of saving a small amount of money and a small probability of paying a fine if caught. In that case, someone who is moderately risk loving might have a reason to take the gamble rather than incur the sure loss of buying a ticket. “Fare evasion is stealing” frames the decision whether or not to buy a ticket as one of honesty, rather than gambling. This idea that this would affect decisions is supported by evidence from the laboratory. James Baldry (1986) gave one group of subjects an income and the option to lay a series of small bets. A second group were given the same income and gambles but the experiment was presented as completing a tax return, with the gambles as small tax evasion decisions, where there was a probability of random audit and financial penalties any time a subject was caught evading. Of course, they were still really laboratory gambles. However, the behavior of the two groups was strikingly different. In the gambling condition, every subject laid the maximum possible bet on every round but, in the tax evasion condition, subjects tried to evade the maximum possible amount in only 38% of decisions, and in 31% of decisions there was no evasion attempt at all. Baldry concluded that the tax evasion scenario had triggered “moral compunctions”. People may be willing to take small gambles, but not when they are framed as a matter of obeying the law, so it matters whether Yarra trams framed the fare evasion decision as a moral decision or as a gamble.

In the decomposed prisoner’s dilemma, game-theoretic reasoning about monetary payoffs conflicts with being helpful or generous. This is sometimes referred to as “might versus

morality” (Liebrand et al, 1986). According to Pruitt (1970) and Colman (1995), the decomposition makes helpfulness, or the moral side of the coin, more salient. Their suggestion is supported by evidence that the way that subjects frame the prisoner’s dilemma correlates with the move they make. Subjects who perceive playing C as co-operative and playing D as non-co-operative are more likely to play C (Baranowski, & Summers, 1972). Similarly, co-operative types (defined as such because they behave co-operatively) tend to frame the dilemma in terms of morality (Liebrand et al, 1986). If moral reasons support a different choice from game-theoretic dominance reasoning and the salience of these reasons can be affected by the presentation of the decision, then people will make different choices in different frames.¹⁰ The reason-based explanation of framing effects is consistent with the evidence of a connection between framing and behavior in prisoner’s dilemmas, but refines it by offering a direction of causality, namely that framing the game in moral terms may lead to co-operative behavior by increasing the perception of, and hence the chance that people act on, moral or other-regarding reasons.¹¹

Relating choices to motivating reasons also provides a connection between valuation and behavior, which I identified above as one of the requirements of a mechanism that could underpin the empirical premise. When an agent acts, we can say that she has a pro-attitude that either is or is based on her motivating reason (depending on our precise technical philosophical take on “motivating reason”).¹² Most philosophers would agree that these pro-attitudes are grounded in considerations about the goodness or value of what is desired. Therefore, when an agent’s motivating reason changes, so does the value that she acts on. Framing can connect motivations and values.

In this section, I have connected motivation crowding to framing and reasons, and therefore to valuations. A caveat is due. I gave two examples that look like both motivation crowding and framing effects. However, in most examples of motivation crowding there is an actual change in the situation, as well as a change in framing, so they are not framing effects.¹³ Framing effects are generally thought to be irrational because a change in description leads to a change in choice without any change in the underlying situation.¹⁴ I am not claiming that all motivation crowding effects are framing effects and I do not mean to make any claim about the rationality of motivation crowding. I am concerned with framing rather than framing effects. The idea is that the process that underlies framing effects, the

process of *framing*, whereby changing the way that people frame a problem changes their motivating reason, can operate in other choice situations, including ones of motivation crowding.

4. A framing explanation of the motivation crowding effect

The paradigm examples of motivation crowding do not involve changes in explicit descriptions. However, the monetary payment may still affect the way subjects frame the situation. Take Deci's (1975, 1971) experiments, where subjects were given puzzles to solve. There were two periods in this experiment. In the second period, subjects were left alone in the room with the puzzles. This was the same for all subjects. In the first period, half of the subjects were paid to solve puzzles and half of the subjects played with them without payment. Deci found that first period activity affected second period behavior, even though naïve theory suggested that the second period was the same for all subjects. The first period activity may have served as an implicit framing task. The puzzles were supposed to be interesting to solve for their own sake. Subjects who were paid were given another way to think about the puzzles: as an activity engaged in to make money. In the second period, the monetary payment was withdrawn. If the subjects who had been paid framed the task of solving them in terms of money, and acted on their monetary motivations, then their reason for solving the puzzles would have gone. Concrete thinkers, who do not generally search for information, would not investigate whether there were other reasons to carry on solving the puzzles.

Deci's experiments involved the withdrawal of payments. Above I said that an explanation of motivation crowding must also explain why the introduction of payment would affect the behavior of agents who are already engaged in a task. When an agent is performing a task that she has intrinsic reasons to do and she is also being paid, then her action is over-determined. There is a sense in which she is over-justified—because she has multiple reasons in favor of her action, not because the price is seen as an instrument of control. If people are one-reason decision-makers, then one of the motivations will become the primary motivating reason, at the expense of any others. In the puzzle solving example subjects started off being paid, so there is reason to think that monetary reasons are salient and that the non-monetary reasons may never strike them. But when payment is introduced,

subjects are already performing the task based on non-monetary motivations. Why should the monetary rather than the non-monetary reason become the motivating reason?

