

CESifo Venice Summer Institute

19 - 24 July 2010



“BEHAVIOURAL WELFARE ECONOMICS”

to be held on **21 - 22 July 2010**
on the island of San Servolo in the Bay of Venice, Italy

An Approach to Sympathy and Commitment

*Mireia Jofre-Bonet, Alistair J. McGuire
and Maria Raikou*



An Approach to Sympathy and Commitment

Mireia Jofre-Bonet* Alistair J. McGuire† Maria Raikou ‡

April 2010

Abstract

There is an extensive literature on altruism, particularly with regard to redistribution. There has been less work undertaken on the notion of commitment, actions undertaken at net cost to individuals for others. The aim of this paper is to take the formal structure suggested by Margolis (1982) and Oswald (1982) a step forward. We retain the basic idea that individuals gain welfare from both selfish aspects of private consumption and the provision of "moral" goods. These moral goods come at a cost in terms of individual effort, but give welfare to others, as well as the individual at hand. Thus, moral goods may also involve some commitment on the behalf of individuals as they are willing to aid provision at a cost to themselves. We derive a number of known outcomes but we explicitly incorporate the contribution to the moral good in terms of effort and retain individual commitment as a means of showing how individual behaviour may return a societal first best level of production in the moral good. That said, the model is incomplete in a number of ways and future efforts will concentrate a fuller specification of the compatibility of individual and societal equilibria, the incorporation of uncertainty and the use of commitment in the support for specific commodities such as health care.

*Contact Author: Mireia Jofre-Bonet. Address: Dept. of Economics, 10 Northampton Square, London EC1V 0HB, UK. Tel. +44 (0) 20 7040 0168. Associate Fellow, LSE Health, LSE. E-mail: mireia.jofre-bonet@city.ac.uk.

†LSE Health, LSE

‡LSE Health, LSE We thank Daniel Fujiwara for comments on an earlier version.

1 Introduction and background

Utility, as an expression of preferences, lies at the heart individual choices in economics. A number of well known and important problems are associated with the specific definition of any individual's utility function. Amongst others these include the level and accuracy of informational holdings by individuals, aggregation issues relating to interpersonal comparisons, inter-dependencies, and the role that distributional concerns play in their formation. This paper is concerned with the latter two issues: specifically externalities and philanthropic behaviour. Individual's pursuing utility maximisation may, for example, incorporate an altruistic externality in their actions as a means of satisfying their distributive concerns. Under general conditions it is easily shown that, presuming Nash conjectures across individuals, free-riding pertains and there is under-provision of goods if reliant on altruistically motivated inter-dependences (Sugden, 1982).

Some have argued that this outcome, and related problems, are a result of defining individual utility in a restrictive manner. In particular the definition of utility as it relates to preferences primarily defined across consumption excludes behaviour based on moral rules other than strict utility maximisation. Sen (1977) for example highlights the importance of commitment in individual actions, where utility maximisation is sacrificed at the expense of some moral behaviour. In fact, Pareto himself distinguished between two distinct concepts: Utility, actually referred to by Pareto as "oiphelinite", was the welfare gain related to the narrow consumption process; whereas Pareto's definition of what he referred to as "utility" was welfare associated with the consumption and relative income levels of others (Pareto, 1916). To Pareto, "utility" could be obtained from the operation of particular social processes, income distribution or the provision of publicly provided goods and services which need not contribute to their own selfishly determined "oiphelinite" or welfare. Sugden (1983) argues that such aspects of welfare may reflect "moral rules (that) might be Kantian, but they need not be. For example, a rich rule utilitarian might give money to the poor because he believed that a general transfer of money from the rich to the poor would increase the sum of happiness, even though it would

decrease his own. Or a rich Rawlsian might give money to the poor because he believed that the inequalities in his society were unjust, even though they benefited him personally. In such cases the giver clearly need not experience any altruistic externality - need not feel any sympathetic emotion - in order to recognise his obligation to give". Of course such motivations need not necessarily be confined to income transfers, as Pareto himself explicitly recognised. Individuals may support interventions such as the public provision of health care because they believe it to be correct to do so (Tobin, 1970), or may commit to certain relationships across individuals or institutions even if suffering costs (Akerlof and Kranton, 2000; 2005). The overarching characteristic of "moral" commodities is that they impart welfare over and above that gained from selfish pursuit.

