

# Awards, incentives and mutual benefit

Robert Sugden<sup>1</sup> 

Received: 30 August 2017 / Accepted: 22 December 2017 / Published online: 6 January 2018  
© The Author(s) 2018. This article is an open access publication

**Abstract** Frey argues that if buyers and sellers of labour understand their relationship merely as exchange, workers' intrinsic motivation can be impaired; this problem can be partially overcome by using awards as a supplementary reward mechanism. I argue that this proposal is self-defeating. In an economy that relies on the division of labour, it is an unavoidable fact that individuals are subject to the will of others; award-giving practices are merely camouflage. However, recognising this fact need not impair anyone's sense of autonomy as a paid worker if participation in market exchanges is understood as expressing intentions for mutual benefit.

**Keywords** Awards · Incentives · Mutual benefit · Intrinsic motivation

**JEL Classification** D9 · B41

This paper reflects on the role of awards in supporting workers' intrinsic motivation to deliver high-quality work. As I will use the word, an *award* is something that may or may not have material value, given gratuitously in recognition of the recipient's merit or achievement, and signalling the giver's esteem for him or her. Academy Awards, Nobel Prizes and Booker Prizes are awards. So are decorations such as the George Cross (awarded in Britain for 'acts of the greatest heroism or of the most conspicuous courage in circumstances of extreme danger') and the Gold Medal of the Royal National Lifeboat Institution (awarded for 'conspicuous gallantry' in saving lives at sea). So is a 'best instructor of the year' award given by a student organisation to a selected teacher. Gratuitousness distinguishes awards from prizes that are awarded to the winners of contests in which individuals compete openly

---

✉ Robert Sugden  
r.sugden@uea.ac.uk

<sup>1</sup> School of Economics, University of East Anglia, Norwich NR4 7TJ, UK

according to well-defined rules. The winner of the women's 100 m sprint in the Olympics is entitled to her gold medal in much the same way that an employee who has delivered her part of a labour contract is entitled to be paid her wage. In contrast, a Nobel Laureate may be *pleased about* or *grateful for* his award, but he is not entitled to it.

I discuss a hypothesis about awards that has been proposed by Frey (1997). Building on work by psychologists who have investigated motivation, Frey argues that if buyers and sellers of labour understand their relationship merely as a market exchange, this perception can impair workers' motivation and sense of autonomy, reducing both their productivity and their happiness. One way of reducing the extent of this problem, according to Frey, is to use awards as a supplementary reward mechanism. I will argue that it is self-defeating to try to use the gratuitousness of awards as a motivating mechanism. In any economy that depends on the division of labour and in which there is freedom of contract, the relationship between employer and employee is necessarily one of exchange, and it is unavoidable that some (typically all) people who are parties to this relationship are subject to the will of others. The giving of awards may camouflage this fact of economic life but cannot erase it. However, I will try to show that recognising this fact need not impair anyone's sense of autonomy.

## 1 Awards

The psychological theory of intrinsic motivation and motivational crowding-out, developed by, for example, Bem (1967), Deci (1971), Lepper and Greene (1978) and Deci and Ryan (1985), has been around for a long time. The essential idea is central to Titmuss's (1970) famous analysis of blood donation—an analysis that, at the time it was written, was well known to economists of social policy. However, it is only relatively recently that the theory has been taken seriously by economists. Arrow (1972), probably the greatest economic theorist of his era, acknowledged Titmuss's argument but expressed puzzlement about it. Gneezy and Rustichini's (2000a, b) report that fines for lateness in collecting children from a day care centre led to an increase in the incidence of lateness is now often cited as evidence of motivational crowding-out, but their favoured explanation of this observation—that it was a consequence of incomplete contracts—was much more conventional.