In order to answer this, we can draw on findings from social and cognitive psychology. From social psychology, we know that the salience of the reward affects motivation crowding. Ross (1975) has shown that a highly salient reward is more detrimental to intrinsic interest than the same reward when it is relatively non-salient. He also showed that reward is less detrimental when the subject's attention is distracted from it. Payments may be salient, especially in workplace examples, where there is a series of incentive payments over time. This may keep attention focused on the payment, reinforcing the monetary motivation. The idea that monetary reasons will come to dominate intrinsic ones is also supported by *attribution theory*, according to which actors are more likely to attribute their behavior to external factors than internal ones (Jones et al, 1972; Heider, 1958). So attribution theory would predict that, if agents are offered payment, then they will attribute their motivation to the payment, rather than any intrinsic motivation they may also have had.

We can also draw on studies of concept usage from cognitive psychology to understand how monetary motivations can crowd out non-monetary ones. Researchers have studied the effect of *priming*, or activating particular representations just before carrying out an action or task. Evidence from the priming paradigm is that frequent use of a concept increases its *accessibility*, defined as the “activation potential of available information” (Higgins, 1996). And, if a concept is not used then, as time passes, it decreases in accessibility. So even if, at one point, the agent is aware of both monetary and non-monetary concepts, and hence monetary and non-monetary reasons, the evidence suggests that a stream of payments can reinforce the monetary way of framing the situation at the expense of the non-monetary one. (An even stronger claim than this is made by Bruner (1957), who postulates that once an agent has categorized a situation, incongruent cues may be “gated out”).¹⁵ This mechanism explains the contention of Lindenberg and Frey (1993), that a “gain frame” will “crowd out” other ways of framing the task. Once financial gain is the motivating reason, then it becomes likely that withdrawal of payment leads to cessation of the activity, by the mechanism described above.¹⁶

Above, I connected reasons to valuations, with reasons being grounded in considerations about the goodness or value of an option. Therefore, an agent's motivating reason reflects the values that

ground her choice. When an agent's motivating reason changes, so does the value that she acts on. As a reason that is not acted on falls out of the picture, so may the values that it is based on. If a gain frame crowds out a normative frame, then the way that the agent values the activity may change too.

So far, I have shown how the framing mechanism can account for the effect of withdrawing payments and the effect of introducing payments, and that it relates behavior change to value change. This is enough for the mechanism to explain how general corruption—in Sandel's (2012) sense—may occur. But it is not yet sufficient to explain Titmuss's puzzle or the empirical premise.

5. Explaining the spillover effect: the breakdown of shared valuations

Standard examples of motivation crowding involve the introduction and withdrawal of payments *to the same agent*. Titmuss's puzzle contained an extra complication, the spillover effect, where payments to one set of agents affected other agents' behavior.

In the paradigm framing effects, a label is introduced or emphasized and that affects behavior. This is also an aspect of Titmuss's puzzle. The introduction of payment for blood in the US introduced both a new action, selling blood, *and* a new label: the act previously described as “not giving” could also be described as “not selling”. The new label is important. An agent might think that one ought to give blood, but that it is fine not to sell it. Therefore, if people are induced to change the way they describe the actions, from giving to selling, they may change their behavior.

Why might agents change the way they describe the action? In the standard motivation crowding examples, I argued that a salient way of framing the situation, sometimes caused by repeated use of a concept, crowds out other ways of thinking about the situation. However, in Titmuss's puzzle, payment is being offered to others. Presumably the payoff on offer is not terribly salient and it is also possible to avoid it. Assimilating Titmuss's puzzle to standard motivation crowding does not seem plausible.

However, there is a mechanism involving framing that would resolve Titmuss's puzzle, which draws on another mode of valuations in Anderson's taxonomy. The value of *shared goods* is dependent on other people enjoying the item according to shared understandings of what they mean; “shared understanding” is left as an intuitive notion (Anderson, 1993). Michael Walzer (1983) and Charles

Taylor (1985) each appeal to a something like a shared understanding. Walzer argues that we should distribute goods in accordance with their *social meanings*, our shared conceptions of what they are and what they are for. He says that social meanings are guides to why we value things and, once we know why we value something, we can think about what distributive mechanism is appropriate for it. Taylor is more concerned with how we interpret social behavior. He introduces notions of “common meanings” and “inter-subjective meanings” which are constitutive of social reality. Between them, these provide a set of common terms of reference that render social practices meaningful to those who participate in them. I will use the terminology of “shared understandings” to refer to this whole cluster of ideas.

My notion of “frames” relates to at least some of what is meant by “understandings”. (Indeed Schick (1991) casts his theory, which is a framing theory of choice, in terms of “understandings”.) The way that an individual frames her decision can affect the reasons she perceives as relevant and how she values her options. So a shared understanding is a common way of framing a situation.