Of course it may, and indeed has been argued that all such behaviour is truly altruistic (Collard, 1983 amongst others). This "warm glow" argument has been challenged, both conceptually and empirically by Andreoni and others (see Andreoni et al, 2005) who suggests that while altruistic behaviour can be seen to exist in experiments even when controlling for "warm-glow" effects, there may well be other, wider motivations for individual behaviour. While the "domain of preferences" remains dominant, there seems enough credence in the concept of moral obligation to attempt to pursue this topic.

In fact Margolis (1982) operationalised this concept, while retaining a general utility approach, suggesting that each individual has two utility functions: One concerned with selfish welfare in which the individual derives satisfaction from private consumption and the other concerned with the welfare derived from group activities. He did not formalise this idea to any great extent however. In an unrelated paper Oswald (1982) did. He essentially shows that these conceptually different utility functions, defined across selfish and non-selfish acts within an individual can support a exposition of selfish and moral behaviour in a manner that is Pareto efficient. He does so through specifying the solution as an internal Nash equilibrium in a co-operative game played by the individual. However, in the Oswald approach the various defined equilibria are defined with respect to marginal utilities and little attention is paid to constraints as his paper merely attempts to highlight the conceptual approach.

In this paper, the formal structure suggested by Margolis (1982) and Oswald (1982) is taken a step forward. We retain the basic idea that individuals gain welfare from both selfish aspects of private consumption and the provision of "moral" goods. These moral goods give welfare to others, as well as the individual at hand and may also involve some commitment on the behalf of individuals in the sense that they are willing to aid provision at an cost (effort) to themselves. The paper is organized as follows: Section 2 sets out the benchmark model and calculates the first-best solution. Section 3 presents the individual solution when no value to the moral good is assumed and the solution when individuals are assumed to attribute some value to the moral good. Section 4, discusses the results and concludes.

2 Benchmark model and first best solution

We model the preferences of an individual over three arguments, a selfish good (s), a moral good (z), and the effort that the contribution to the moral good requires from the individual, c , by extending Oswald's (1982) model . The moral good, z , is a function of the sum of all individuals contributions, $z = z(\sum c_k)$.

As in Oswald's (1982) model, we assume that an individual i has a two-part utility function. The first part, $U_i(\cdot)$, depends only on the "selfish" good she consumes and it defines the optimal consumption of this good. The second part of her utility function, $\Psi_i(\cdot)$, is used to choose the optimal contribution to the moral good z and it depends not only on the individual's own arguments but also on other individuals' utility functions. Thus, we assume that utility is separable in the selfish and the moral good, and the overall utility function of the individual can be thought as:

$$V_i(s_i, c_i) \cong U_i(s_i) + \psi_i(z(\sum c_k), c_i), \psi_j(z(\sum c_k), c_j)) .$$

Thus, the individual's problem is to choose the optimal amount of selfish good, s_i , and the optimal contribution towards the moral good, c_i , so that her two-part utility function is maximized, i.e. the individual chooses s_i such that $s_i^{IN} \in \arg \max\{U_i = U_i(s_i)\}$ and chooses $c_i^{IN} \in \arg \max\{\psi_i = \psi_i(z(\sum c_k), c_i), \psi_j(z(\sum c_k), c_j)\}$. This solution, (s_i^{IN}, c_i^{IN}) , is what Oswald (1982) calls the *internal Nash equilibrium* in the sense that: $V_i(s_i^{IN}, c_i^{IN}) \geq V(s_i, c_i^{IN})$ for $\forall s_i \neq s_i^{IN}$ and $V_i(s_i^{IN}, c_i^{IN}) \geq V_i(s_i^{IN}, c_i)$ for $\forall c_i \neq c_i^{IN}$.

For simplicity, we assume that there are only two individuals, i and j , and that the moral good depends on the sum of both individuals' contributions, $z = z(c_i + c_j)$. We also assume the usual concavity conditions on both parts of the utility function so that: $\frac{\partial U_i}{\partial s_i} \geq 0$, $\frac{\partial^2 U_i}{\partial s_i^2} \leq 0$, $\frac{\partial \psi_i}{\partial z} \geq 0$, $\frac{\partial^2 \psi_i}{\partial z^2} \leq 0$, $\frac{\partial \psi_i}{\partial c_i} \leq 0$, $\frac{\partial^2 \psi_i}{\partial c_i^2} \geq 0$, and $\frac{\partial \psi_i}{\partial \psi_j} \geq 0$, the final condition reflecting the fact that individual i sees their utility increasing as individual j 's utility increases.