Frey (1994, 1997) was a pioneer in applying the theory of intrinsic motivation to economic problems. His analysis draws on Deci and Ryan's distinction between *extrinsic* and *intrinsic* motivation. According to the taxonomy proposed by Ryan and Deci (2000), an activity is extrinsically motivated if it is done 'in order to attain some separable outcome'; it is intrinsically motivated if it is done 'for its inherent satisfactions ... for the fun or challenge entailed rather than because of external prods, pressures, or rewards'. Extrinsic motivations can be classified on an *internal/external* scale. They are more internal, and so more similar to intrinsic motivations, the greater the degree to which they go together with perceptions of *competence*, *autonomy* and *relatedness*. ('Competence' is self-explanatory. 'Autonomy' is the sense of an 'internal locus of causality'—of being the cause of one's own actions.

‘Relatedness’ is the sense that one’s actions are valued by ‘significant others’ with whom one feels connected.) The hypothesis of motivational crowding-out is that if an activity has both extrinsic and intrinsic rewards, people tend to attribute extrinsic motivations both to other people’s and (crucially) to their own engagement in that activity, undermining intrinsic motivation. Thus, in Titmuss’s example, unpaid blood donation is attributed to motivations of altruism, reciprocity or public spirit. But if payments are introduced, giving blood (even by people who would have donated without being paid) tends to be attributed to the extrinsic reward provided by the payment.

Given this conceptual scheme, it is difficult for any form of paid labour to be intrinsically motivated, although in the right circumstances a worker might have feelings of competence, autonomy and relatedness. The problem is fundamental to the whole idea of a market economy based on the division of labour. A worker contributes specialised labour to the production of some particular good in return for a wage that gives her a menu of opportunities to consume goods that other workers have produced. If this is how she understands what she is doing, her work is an activity that is done to obtain a *separable outcome* or *external reward*, namely the goods that other workers produce. She can be intrinsically motivated only if she would have wanted to do the work even if she had not been paid for it. But, to make sense of that idea, we have to imagine a situation in which there was no *prod* or *pressure* for her to spend her working day on any other, unwanted kind of labour. So it seems that we must imagine that, whether she worked or not, she would have been paid the same wage. Her intrinsic motivation would then be revealed by her choosing to work solely for its fun or challenge.

Contributors to the literature of intrinsic motivation have tried to find ways of combining the economic reality of employment as an exchange relationship with the perception that there is gratuity on both sides—that, at least to some degree, the worker supplies her labour without intending to receive payment for it, and the employer pays the worker’s wage without intending to receive her labour in return. A variant of this idea can be found in Akerlof’s (1982) model of labour contracts as ‘partial gift exchange’. (In Akerlof’s model, the effort expended by workers is hard to monitor. An employer pays a worker more than her reservation wage; the worker responds by supplying effort in situations in which effort cannot be monitored.)

Developing an idea from the psychological literature, Frey distinguishes between two types of intervention in a person’s working environment:

Crowding-out effects take place... when an external intervention is perceived to be *controlling*. In contrast, when the intervention by the principal is understood to be *supportive*, intrinsic work motivation is unaffected or may possibly even rise...

When employees get rewarded only if they have performed exactly according to their chief’s directions, their intrinsic work motivation is negatively affected. The more a reward is contingent on the performance desired by the principal, the more strongly the locus of control is shifted from intrinsic to extrinsic incentives, and the more work morale is crowded out. (1997, pp. 93–94)

If rewards are separated from specific components of the general activity that is being rewarded, a perception of gratuitousness can be maintained. One way in which this separation can be achieved is through suitably chosen gifts:

Non-monetary gifts in kind (for example, in the form of flowers, books or chocolates) constitute a conscious effort to disassociate the reward from any particular performance. Rather, these gifts are chosen so that the person's self-esteem is acknowledged, thereby bolstering the recipient's intrinsic work motivation. (1997, p. 95)

The suggestion is that, by supporting intrinsic motivation, gifts can serve a useful economic function. They can be perceived as gratuitous acknowledgements of merit, independent of the economic value actually created by the recipient, while in fact inducing workers to create economic value.