However, an important feature of shared understandings or shared frames is that they entail more than just everyone having the same individual understanding or frame. There is a thicker sort of shared-ness at work here. At the very least, as well as having the same understanding, individuals also have to believe correctly that others have the same understanding. This is similar to the philosophical idea of “mutual belief”. Taylor (1985, p.197) has an even stronger idea of shared understanding. He says of common meanings that they are not just shared in the sense that everyone has them, nor are they simply known to be shared, but they are also common in the sense of “being in the common reference world.” By this, Taylor means that the sharing is communally sustained.

While a frame is a feature of an individual’s psychology, if a shared understanding exists then it is a fact about the world, albeit one that is based on the psychology of the individuals within it. There is a normative aspect to understandings: someone might think that we ought to understand or frame a good or service in a particular way, related to particular values that she has and thinks others ought to have. However, whilst an individual might think that a good or service ought to be a shared good, presumably related to the way that she understands or frames it as an individual, she cannot decide that there will be a shared understanding and that the good will be valued in that way, independently of how other people understand it.

Selling blood is a market transaction. Markets are individualistic modes of interaction, in the sense that participating in a market exchange only requires knowing one's individual valuation of a good and prices reflect individual preference orderings without reference to the underlying reasons for which the good or service is valued. Therefore, in the marketplace, there is no need for shared values or shared understandings of the items that are for sale. (There may be a shared understanding of the situation, that this is a market transaction. However, a part of that understanding is that individuals can have qualitatively different valuations of the goods or services being exchanged, i.e. they can value the goods for different reasons.) In contrast, giving blood is an example of what Taylor calls a "constitutive practice". Implicit within them "is a certain vision of the agent and his relation to others and society" and they "require that one's actions and relations be seen in the light of this picture and the accompanying norms" (Taylor, 1985, pp. 34-35).

Titmuss makes a suggestion which would add detail to Taylor's picture, that there are competing conceptions of blood, as a gift that one gives and as a commodity that one sells. Gifts are thick normative concepts and gift giving is part of a complex social web. The giving and receiving of gifts brings a host of obligations that vary between cultures (Mauss, 1954). Viviana Zelizer studied domestic transactions, gifts and charitable donations in the US between 1870 and 1930 (Zelizer, 1997). She found that people maintain strong distinctions among entitlements, gifts and payments, considering each to define different kinds of social relations and meaning systems. It is important to distinguish a gift from a market transaction, and important to make gifts appear personal in order to avoid the interpretation that the transaction is merely payment for services. It is important to convey that something is understood as a gift, not a payment.

Shared understandings may be fragile because they rely on mutual beliefs that others understand the good or service in the same way. In large scale situations, such as blood donation, we cannot discover everyone's beliefs by communication. Rather we infer their beliefs from their actions. As Taylor puts it, social actions are communicative acts. The existence of a market marks something out as a commodity, the creation of a market for blood is a sign that the shared understanding that is necessary for giving does not exist, and the participation of people in that market is a further sign. The more people who are not participating, or the more salient the option of non-participation, the more

the existence of the shared understanding is challenged. There is evidence of the stability of monetary frames (Gneezy & Rustichini, 2000). This is no wonder because, in general, mutual belief is easier to break down than create.

If markets cause shared understandings to break down, they will also destroy shared valuations and prevent people from enjoying shared goods. Although one may think that blood *should* be a gift, and even decide to treat it as a gift, one cannot decide that it *is* a gift independent of what other people in the community are doing. So if the shared understanding is threatened by the behavior of some individuals, it may not be possible for others to give it as a gift. Thus the spillover effect is explained, as is Titmuss's assertion that "private market systems ... deprive men of their freedom to give" (Titmuss, 1970, p.239). The crowding out of motivation crowds out frames; shared valuations depend on shared understandings, so the crowding out of frames can cause shared understandings and, hence, shared valuations to break down. Connecting the crowding out of behavior with the crowding out of shared valuations shows how some people's market transactions can affect other people's valuations of goods and services, hence providing a mechanism that could underpin the empirical premise.

Although I rejected the simple over-justification and signaling theories of motivation crowding, this framing mechanism has elements of both. The sense of over-justification is not that of being controlled, but of actions that are over-determined by the agents' reasons. So it is not that agents cease being intrinsically motivated because they dislike losing their autonomy, but that they tend to be one-reason decision-makers so, when their actions are over-determined, some reasons tend to drop from sight. Whilst the payment for blood does not send a signal that the activity is unpleasant, nor the action of donating send a signal that the donor is an altruist, the act of selling blood is a signal that a shared understanding of blood as a gift does not exist, so giving is not appropriate.

The account I have given contains a rather crude account of shared understandings. There are at least two notable simplifications, which turn out to be related: that people are homogeneous, all either having the same understanding or not, and that shared understandings are binary in nature, either existing or not. In fact, people are heterogeneous and there may be vagueness around where a shared understanding passes into or out of existence. These two points are connected. If agents are heterogeneous, so different agents can be using different frames, then there is an issue of how many

agents have to share an understanding in order for a shared understanding to exist. In turn, resolving this vagueness is one source of individual differences, as different individuals may have different thresholds of how many others have to share the understanding before they consider a shared understanding to exist.¹⁷ Some agents may be more tenacious in their shared understandings and less likely to relinquish them in the face of contrarian evidence. For instance, although I said above that a single individual cannot *make* gift be the social meaning of blood independently of the understandings of others in the community, she could potentially decide unilaterally to *treat* blood as a gift, even in the face of evidence that others do not treat it that way.