2.1 First Best Solution

We find the first best solution by maximizing the welfare function of this economy subject to a minimum level of moral good provision \bar{z} , such that $z(c_i + c_j) \geq \bar{z}$.

For simplicity, we choose a welfare function which is the sum of both individuals' utility functions:

$$W(s_i, s_j, c_i, c_j) = [U_i(s_i) + \psi_i(z(c_i + c_j), c_i), \psi_j(z(c_i + c_j), c_j)] \\ + [U_j(s_j) + \psi_j(z(c_i + c_j), c_j), \psi_i(z(c_i + c_j), c_i)]$$

Thus, the first best solution is $(s_i^{FB}, s_j^{FB}, c_i^{FB}, c_j^{FB}) \in \arg \max W(s_i, s_j, c_i, c_j)$ subject to $z(c_i + c_j) \geq \bar{z}$.

The associated Lagrangian being $\mathcal{L}(s_i, s_j, c_i, c_j, \lambda) = W(s_i, s_j, c_i, c_j) - \lambda(z(c_i + c_j) - \bar{z})$ with correspondent first order conditions (FOC):

$$\frac{\partial \mathcal{L}(\cdot)}{\partial s_k} = \frac{\partial U_k(\cdot)}{\partial s_k} = 0 \text{ for } k = i, j$$

$$\frac{\partial \mathcal{L}(\cdot)}{\partial c_i} = \frac{\partial \psi_i(\cdot)}{\partial c_i} + \frac{\partial \psi_i(\cdot)}{\partial z} \cdot \frac{\partial z}{\partial c_i} + \frac{\partial \psi_i(\cdot)}{\partial \psi_j} \frac{\partial \psi_j(\cdot)}{\partial z} \cdot \frac{\partial z}{\partial c_i} - \lambda \frac{\partial z}{\partial c_i} = 0$$

$$\frac{\partial \mathcal{L}(\cdot)}{\partial c_j} = \frac{\partial \psi_j(\cdot)}{\partial c_j} + \frac{\partial \psi_j(\cdot)}{\partial z} \cdot \frac{\partial z}{\partial c_j} + \frac{\partial \psi_j(\cdot)}{\partial \psi_i} \frac{\partial \psi_i(\cdot)}{\partial z} \cdot \frac{\partial z}{\partial c_j} - \lambda \frac{\partial z}{\partial c_j} = 0$$

$$\frac{\partial \mathcal{L}(\cdot)}{\partial \lambda} = z(c_i + c_j) - \bar{z} = 0$$

$$\text{i.e. } \frac{\partial \psi_i}{\partial c_i} = \left[\lambda - \left(\frac{\partial \psi_i}{\partial z} + \frac{\partial \psi_i(\cdot)}{\partial \psi_j} \frac{\partial \psi_j}{\partial z} \right) \right] \cdot \frac{\partial z}{\partial c_i} \text{ and } \frac{\partial \psi_j}{\partial c_j} = \left[\lambda - \left(\frac{\partial \psi_j}{\partial z} + \frac{\partial \psi_j(\cdot)}{\partial \psi_i} \frac{\partial \psi_i}{\partial z} \right) \right] \cdot \frac{\partial z}{\partial c_j} .$$

Thus given that $\lambda > 0$, $\frac{\partial z}{\partial c_i} > 0$, and $\frac{\partial \psi_i}{\partial z}, \frac{\partial \psi_j}{\partial z} \leq 0$, then $\frac{\partial \psi_i}{\partial c_j} = 0$ whenever $\lambda = \frac{\partial \psi_i}{\partial z} + \frac{\partial \psi_i(\cdot)}{\partial \psi_j} \frac{\partial \psi_j}{\partial z}$. That is, the marginal disutility of the contribution towards the moral commodity for each individual, $[\frac{\partial \psi_k}{\partial c_k}]$, is minimised when the marginal societal value of the moral good $[\lambda]$ is equal to its marginal benefit $[\left(\frac{\partial \psi_k}{\partial z} + \frac{\partial \psi_k(\cdot)}{\partial \psi_{-k}} \frac{\partial \psi_{-k}}{\partial z}\right)]$.