In the passages I have just quoted, Frey is discussing hierarchical workplace relationships between supervisors (or 'chiefs') and their supervisees. A 'controlling' relationship is being pictured as one in which a supervisor gives *directions* to a supervisee, and rewards the supervisee for acting according to those directions. In other words, the supervisor uses the reward mechanism as a means of intentionally controlling the actions of the supervisee. In contrast, the use of a gift as a reward is supposed to signal a relaxation of that control.

Frey extends this analysis to the case of 'honours and prizes', where hierarchical relationships may be less explicit:

A lot of trouble is taken *not* to grant these rewards as a response to a particular performance, but rather as a recognition of a person's dedication to his or her work or career in general. Titles, orders and other honours (such as honorary doctorates) are normally given for one's lifelong work, and are thus perceived by the recipients as a recognition of their competence, raising intrinsic motivation. (1997, p. 95)

Take an example of an academic honour. Fellowships of the British Academy are honours to which neither material rewards nor particular duties are attached. Their recipients are scholars whom the existing Fellows judge to have achieved 'distinction in the humanities and social sciences'. The Academy is not the employer of potential recipients; it is merely an institution that decides whether to honour them. Still, one might say that there is a hierarchical relationship between the two. To the extent that Fellowships are valued by their recipients, the Academy controls a reward mechanism that might be used to direct scholars towards particular activities. The Academy's (rather secretive) procedures are careful not to specify explicit criteria for Fellowship awards that potential recipients might try to meet. Thus, as Frey claims is typical in the award of honours, reward is disassociated from particular performance, in much the same way as when a supervisor gives a box of chocolates to a meritorious supervisee.

But Frey also applies his analysis of intrinsic motivation to a very different case—the payment of labour in a competitive market:

A monetary payment received through the functioning of the market constitutes a case where the reward depends on performance; in a perfectly competitive market the reward (wage rate) depends exactly on the marginal product performed. The price system therefore tends to substitute intrinsic with extrinsic motivation. On the other hand, a market reward may also indicate competence and then tends to raise work morale. (p. 94)

Here the argument is not about hierarchical relationships of control. (In a perfectly competitive market, no individual agent has any control over prices.) The claim is that intrinsic motivation is undermined if the worker perceives his wage as the exact value of his services to a buyer, but not if he perceives it as an acknowledgement of his competence, independent of the actual value of what he does. Perhaps what Frey has in mind in the second case are the kinds of career-long salary scales that were once characteristic of the professions—scales that are restricted to people who enter a profession with the required qualifications, and then reward those people according to their seniority rather than their actual performance. The suggestion seems to be that, in order to preserve workers' sense of autonomy, the relationship between buyers and sellers of labour needs to be construed as two-sided gift-giving rather than as the exchange of things that have equal market value.

It seems that, for Frey, there is an equivalence between (on the one hand) payment-by-results incentive systems in hierarchical supervisor–supervisee relationships within firms, and (on the other) relationships between buyers and sellers of labour in competitive markets. By transferring the results of his analysis from the first case to the second, he is presenting *markets* as hostile to intrinsic motivation. I want to argue that the two cases are *not* equivalent.

## 2 Incentives

One way in which one might claim to find an equivalence between these cases is by saying that, in a competitive labour market, the price system is a payment-by-results *incentive mechanism*. This may be what Frey means when he says that the wage received by each worker depends exactly on the marginal product performed. If wages have this property, workers *are* paid according to the results of their labour, and in the language of modern economics, wages *are* incentives. Nevertheless, I think there would be some slippage of meaning in using the word 'incentive' both for payment schemes designed to induce specific performances from workers in a hierarchical firm and for wage rates in a competitive market.

As Grant (2002) has pointed out, the meaning of 'incentive' has evolved significantly since economics began to emerge as a discipline in the eighteenth century. In its original meaning, an 'incentive' could be any stimulus that incited or aroused a person to feeling or action. For example, consider the following passage, which is part of Hume's (1739–40/1978: 394–396) discussion 'Of the amorous passions, or of love betwixt the sexes':

The appetite of generation, when confin'd to a certain degree, is evidently of the pleasant kind, and has a strong connexion with all the agreeable emotions.