Shared understandings are relative to a community. A further issue is what are the boundaries of the community. If an understanding is not shared in the wider community, one response for individuals who want to institute it as a shared understanding is to create a sub-community with like-minded people, where their preferred understanding is shared. However, if this route is not taken then, whilst people may believe that it is appropriate for something to be a shared good, they may not be able to sustain it as such.

We are now in a position to answer Mack's (1989) challenge. He asked why, if non-market relations are so much better than commodified ones, do people not continue using the non-commodified relations after markets are introduced? The answer is that, if the non-market relation relies on a shared understanding and the introduction of a market breaks down that shared understanding, then the non-market relation may no longer exist. It is also perfectly consistent to acknowledge that the non-market relation would be preferable, even though it is not available. Further, people sometimes do try to influence shared valuations or form sub-communities where the shared valuation exists.

I have offered a mechanism that can explain both Titmuss's puzzle and the empirical premise. Of course, this is not the only mechanism that might explain Titmuss's puzzle. For instance, the introduction of the market might change people's expectations about what others will do. If people care about the total amount of blood collected and they think that paying for blood will make others more likely to give, then they would withdraw their own contribution. However, this explanation is inconsistent with the data on charitable giving (Sugden, 1982). People want to make a contribution

themselves, sometimes referred to as getting a “warm glow”, as well as wanting to reach a high total contribution level (Andreoni, 1990). In any case, my aim is not to argue that the framing explanation of Titmuss’s puzzle is the only solution, it is to argue that the framing explanation is needed both to explain Titmuss’s puzzle about behavior change and to connect that behavior change to value change, as required for the empirical premise. To the extent that you think the correct explanation of Titmuss’s puzzle involves, e.g., expectations, and does not connect behavior to values, so it cannot explain the empirical premise, then so much the worse for commodification arguments.

6. Limiting the scope of commodification arguments: why some shared valuations are vulnerable to the market but others are not

I have shown that a framing mechanism can explain motivation crowding, solve Titmuss’s puzzle, and support the empirical premise in the case of markets for blood. I also addressed the explicit challenge that Mack (1989) posed. Implicit in his example was a further, more vexing challenge: A plausible commodification argument needs to explain why markets devalue gifts of blood but not gifts of sweaters. I said that a mechanism that gives an answer to the question of *why* commodification occurs will also give guidance as to *what* goods may be commodified because it may place conditions on the type of goods that are vulnerable to commodification. I turn to that issue now.

In Anderson’s typology both blood and sweaters may be shared goods, whose value depends on other people enjoying the item according to shared understandings of what they mean. Markets may drive out shared goods if they cause shared understandings to break down. I explained how this might happen using the example of blood. However, it is not plausible that this mechanism would operate in the case of sweaters.

When Mack changed the example, he made a relevant change to the type of good under discussion. Donating blood involves anonymous giving on a large social scale whereas, in the case of Mack’s sweater, the gift involves only him and his wife. There are many more people involved in the shared understanding underlying blood donation, there is less possibility of communication, and there is less background knowledge of the motivations and values of others involved in the exchange. Between spouses, gifts can be personalized, by giving something that the receiver wants, needs or would

like.¹⁸ A paradigm gift reflects the giver's thoughtfulness about and knowledge of the receiver. A gift between spouses is also given in the context of an on-going relationship where there are other opportunities for showing love and respect, so the parties come to the gift exchange with background knowledge of the other's love and esteem. Hence, unlike knowing the price of blood, knowing the price of a gift from one's spouse does not cause one to doubt the nature of the relationship.

Hence, the framing mechanism does not provide unqualified support for commodification arguments because of the difference between small scale and large scale social exchange. Small scale exchanges allow some goods (like sweaters) to be partially commodified, with people knowing their price but still sustaining a shared valuation (for instance, as a symbol of a loving relationship). In large scale exchange, there is greater potential for lack of mutual belief about a shared understanding of a good or service. For this reason, we might say that the shared understandings underpinning large scale shared goods are likely to be weaker and more easily eroded than those underpinning smaller scale shared goods.

The scope of commodification arguments is limited to goods and relationships that depend on relatively fragile shared understandings. In order to decide whether commodification arguments apply to any particular good, we would have to look in detail at the context of the relationships and the shared valuations involved, in order to assess the robustness of the shared understanding that underpins them, and decide whether the shared understanding was vulnerable to the market.