3 PRIVATE SOLUTIONS

In this section we find the equilibrium of our two individuals' society in the following situations: First, when individuals do not confer any value to the moral good, so that they maximize their own utilities without internalizing that the moral good should be above any critical level. Second, we solve the model when they take into account as a restriction that the individual does value the moral good above an individually determined minimal level. Third, we assume identical preferences. Lastly, we assume individuals incur a commitment that costs over and above the private costs associated with the production of the moral good.

(i) We begin with a state of the world where individuals recognise the existence of a "moral" good and indeed that there may be an aggregate gain in welfare if this

good is provided above a minimum level, but do not attempt to ensure this minimum level is obtained. In other words, individuals recognise the existence of moral goods but do not ensure their production. The individual's equilibrium in this case follows from the benchmark model described above. Thus, the problem of each individual i will be to choose $s_i \in \arg \max \{V_i(s_i, c_i) = U_i(s_i) + \psi_i(z(c_i + c_j), c_i), \psi_j(z(c_i + c_j), c_j)\}$ and choose $c_i \in \arg \max \{V_i(s_i, c_i) = U_i(s_i) + \psi_i(z(c_i + c_j), c_i), \psi_j(z(c_i + c_j), c_j)\}$. Given our assumptions, this optimisation problem has associated first order conditions (FOCs):

$$\frac{\partial V_i}{\partial s_i} = \frac{\partial U_i}{\partial s_i} = 0$$

$$\frac{\partial V_i}{\partial c_i} = \frac{\partial \psi_i}{\partial z} \cdot \frac{\partial z}{\partial c_i} + \frac{\partial \psi_i}{\partial c_i} + \frac{\partial \psi_i}{\partial \psi_j} \frac{\partial \psi_j}{\partial z} \cdot \frac{\partial z}{\partial c_i} = 0$$

$$\text{i.e. } \frac{\partial \psi_i}{\partial c_i} = - \left(\frac{\partial \psi_i}{\partial z} + \frac{\partial \psi_i}{\partial \psi_j} \frac{\partial \psi_j}{\partial z} \right) \cdot \frac{\partial z}{\partial c_i} < 0$$

The private equilibrium indicates that -when individuals give no value to the moral good - their levels of contribution towards the moral good is such that its marginal disutility $[\frac{\partial \psi_i}{\partial c_i}]$ is negative. Since the societal value of the moral good $[\lambda]$ is assumed to be positive, the private solution yields a contribution level that is less than the first best outcome. In other, words this private equilibrium replicates the general finding that individuals will free-ride in the provision of moral goods and under-provision from the optimum position occurs.

(ii) In the second scenario, individuals care about the level of provision of the moral good and attempt to ensure that a minimum level, \tilde{z} , is achieved. In this state of the world, individuals consider that there is some opportunity cost to the lack of provision of the moral good and aim to ensure some positive level of this good is provided. Thus, individual i chooses s_i , and c_i such that they maximise the following Lagrangian and satisfy the corresponding FOCs:

$$\mathcal{L}(s_i, c_i) = V_i(s_i, c_i) - (\lambda(z(c_i + c_j) - \tilde{z})) = U_i(s_i) + \psi_i(z(c_i + c_j), c_i) + \psi_j(z(c_i + c_j), c_j) - \lambda(z(c_i + c_j) - \tilde{z})$$

$$\frac{\partial \mathcal{L}_i}{\partial s_i} = \frac{\partial U_i}{\partial s_i} = 0$$

$$\frac{\partial \mathcal{L}_i}{\partial c_i} = \frac{\partial \psi_i}{\partial z} \cdot \frac{\partial z}{\partial c_i} + \frac{\partial \psi_i}{\partial c_i} + \frac{\partial \psi_i}{\partial \psi_j} \frac{\partial \psi_j}{\partial z} \cdot \frac{\partial z}{\partial c_i} - \tilde{\lambda} \cdot \frac{\partial z}{\partial c_i} = 0$$

$$\frac{\partial \mathcal{L}(\cdot)}{\partial \lambda} = z(c_i + c_j) - \tilde{z} = 0$$

$$\text{Thus, } \frac{\partial \psi_i}{\partial c_i} = [\tilde{\lambda} - \left(\frac{\partial \psi_i}{\partial z} + \frac{\partial \psi_i}{\partial \psi_j} \frac{\partial \psi_j}{\partial z} \right)] \cdot \frac{\partial z}{\partial c_i}.$$

Again, whenever $\tilde{\lambda} < \frac{\partial \psi_i}{\partial z} + \frac{\partial \psi_i}{\partial \psi_j} \frac{\partial \psi_j}{\partial z}$, then $\frac{\partial \psi_i}{\partial c_i} < 0$. Or else, if $\tilde{\lambda} = \frac{\partial \psi_i}{\partial z} + \frac{\partial \psi_i}{\partial \psi_j} \frac{\partial \psi_j}{\partial z}$ then $\frac{\partial \psi_i}{\partial c_i} = 0$.