Joy, mirth, vanity, and kindness are all incentives to this desire; as well as music, dancing, wine, and good cheer.

Music, wine and dancing are stimuli that tend to induce amorousness; as such, they are *incentives* to amorousness.

Interestingly (if my word-search software is to be believed), the word ‘incentive’ does not appear anywhere in Smith’s *Theory of Moral Sentiments* (1759/1976) or his *Wealth of Nations* (1776/1976). However, it does appear occasionally in nineteenth-century economics, with something like what I take to be its central meaning in modern economics. This is a specialisation of the original meaning. To say that, for person  $i$ ,  $x$  is an incentive to do  $y$  is to say that, were  $i$  to do  $y$ , he would get  $x$  as a *separable private benefit*, and that the prospect of this benefit tends to induce  $i$  to do  $y$ . For example, when John Stuart Mill discusses inheritance in his *Principles of Political Economy* (1871/1909, p. 226), he says:

No doubt, persons have occasionally exerted themselves more strenuously to acquire a fortune from the hope of founding a family in perpetuity; but the mischiefs to society of such perpetuities outweigh the value of this incentive to exertion, and the incentives in the case of those who have the opportunity of making large fortunes are strong enough without it.

Notice that there is no implication that anyone *intends* that the prospect of founding a family fortune induces exertion; this is just a property that the economic system happens to have.

Grant (2002) is particularly interested in an even more specialised meaning of ‘incentive’, which she traces to American writings about the scientific management of production in the ‘Taylorist’ era of the early twentieth century. Grant offers the following definition of this notion of ‘incentive’:

An offer is made which is an extrinsic benefit or a bonus, neither the automatic consequence of an action nor a deserved reward or compensation. The offer is usually made in the context of an authority relationship... The offer is a discrete prompt expected to elicit a particular response. (p. 135)

As in the nineteenth-century usage, an incentive is a separable private benefit which tends to induce a specific response. But in addition: it occurs in an authority relationship; it is set up intentionally by the person in authority; and the response that is incentivised is intended by that person. Grant contrasts incentives (so understood) with ‘market forces’, which are viewed as impersonal and unintended (p. 134). She makes a parallel contrast between wages as incentives and wages as ‘compensation’, maintaining that it is meaningful to talk about fair compensation but not about fair incentives.

Grant is surely right in identifying a distinct, relatively recent and now very common meaning of ‘incentive’, and her account of its Taylorist origins is convincing. I think she overstates the degree to which modern economics is committed to this meaning: my sense is that the nineteenth-century and Taylorist meanings are now used in parallel. But for my purposes, what matters is Grant’s contrast between the two conceptions of incentives, when considered in the context

of intrinsic motivation. It is reasonable to describe Taylorist incentives as ‘external interventions’ that are perceived as ‘controlling’. But does this description apply to the market forces that determine wage rates in a competitive market, as Frey seems to claim it does? I am not persuaded.

### 3 Autonomy

The weak link in this chain of argument is between the concept of a ‘separable outcome’ and that of a ‘controlling intervention’. Separability of outcomes is part of the definition of extrinsic motivation, and extrinsic motivation is classified as not fully autonomous. Being subject to someone else’s control clearly implies some loss of autonomy. But it is not obvious that separability *in itself* is incompatible with autonomy.

Think of someone doing an ordinary useful job, perhaps as a check-out assistant at a supermarket or as the driver of a delivery van. She can see that what she is doing is useful to other people, but she wouldn’t do the work if she wasn’t paid. The wage she receives is equal to the value that those other people place on the services she provides. Is this wage equivalent to a *prod*? Does it compromise her autonomy? If the answer to the second question is ‘Yes’, how are we to imagine an economy of agents who *are* autonomous?