However, this consideration may lead us to doubt whether commodification arguments are sound in one of the other cases that Anderson and Radin discuss, namely the market for sexual services. The shared understanding between the participants in intercourse only needs to be shared by them, and the shared understandings of dyads and small groups are relatively strong. In George Orwell's *1984*, the Party tries to impose the shared understanding that sex is only for procreation, but, ultimately, it cannot prevent Winston Smith and Julia from having a meaningful sexual relationship based on their pairwise shared understanding of what they are doing. (Although the Party can take steps to destroy the relationship when it discovers it.) Regarding prostitution, most people know that sexual intercourse is something that one could purchase. However, their sexual relationships need not become commodified because the shared understanding of the participants about the meaning of the

interaction is likely to be strong enough to maintain intercourse as a shared good, even if others are buying and selling sex. Therefore it seems likely that Radin (1989) is wrong when she says that “the existence of some commodified sexual interactions will contaminate or infiltrate everyone’s sexuality so that all sexual relationships will become commodified.” (p.182) The existence of prostitution does not prevent couples enjoying the shared and personal goods of intimate sexual relationships.¹⁹

I have tackled the problem of commodification from the perspective of philosophers who study practical reasoning, which is the tradition that Anderson (1993) and Radin’s (1989, 1987) arguments hail from. This is complementary to approaches taken in sociology and, indeed, I have made use of some sociological research. Whereas the practical reasoning approach focusses on the mental states of individuals, the sociological approach focusses on the study of institutions, understood as systems that are governed by rules and norms, and which embed values. In that approach, Titmuss (1970) is read as saying that different institutions of exchange encourage different behaviors and valuations. A market is one such institution. Instead of contrasting market exchanges with non-market exchanges, one can investigate the nuance of exchange and the ways in which people involved in markets (or other institutions) structure and mark their transactions in order to support or obscure social meanings of goods (Rossman, 2014; Almeling, 2011; Healey, 2010; Zelizer, 2010). So where I have talked of frames and making efforts to frame a good in a particular way, market vs non-market, a sociologist might substitute talk of the ways in which a particular market employs social relations in order to sustain social meanings of goods. In the framework I have used, this would be operationalized as supporting or occluding frames. The two approaches differ in that the market as an institution is a hybrid in a way that the market frame is not. One additional insight from the perspective from sociology is that participants in markets may take steps to prevent the transactions being viewed through the market frame. Thus we have yet another potential bulwark against commodification arguments. Contra Radin (1987) commodification is not a domino effect, where pushing over the first domino inevitably causes all the rest to topple, though it may exert pressures in that direction.

7. Conclusion

I have connected political philosophy arguments about commodification to an empirical literature, especially regarding motivation crowding. I presented a mechanism by which markets might break down shared understandings and hence prevent people from enjoying shared goods, in Anderson's (1993) sense. The crowding out of motivation crowds out frames; shared valuations depend on shared understandings, so the crowding out of frames can cause shared understandings and, hence, shared valuations to break down. This mechanism can explain the effect of withdrawing payments, found in Deci's original experiments, and the effect of introducing payments; it connects behaviour to valuations and shows how some people's behavior can spillover and affect other people's valuations,

However, whether or not a particular shared understanding is corroded by the market depends on how fragile it is. It is plausible that the shared understandings involved in large scale social exchange are fragile and hence vulnerable to this mechanism. However, some shared understandings are more robust, especially those existing between a lesser number of people and in the context of long-term relationships. For this reason, it seems doubtful that prostitution breaks down the shared valuations involved in long-term sexual relationships, contra the claims of Anderson and Radin.

The project of connecting arguments from political philosophy to the empirical literature is important for all those who worry about the adverse effect of markets. This is most obviously true for non-essentialist arguments against markets, which point only to their role in sustaining harmful outcomes, for example, Satz's (2010) argument that prostitution is problematic because it is a "theatre of inequality", which contributes to sustaining the subordination of women. However, empirical evidence can also be relevant to essentialist arguments. Commodification arguments are a type of essentialist argument that also relies on a mechanism to propagate essentialist effects. I used moral psychology to explain *why* some people's trading a good or service can debase it for third parties and showed how this can put conditions on *what* goods and services may be commodified, with implications for political theory.

Although I have drawn on an empirical literature in order to devise a mechanism, the mechanism I presented has yet to be empirically tested. To my knowledge, there are no existing experiments examining motivation crowding (of behavior) in third parties, let alone the crowding out

of valuations in third parties. Obviously, it is important for commodification arguments that the hypothesized mechanism actually operates. But testing it is a task for another day and another paper.

References

Akerlof, G.A. and Dickens, W.T. (1982) The economic consequences of cognitive dissonance. *The American economic review*, 72(3): 307-319.

Almeling, R. (2011). Sex cells: The medical market for eggs and sperm. Univ of California Press.

Alvarez, Maria, "Reasons for Action: Justification, Motivation, Explanation", The Stanford Encyclopedia of Philosophy (Summer 2016 Edition), Edward N. Zalta (ed.), URL = <<http://plato.stanford.edu/archives/sum2016/entries/reasons-just-vs-expl/>>.

Anderson, E. (1993) *Value in ethics and economics*. Harvard University Press.

Andreoni, J. (1990) Impure altruism and donations to public goods: A theory of warm-glow giving. *The economic journal*, 100(401): 464-477.

Arrow, K.J. (1972) Gifts and exchanges. *Philosophy & Public Affairs* 1(4): 343-362.

Bacharach, M. (2003) Framing and cognition: the bad news and the good. In: N. Dimitri, M. Basili & I. Gilboa (eds.), *Proceedings of ISER Workshop XIV: Cognitive Processes in Economics*, London: Routledge.