Here the first best societal optimum is not guaranteed. It will only occur if the shadow marginal value of the moral good $[\tilde{\lambda}]$ perceived by the individual coincides with the societal shadow marginal value of the moral good $[\lambda]$ and thus their utility with respect to private contribution to the moral good is minimised $[\frac{\partial \psi_i}{\partial c_i} = 0]$, otherwise there may be under- or over-provision. Given that it is likely that the shadow value of the moral good is liable to be under-valued by the individual compared to the social optimum, [i.e $\tilde{\lambda} < \lambda$], then it is likely that the moral good will be under-provided.

(iii) In this third sub-scenario individuals not only care about the level of provision of the moral good and attempt to ensure that a minimum level, \tilde{z} , is achieved, but they have identical preferences. Individuals again consider there is some opportunity cost to the lack of provision of the moral good and ensure some positive level of this good is provided.

Since both individuals have identical preferences, $c_i = c_j$, so $z = z(2c)$. Thus, the private solution subject to a minimal level \tilde{z} will maximise the following Lagrangian and satisfy its corresponding FOC:

$$\begin{aligned} \mathcal{L}(s_i, c_i) = & V_i(s_i, c_i) - (\lambda(z(2c) - \tilde{z})) = U_i(s_i) + \psi_i(z(2c), c_i) + \psi_j(z(2c), c_j) - \\ & - \lambda(z(2c) - \tilde{z}) \end{aligned}$$

$$\frac{\partial \mathcal{L}_i}{\partial s_i} = \frac{\partial U_i}{\partial s_i} = 0$$

$$\frac{\partial \mathcal{L}_i}{\partial c_i} = \frac{\partial \psi_i}{\partial z} \cdot 2 + \frac{\partial \psi_i}{\partial c_i} + \frac{\partial \psi_i}{\partial \psi_j} \frac{\partial \psi_j}{\partial z} \cdot 2 - \tilde{\lambda} \cdot 2 = 0$$

$$\frac{\partial \mathcal{L}()}{\partial \lambda} = z(c_i + c_j) - \tilde{z} = 0$$

Thus, this means that $\frac{\partial \psi_i}{\partial c_i} = 2[\tilde{\lambda} - \left(\frac{\partial \psi_i}{\partial z} + \frac{\partial \psi_i}{\partial \psi_j} \frac{\partial \psi_j}{\partial z}\right)]$.

Therefore, when $\tilde{\lambda} = \lambda$, that is the shadow (marginal) value of moral good [$\tilde{\lambda}$] perceived by the individual coincides with the societal shadow marginal value of the moral good [λ] and the preferences are the same, $\frac{\partial \psi_i}{\partial c_i}^{FB} = \frac{1}{2} \frac{\partial \psi_i}{\partial c_i}^{PC}$ if $\frac{\partial \psi_i}{\partial c_i} < 0$, and the moral good will be undersupplied. However, if the individuals minimise their private contributions to the moral good [i.e., if $\frac{\partial \psi_i}{\partial c_i} = 0$] then when the shadow valuations are equal, $\tilde{\lambda} = \lambda$, the societal first best is returned. That said, from an individual perspective, individuals may well under-value the shadow value of the moral good and there is no guarantee that $\tilde{\lambda} = \lambda$. In other words, once again the first best is not guaranteed. Thus, even with identical preferences, given that it is probable that the shadow value of the moral good is liable to be under-valued by the individuals compared to the social optimum, [i.e. $\tilde{\lambda} < \lambda$], the moral good is likely to be under-provided.