Marx and Engels (1845) famously imagine such an economy as a communist Utopia:

...in communist society, where nobody has one exclusive sphere of activity but each can become accomplished in any branch he wishes, society regulates the general production and thus makes it possible for me to do one thing today and another tomorrow, to hunt in the morning, fish in the afternoon, rear cattle in the evening, criticise after dinner, just as I have a mind, without ever becoming hunter, fisherman, herdsman or critic.<sup>1</sup>

The essential idea is that everyone works according to his or her intrinsic motivation, but somehow, everyone’s needs are satisfied. No one intends to be useful to other people; usefulness just happens. (We are told that ‘general production’ is ‘regulated’, but the regulations somehow allow everyone to spend his time however he likes.) Perhaps the best way to make sense of this picture is to suppose that, under communism, technology has developed so far ahead of human wants that the marginal product of labour has zero value. The only remaining value of work is as a form of amusement. (Whether this vision is Utopian or dystopian is an interesting moral question, but that is not my concern here.)

If Marx and Engel’s ideal of an economy of intrinsically motivated workers is coherent at all, that is because they assume away the problem of scarcity. But in a world of scarcity, pursuing that ideal leads to an impossibility. Intrinsic motivation is as an ideal of autonomy, construed as independence of the will of others. The

---

<sup>1</sup> This passage appears in Part IA of *The German Ideology*, in the subsection ‘Private property and communism’.

intrinsically motivated worker is independent of the will of others, because as a supplier of labour she does as she likes—‘just as she has a mind’. But it is not possible for everyone on both sides of a market (or whatever is the analogue of a market under communism) to do as they like. ‘Doing as you like’ in a market amounts to setting your own terms of trade—for example, choosing the type of work you do *and* choosing that the wage for that work gives you an acceptable standard of living, or choosing which products to consume *and* choosing that those products are acceptably cheap to buy. Since trade is a relationship between people, it is not possible for everyone to set their own terms of trade. If exchange is to be possible, we cannot all be independent of the will of others.

This fundamental problem cannot be avoided by the use of awards. To say that awards are gratuitous is to say that they are subject to the will of the awarder. Just as the earnings of someone who produces consumer goods depend on the tastes and judgements of potential customers, the receipt of an award depends on the tastes and judgements of the awarder. Thus, *wanting* to receive an award makes you subject to the will of others. The practice of tipping illustrates the point. Because restaurant waiters can expect to receive tips, the market wage for waiters is less than it would otherwise be: the expected value of tips is a non-contractual component of the payment that waiters receive for their labour. The practice of tipping reduces the autonomy of waiters by making their rewards dependent on the will of the particular customers they serve. Correspondingly, it increases the autonomy of the customers, by giving them some range of choice over their terms of trade. (Within limits, they can choose not only what to buy, but also how much to pay for it.)

Remember that what is at issue is whether awards can be used to increase workers’ intrinsic motivation. If a practice of award-giving is to have this effect, it is essential that awards are evaluated positively by workers (whether for their material value or as tokens of esteem). It is fundamental to Frey’s concept of award-giving that awards are understood by workers as acknowledging ‘meritorious’ performance of a kind that is at least partly under each worker’s control and that (in fact, but not necessarily as understood by workers) is indicative of intrinsic motivation.

One possibility is that awards are unanticipated—they acknowledge previous meritorious performance by workers who, at the time of that performance, were unaware of the possibility that it might lead to an award. (For example, a professional institution might institute an annual award for lifetime achievement at the same time as making the first award. The first award, but not the later ones, is unanticipated.) The existence of an unanticipated award cannot affect the motivation that lay behind the performance that is being acknowledged and might enhance the recipient’s subsequent intrinsic motivation. In this special case, arguably, the autonomy of potential recipients is not compromised. But if we are concerned with *continuing practices* of award-giving, this case is of little relevance.