Baldry, J. (1986) Tax Evasion is not a Gamble: A report on two experiments. *Economics Letters*, 22(4): 333-335.

Baranowski, T.A. and Summers, D.A. (1972) Perception of response alternatives in a Prisoner's Dilemma game. *Journal of Personality and Social Psychology*, 21(1): 35.

Becker, G. (1968) Crime and Punishment: An economic approach. *Journal of Political Economy*, 76(2): 169-217.

Benabou, R. and Tirole, J. (2006) Incentives and Prosocial Behavior. *American Economic Review* 96(5): 1652-1678.

Benabou, R. and Tirole, J. (2003) Intrinsic and Extrinsic Motivation. *Review of Economic Studies* 70(3): 489-520.

Bicchieri, C. (2005) *The grammar of society: The nature and dynamics of social norms*. Cambridge: Cambridge University Press.

Bruner, J. (1957) On Perceptual Readiness. *Psychological Review*, 64(2):123-152.

Cerasoli, C. P., Nicklin, J. M., & Ford, M. T. (2014). Intrinsic motivation and extrinsic incentives jointly predict performance: A 40-year meta-analysis. *Psychological bulletin*, 140(4), 980.

Cohn, A., Fehr, E. and Maréchal, M.A., (2014). Business culture and dishonesty in the banking industry. *Nature* 516(7529) .86-89

Colman, A. M. (1995). *Game Theory and its applications in the social and biological sciences* (2nd ed.). Oxford: Butterworth-Heinemann & London: Routledge.

Cookson, R. (2000) Framing effects in public goods experiments. *Experimental Economics* 3(1):55-79.

Deci, E. L (1971) Effects of externally mediated rewards on intrinsic motivation. *Journal of personality and Social Psychology* 18(1): 105-115

Deci, E. L. (1975) *Intrinsic motivation*. New York: Plenum Publishing Co.

Deci, E. L., Koestner, R., & Ryan, R. M. (1999) A meta-analytic review of experiments examining the effects of extrinsic rewards on intrinsic motivation. *Psychological bulletin* 125(6): 627-668.

Deci, E. L., & Ryan, R. M. (2000). The "what" and "why" of goal pursuits: Human needs and the self-determination of behavior. *Psychological inquiry*, 11(4), 227-268.

Dietrich, F., & List, C. (2016). Reason-based choice and context-dependence: An explanatory framework. *Economics and Philosophy*, 32(02), 175-229.

Fischhoff, B., Slovic, P. and Lichtenstein, S., 1978. Fault trees: Sensitivity of estimated failure probabilities to problem representation. *Journal of Experimental Psychology: Human Perception and Performance* 4(2): 330.

Frey, B. S. (1997a) *Not Just for the Money* London: Edward Elgar.

Frey, B. S. (1997b) A Constitution for Knaves Crowds Out Civic Virtues. *The Economics Journal* 107(443): 1043-1053.

Frey, B. S., & Jegen, R. (2001) Motivation crowding theory. *Journal of economic surveys* 15(5): 589-611.

Frisch, D. (1993) Reasons for framing effects. *Organizational Behavior and Human Decision Processes* 54(3): 399-429.

Gneezy, U. and Rustichini, A. (2000) A fine is a price. *J. Legal Stud.* 29: 1-17.

Gold, N. and List, C. (2004) Framing as path dependence. *Economics and Philosophy* 20(02):253-277.

Granovetter, M., (1978). Threshold models of collective behavior. *American journal of sociology* 83(6): 1420-1443.

Healy, K. (2000) Embedded Altruism: Blood Collection Regimes and the European Union's Donor Population. *American Journal of Sociology* 105(6): 1633-1657.

Healy, K. (2010). Last best gifts: Altruism and the market for human blood and organs. University of Chicago Press.

Heider, F. (1958) *The Psychology of Interpersonal Relations*. New York: Wiley.

Heyman, J., & Ariely, D. (2004). Effort for payment a tale of two markets. *Psychological science*, 15(11), 787-793.

Higgins, E. T. (1996) Knowledge Activation: Accessibility, Applicability, and Salience. In: E. T. Higgins & A. W. Kruglanski (eds.), *Social psychology: Handbook of basic principles*. New York: The Guildford Press.

Jones, E.E., Kannouse, D., Kelley, R., Nisbett, R., Valins, S. and Weiner, B (eds.) (1972) *Attribution: Perceiving the causes of behavior*. Morristown, NJ: General Learning Press.

Kahneman, D. and Tversky, A. (1979) Prospect theory: An analysis of decision under risk. *Econometrica* 47(2): 263-291.

Komorita, S.S. (1987) Cooperative choice in decomposed social dilemmas. *Personality and Social Psychology Bulletin* 13(1): 53-63.

Kreps, D.M. (1997) Intrinsic motivation and extrinsic incentives. *The American Economic Review* 87(2): 359-364.

Kühberger, A. (1998) The influence of framing on risky decisions: A meta-analysis. *Organizational behavior and human decision processes* 75(1): 23-55.

Lepper, M. Greene, D. and Nisbett, R. E. (1973) Undermining Children's Intrinsic Interest with Extrinsic Rewards *Journal of Personality and Social Psychology* 28(1973): 129-137.

Liebrand, W.B., Jansen, R.W., Rijken, V.M. and Suhre, C.J., (1986) Might over morality: Social values and the perception of other players in experimental games. *Journal of Experimental Social Psychology* 22(3): 203-215.

Lindenberg, S. and Frey, B.S. (1993) Alternatives, frames, and relative prices: A broader view of rational choice theory. *Acta sociologica* 36(3): 191-205.

Mack, E. (1989) Dominos and the Fear of Commodification. In: J. Pennock & J. Chapman (eds.), *Markets and Justice* New York: New York University Press, pp.198-225.

Mauss, M. (1954) *The Gift: The Form and Reason for Exchange in Archaic Societies*. New York and London: W. W. Norton & Co.

Messick, D.M. and McClintock, C.G. (1968) Motivational bases of choice in experimental games. *Journal of experimental social psychology* 4(1): 1-25.

Mijović-Prelec, D., & Prelec, D. (2010). Self-deception as self-signalling: a model and experimental evidence. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, 365(1538), 227-240.

Montgomery, H. (1983) Decision rules and the search for a dominance structure: Towards a process model of decision making. *Advances in psychology* 14: 343-369.

Pruitt, D. (1967) Reward structure and cooperation: The decomposed prisoner's dilemma game. *Journal of Personality and Social Psychology* 7(1): 21-7.

Pruitt, D.G. (1970) Motivational processes in the decomposed Prisoner's Dilemma game. *Journal of Personality and Social Psychology* 14(3): 227-38.

Rabin, M. (1993) Incorporating fairness into game theory and economics. *The American economic review* 83(5): 1281-1302.

Radin, M.J. (1987) Market-inalienability. *Harvard law review* 100(8): 1849-1937.

Radin, M.J. (1989) Justice and the market domain. *Nomos*, 31, In J. W. Chapman & J. R. Pennock (eds.), *Markets and Justice* New York: New York University Press, pp.165-197.

Ross, M., (1975). Salience of reward and intrinsic motivation. *Journal of Personality and Social Psychology*, 32(2): 245-254.

Rossman, G. (2014). Obfuscatory relational work and disreputable exchange. *Sociological Theory*, 32(1), 43-63.

Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American psychologist*, 55(1), 68.

Sandel, M.J.,(2012). *What money can't buy: the moral limits of markets*. Macmillan.

Satz, D. (1995) Markets in women's sexual labor. *Ethics* 106(1): 63-85.

Satz, D. (2010). *Why some things should not be for sale: The moral limits of markets*. Oxford University Press.

Schick, F. (1991). *Understanding action: An essay on reasons*. Cambridge University Press.

Schick, F. (2003) *Ambiguity and logic*. Cambridge: Cambridge University Press.

Seabright, P. (2002) "Blood, Bribes and the Crowding-Out of Altruism by Financial Incentives,"
unpublished manuscript, University of Toulouse.

Shafir, E., Simonson, I. and Tversky, A. (1993) Reason-based choice. *Cognition* 49(1-2): 11-36.

Skyrms, B. (1998) "Review of Frederick Schick's Making Choices" *The Times Literary Supplement*. 949: 30

Slovic, P., Fischhoff, B., and Lichtenstein, S. (1988) Response Mode, Framing, and Information-Processing Effects in Risk Assessment. In: D. Bell, H. Raiffa & A. Tversky (eds.), *Decision making: Descriptive, normative and prescriptive interactions* Cambridge: Cambridge University Press.

Solow, R (1971) Blood and Thunder. *Yale Law Review* 80 (8): 1696-1711.

Sugden, R., (1982) On the economics of philanthropy. *The Economic Journal* 92(366): 341-350.

Taylor, C. (1985) Interpretation and the Sciences of Man. In: his *Philosophy and the Human Sciences*.
Cambridge: Cambridge University Press.

Titmuss, R. M. (1970). *The gift relationship*. London: George Allen & Unwin Ltd,

Tversky, A. and Kahneman, D. (1981) The framing of decisions and the psychology of choice. *Science* 211(4481): 453-458.

Underhill, K. (2016). When Extrinsic Incentives Displace Intrinsic Motivation: Designing Legal Carrots and Sticks to Confront the Challenge of Motivational Crowding-Out. *Yale J. on Reg.*, 33, 213.

Walzer M. (1983) *Spheres of justice: A defense of pluralism and equality*, New York: Basic Books.

Ward, A. and Ross, L., (1997). Naive realism in everyday life: Implications for social conflict and misunderstanding. In: Reed, E., Turiel, E., and Brown, T. (eds) *Values and knowledge*. Psychology Press, pp. 103-135.

Weirich, P. (2010) Utility and framing. *Synthese* 176(1): 83-103.

Zelizer, V.A.R. (1997) *The social meaning of money*. Princeton: Princeton University Press.

Zelizer, V. A. (2010). *Economic lives: How culture shapes the economy*. Princeton University Press.