If an award-giving practice is known about in advance of opportunities for the meritorious performances it acknowledges, and if workers value the receipt of awards, that practice makes workers subject to the will of the award-giver. Take the case of a writer working on a novel, having signed a contract with a publisher that will give her a percentage royalty on the value of sales. If she thinks of her activity in terms of market exchange, the income she will receive in return for her work



depends on the tastes of the reading public. If she prefers to earn more rather than less, she is subject to the will of those potential readers. She might try to use this to her advantage, by tailoring her book to what she believes to be the tastes of the reading public. Alternatively, she might write the book as she would most like it to be, recognising that she may be sacrificing royalties by doing so. And there is a third possibility: she might think that it is her job as an author to write books that readers want to buy. In other words, the way she most wants the book to be *just is* the way that would be most successful in inducing readers to buy it.

Now suppose that this writer is also aware that her novel has a chance of winning a prestigious literary prize. The award of the prize would give her acclaim as an artist. (It might also increase her sales, but let us leave that aside). If she enjoys acclaim, she is subject to the will of the members of the award-giving committee. She might try to use this to her advantage, by tailoring her book to what she believes to be the tastes of the literary critics who serve on award-giving committees. Or she might write the book as she would most like it to be, recognising that she may be sacrificing acclaim by doing so. Or she might want to achieve greatness as a novelist and think that the criterion of greatness is the receipt of critical acclaim: the way she most wants the book to be *just is* the way that would win most acclaim.

The point of this example is that royalties and acclaim feature in the writer's deliberations in essentially the same way. As Brennan and Pettit (2000) point out, there is an economy of esteem as well as an economy of goods and services. In both economies, each of us is subject to the will of others. But being subject to the will of others is not the same thing as being subject to another person's *control*, as a worker on a Taylorist production line or in an electronically monitored call centre might reasonably be said to be.

## 4 Mutual benefit

I suggest that we need to find a way of understanding autonomy that is compatible with the fact that in economic life, each of us is dependent on the will of others. The solution I propose is explained more fully by Bruni and Sugden (2008, 2013). But in this paper, I will use Immanuel Kant as its advocate.<sup>2</sup>

In 1784, Kant was organising his thoughts for *Groundwork for the Metaphysics of Morals* (1785/2002). A salient feature of this book is its 'practical imperative'. This imperative, which Kant presents as a dictate of reason, is: 'Act so that you use humanity, as much in your own person as in the person of every other, always at the same time as end and never merely as means'. In the same year, Kant was having his first house built in Königsberg. The house was rather grand (it included a lecture hall). Kant's ability to afford this was probably the result of a long-standing and profitable investment partnership with Joseph Green, an English merchant based in Königsberg. Kant and Green were also friends who met every afternoon to discuss philosophy. So Kant was a direct beneficiary of the market economy of his time and

<sup>2</sup> My information about the episode of Kant and the builders is derived from Kuehn (2001) and Willaschek (2017). I was directed to it by Christian Schubert, who recognised its connection with Bruni's and my account of market virtue.

had a continuing opportunity to discuss the practice of business with a successful and philosophically sophisticated merchant.<sup>3</sup> It seems to have occurred to Kant that he might be treating the builders of his house merely as means to his end, rather than as ends in themselves.

One day, he wrote to his architect, in relation to some problem that needed to be sorted out (probably an over-delivery of bricks): ‘Of course my wish is to save money, but in a way that the workers are also satisfied’. Next day, he gave his first lecture on Natural Right (Kant 1784/2003). Here is how he begins (I have added the italics):

All of nature, as far as it is within the reach of his power, is subjected to the will of man, with the exception of other men and reasonable beings. From the point of view of reason, the things in nature can only be regarded as means to ends, but man alone can himself be regarded as an end. It is impossible to conceive of a value of other things other than by regarding them as means to further ends. .... However, man is never merely a means; rather he is at the same time an end. For example: *If a mason serves me as a means to building a house, I serve him, in turn, as a means to acquire money.*

A little later, he repeats the idea:

... The world, as a system of ends, finally has to contain a purpose, and this is the reasonable being. If there existed no end, the means would serve no purpose and would have no value. – Man is an end. It is therefore contradictory that he should be a mere means. – *If I am making a contract with a servant, he has to be an end as well, just as I am, and not merely a means. He must also will.*