¹ The set of goods with use value is a strict subset of extrinsic goods, as there are extrinsic goods that do not have use value. For instance, a bracelet that is only valued as a token of a friendship is an extrinsic good that is not valued as a means to an end.

² Anderson and Radin both talk of female prostitutes and male clients; neither considers male prostitutes, female clients, or same sex prostitution hence nor do I.

³ It should be noted that Deci and Ryan's (2000) self-determination theory still includes feelings of competence, as well as autonomy; it posits that individuals have psychological needs for autonomy, competence and relatedness (the development and maintenance of close personal relationships with others).

⁴ There is no consensus on how exactly these terms are used. In the same way that motivation crowding has been used to refer both to an empirical effect and to a hypothesis about the cause of that effect, the term "over-justification" was coined by Lepper, Greene and Nisbett (1973) who defined it similarly to the way we now understand motivation crowding, as "the proposition that a person's intrinsic interest in an activity may be undermined by inducing him to engage in that activity as an explicit means to some extrinsic goal", but it has also been used as an explanation rather than just a re-statement of motivation crowding theory (Frey, 1997a; Underhill, 2016).

⁵ In fact, the asymmetry is the other way round: donors have information about the quality of their blood that the authorities lack. This does create a problem, namely that authorities must trust donors not to give blood if they have infectious diseases. Titmuss also found that the amount of infected blood donated in the US was higher than in the UK. Maybe not so surprising when one considers that a norm of market transactions is *caveat emptor*, or buyer beware!

⁶ This is in stark contrast with London Underground's campaign in the 1990s, when the British government increased the cost of crime, introducing the London Regional Transport (Penalty Fares) Act, which passed in 1992, and ran a series of radio commercials with the tag line "Get a ticket, not a criminal record".

⁷ This is of particular relevance here because blood donation is a public good and could be modeled as an n -person prisoners' dilemma.

⁸ However, we cannot assume that a player would see any given decomposition from the parent game because, if a prisoner's dilemma is decomposable, then there are an infinite number of possible decompositions (Messick & McClintock, 1968).

⁹ It is possible to translate between the "value-based" model given by Tversky and Kahneman and the "reason-based" tradition. Roughly, what Kahneman and Tversky describe as a change in curvature of the utility function becomes a difference in how people value the options. See also Gold & List (2004).

¹⁰ This can also explain the higher cooperation rates when the prisoner's dilemma is called the "community game" rather than the "Wall Street game" (Ward & Ross, 1997).

¹¹ I was once offered, in the spirit of a competing explanation, the idea that the change in framing in the decomposed prisoner's dilemma changes a player's expectations about what the other player will do. For example, in the Rabin (1993) model of reciprocal fairness, agents want to be kind to agents who they expect will be kind to them. So if the decomposed dilemma leads a player to expect that her co-player will cooperate and she cares about Rabin-kindness, then the change in expectations could lead her to cooperate. Alternatively, the framing may trigger a social norm, whereby it alters the expectation that others will cooperate, and people have a conditional preference that they cooperate given that others will too (Bicchieri, 2005). But why would decomposing the dilemma increase a player's expectation that her co-player will be kind, except because the player is more likely to see the opportunity for kindness? In other words, there can only be a change in expectations because people realise that framing affects behavior. The change in expectations is parasitic on the possibility of changes in frame. Of course, sometimes changing expectations can change behavior without a change in frame but, equally, there is recent evidence that framing can affect behavior without changing expectations (Cohn, Fehr, & Maréchal, 2014).

¹² For an overview of the philosophical complexity of reasons, including the intriguing possibility that some philosophers would prefer to classify the "motivating reasons" that I discuss as "explanatory reasons", see Alvarez (2016).

¹³ Hence I side-step debates about what constitutes a framing effect. For more on different uses of the term see Kühberger (1998) and Frisch (1993).

¹⁴ One way of putting this is that the sort of selective seeing of a situation that is involved in framing is irrational, that rationality requires us to see all possible ways of framing it (Skyrms, 1998). That requirement is imposed by orthodox decision theory. However, in the prisoner's dilemma, if there is one way of decomposing a matrix then there are an infinite number of possible decompositions. This casts some doubt on whether it really is irrational not to see all the decompositions, but there is no space to pursue that argument here.

¹⁵ Bruner (1957) does not say how or why gating out occurs but, in cognitive psychology, there is a well known effect called *assimilation*, where an agent perceives an object's attributes as more typical of the category that is being used than it actually is. Of course, some incongruous cues may be so insistent that the agent cannot ignore them.

¹⁶ So it may be the case that treating workers as though they are extrinsically motivated will actually cause them to be so motivated, creating the need for incentives and rewards where none existed before. Further discussion of the idea that designing institutions as if people are knaves causes them to behave as such can be found in Frey (1997b).

¹⁷ For threshold models of behavior, as studied in sociology, see Granovetter (1978).

¹⁸ It is interesting that money is an acceptable gift only for inter-generational transfers between family members. Even in casual gifts, between acquaintances, there is some onus to get something appropriate and to have invested some little effort in its purchase (Zelizer, 1997).

¹⁹ This is not to deny that widespread prostitution might have a society-wide effect on how men see women, as argued by Satz (1995). But that is a different argument.