For my present purposes, I am not concerned with Kant’s rationalistic metaphysics. What is important is his understanding of the relationship between a buyer and a seller of labour, as expressed both in the lecture and in the letter to the architect. As I read him, Kant is saying that the parties to a market relationship should not treat one another only as means to their own ends. Rather, each party should intend mutual benefit. In his relationship with the builders, Kant should *at the same time* treat the builders as means to his end (getting a house) *and* treat himself as a means to their ends (getting money). Although Kant does not say this explicitly, it is natural to infer that he thinks that the builders’ intentions should reciprocate his: they should at the same time treat Kant as a means to their end of getting money and treat themselves as means to Kant’s end of getting a house.

Presumably, the builders are not acting only on intrinsic motivation in the sense defined by Deci and Ryan. Although they might feel pride in their competence in exercising their craft, they are not building the house merely for fun or challenge. Nor, although they might feel goodwill towards Kant as a prominent resident of their city, are they doing the work because Kant is a significant other in their lives

<sup>3</sup> Unsurprisingly, Kant seems to have had some interest in the ethics of business. In the *Groundwork for the Metaphysics of Morals*, Kant (1785/2002, p. 13) uses ‘honesty’ in business (represented by non-discriminatory pricing) as his leading example of the difference between an action that is performed for the sake of the moral law and one that merely conforms to that law.

who would like to have a house built. They are building the house because they are being paid to do so. But they can recognise this without thinking of the payment as a *prod* by which Kant *controls* them. Reciprocally, Kant is paying their wages because they are building his house, but he can recognise this without thinking of their work as a *prod* by which he is controlled into paying. The builders and Kant can think of themselves as *voluntary co-participants in a scheme of mutually beneficial cooperation*.

Can they think of their relationship in this way even if (as in the case discussed by Frey) the labour market is perfectly competitive? Suppose that if any of Kant's builders had not contracted to work for him, that worker could have earned a similar income by working for someone else. And suppose that if Kant had not employed these particular builders, he could have employed other, similarly efficient workers at a similar rate of pay. In this case, the *marginal* benefit that each party derives from his participation in a specific two-person labour market relationship is very small. Suppose also that there is no gratuitousness of the kind that features in Akerlof's model of partial gift exchange: Kant pays no more than he needs to pay to get the work done, and the builders do no more work than they are paid to do. But it still seems right to say that the relationship between Kant and his builders is one of mutually beneficial cooperation.

To see why, one needs to understand what it means for someone to think of himself as playing his part in a larger scheme in which others play their parts too. In thinking in this way, the individual does not identify his action in terms of its marginal consequences; he identifies it as a part of the scheme as a whole and then considers the consequences of that scheme. [This is the essential idea behind the concept of team reasoning, as developed by Sugden (1993) and Bacharach (2006).] Consider one of the builders, say Franz, laying bricks at the building site. If he is asked what he is doing, he can say that he is playing his part in the building of Kant's house. If he is asked what use this work is, he can say it is useful to Kant because Kant will be able to live in the house after it is built. He can also say that it is useful to him, because working as a builder is how he earns his living. He is playing his part in a building project which, taken as a whole, will benefit Kant. And the time he spends on Kant's house is part of his working week in an occupation which, taken as a whole, benefits him. All this is true, even if it is also true that if Franz had not taken on the work, Willy would have been employed instead, and the house would still have been built. Equally, it is true even if, had Kant not employed him, Franz would have worked the same hours for someone else, at the same rate of pay.

In his lecture on Natural Right, Kant is claiming that doing work for payment does not necessarily involve a loss of autonomy; and I agree with him. There is no need to camouflage the fact that a market economy is a network of mutually beneficial exchanges. One can be autonomous in a market economy by intending mutual benefit in one's trading relationships.

**Acknowledgements** This is an expanded version of a keynote lecture given at an HEIRS (Happiness, Economics and Interpersonal Relations) workshop on 'Prizes and Virtues' at LUMSA University, Rome, in April 2017. I thank participants at that conference and Geoffrey Brennan for constructive comments, and Christian Schubert for suggesting the relevance of Kant's relationship with his builders.

**Funding** This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme, Grant Agreement No. 670103.

### Compliance with ethical standards

**Conflict of interest** The author declares that he has no conflict of interest.

**Open Access** This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

## References

- Akerlof G (1982) Labor contracts as partial gift exchange. *Q J Econ* 97:543–569
- Arrow K (1972) Gifts and exchanges. *Philos Public Aff* 1:343–362
- Bacharach M (2006) *Beyond individual choice*. Princeton University Press, Princeton
- Bem D (1967) Self-perception: an alternative interpretation of cognitive dissonance phenomena. *Psychol Rev* 74:183–200
- Brennan G, Pettit P (2000) The hidden economy of esteem. *Econ Philos* 16:77–98
- Bruni L, Sugden R (2008) Fraternity: why the market need not be a morally free zone. *Econ Philos* 24:35–64
- Bruni L, Sugden R (2013) Reclaiming virtue ethics for economics. *J Econ Perspect* 27:141–164
- Deci E (1971) Effects of externally mediated rewards on intrinsic motivation. *J Pers Soc Psychol* 18:105–115
- Deci E, Ryan R (1985) *Intrinsic motivation and self-determination in human behavior*. Plenum, New York
- Frey B (1994) How intrinsic motivation is crowded in and out. *Ration Soc* 6:334–352
- Frey B (1997) *Not just for the money: an economic theory of personal motivation*. Edward Elgar, Cheltenham
- Gneezy U, Rustichini A (2000a) A fine is a price. *J Leg Stud* 29:1–17
- Gneezy U, Rustichini A (2000b) Pay enough or don't pay at all. *Q J Econ* 115:791–810
- Grant R (2002) The ethics of incentives: historical origins and contemporary understandings. *Econ Philos* 18:111–139
- Hume D (1739–40/1978) *A treatise of human nature*. Oxford University Press, Oxford
- Kant I (1784/2003) *Natural right*. Lecture notes taken by Gottfried Feyerabend, translated by Lars Vinx. [http://www.academia.edu/5541693/English\\_Translation\\_of\\_the\\_Naturrecht\\_Feyerabend](http://www.academia.edu/5541693/English_Translation_of_the_Naturrecht_Feyerabend). Accessed 16 April 2017
- Kant I (1785/2002) *Groundwork for the metaphysics of morals*. Translated by Allen Wood. Yale University Press, New Haven
- Kuehn M (2001) *Kant: a biography*. Cambridge University Press, Cambridge
- Lepper M, Greene D (eds) (1978) *The hidden costs of reward: new perspectives on the psychology of human motivation*. Erlbaum, Hillsdale
- Marx K, Engels F (1845) *The German ideology*. <https://www.marxists.org/archive/marx/works/1845/german-ideology/ch01a.htm>. Accessed 24 Aug 2017
- Mill JS (1871/1909) *Principles of political economy*. Longmans, London. First edition 1848
- Ryan R, Deci E (2000) Intrinsic and extrinsic motivations: classic definitions and new directions. *Contemp Educ Psychol* 25:54–67
- Smith A (1759/1976) *The theory of moral sentiments*. Oxford University Press, Oxford
- Smith A (1776/1976) *The wealth of nations*. Oxford University Press, Oxford
- Sugden R (1993) Thinking as a team: toward an explanation of nonselfish behavior. *Soc Philos Policy* 10:69–89
- Titmuss R (1970) *The gift relationship*. Allen and Unwin, London
- Willaschek M (2017) *Die Vulkane im Monde und das moralische Gesetz in mir. Immanuel Kant mit 60 Jahren*. Forthcoming in *Festschrift for Birgit Recki*, edited by Stefan Waller