(iv) A fourth scenario is when individuals commit to a certain minimum contribution \hat{c} in providing the moral good. Thus, the private solution subject to a minimal level of provision of the moral good, \hat{z} , will optimise the following Lagrangian and associated FOCs:

$$\begin{aligned} \mathcal{L}(s_i, c_i) = & V_i(s_i, c_i) - (\lambda(z(c_i+c_j)-\hat{z}) - \rho(c_i-\hat{c})) = U_i(s_i) + \psi_i(z(c_i+c_j), c_i) + \psi_j(z(c_i+c_j), c_j) \\ & - \lambda(z(c_i+c_j)-\hat{z}) - \rho(c_i-\hat{c}). \end{aligned}$$

$$\frac{\partial \mathcal{L}_i}{\partial s_i} = \frac{\partial U_i}{\partial s_i} = 0$$

$$\frac{\partial \mathcal{L}_i}{\partial c_i} = \frac{\partial \psi_i}{\partial z} \cdot \frac{\partial z}{\partial c_i} + \frac{\partial \psi_i}{\partial c_i} + \frac{\partial \psi_i}{\partial \psi_j} \frac{\partial \psi_j}{\partial z} \cdot \frac{\partial z}{\partial c_i} - \hat{\lambda} \cdot \frac{\partial z}{\partial c_i} - \rho = 0$$

$$\frac{\partial \mathcal{L}()}{\partial \lambda} = z(c_i + c_j) - \hat{z} = 0$$

So, $\frac{\partial \psi_i}{\partial c_i} = \left[\hat{\lambda} - \left(\frac{\partial \psi_i}{\partial z} + \frac{\partial \psi_i}{\partial \psi_j} \frac{\partial \psi_j}{\partial z}\right)\right] \cdot \frac{\partial z}{\partial c_i} - \rho$.

Now, again, since $\rho > 0$, whenever $\widehat{\lambda} < \frac{\partial \psi_i}{\partial z} + \frac{\partial \psi_i}{\partial \psi_j} \frac{\partial \psi_j}{\partial z}$, then $\frac{\partial \psi_i}{\partial c_i} = -\rho < 0$ even if $\widehat{\lambda} = \frac{\partial \psi_i}{\partial z} + \frac{\partial \psi_i}{\partial \psi_j} \frac{\partial \psi_j}{\partial z}$, which means that the contribution will be higher under this fourth scenario than in the other private scenarios shown.

Therefore, even if individuals under-value the shadow value of the moral good compared to the optimal valuation [$\widehat{\lambda} < \lambda$], the first-best optimum can still be returned as the shadow cost of commitment [ρ] can compensate for that so that $\left[\widehat{\lambda} - \left(\frac{\partial \psi_i}{\partial z} + \frac{\partial \psi_i}{\partial \psi_j} \frac{\partial \psi_j}{\partial z} \right) \right] \cdot \frac{\partial z}{\partial c_i} - \rho = \left[\lambda - \left(\frac{\partial \psi_i}{\partial z} + \frac{\partial \psi_i}{\partial \psi_j} \frac{\partial \psi_j}{\partial z} \right) \right] \frac{\partial z}{\partial c_i}$. i.e., $\rho = \frac{\left[\lambda - \left(\frac{\partial \psi_i}{\partial z} + \frac{\partial \psi_i}{\partial \psi_j} \frac{\partial \psi_j}{\partial z} \right) \right] \frac{\partial z}{\partial c_i}}{\left[\widehat{\lambda} - \left(\frac{\partial \psi_i}{\partial z} + \frac{\partial \psi_i}{\partial \psi_j} \frac{\partial \psi_j}{\partial z} \right) \right] \cdot \frac{\partial z}{\partial c_i}} = \frac{\lambda - \left(\frac{\partial \psi_i}{\partial z} + \frac{\partial \psi_i}{\partial \psi_j} \frac{\partial \psi_j}{\partial z} \right)}{\widehat{\lambda} - \left(\frac{\partial \psi_i}{\partial z} + \frac{\partial \psi_i}{\partial \psi_j} \frac{\partial \psi_j}{\partial z} \right)}$, which is likely to be greater than 0, even given that the individual under-values the shadow value of the moral good.

Indeed, even if the individuals' under-value the shadow value of the moral good, the societal first best may be returned in a similar manner if the shadow cost of commitment compensates adequately.

Thus, whilst the first best optimum is not guaranteed under commitment, our approach returns the sufficiency we require to accommodate commitment into a traditional framework, which is extended only by the "two-individuals within one" concept proposed by Margolis and Oswald. In this setting, if the shadow cost of commitment [ρ] is chosen appropriately, under-provision will not occur as the committed individual will ensure movement towards the societal optimum.

4 DISCUSSION AND CONCLUSION

It has long been recognised that utility maximisation across arguments confined to the selfish expression of preferences has difficulty expressing the full range of human behaviour. In this paper one such difficulty has been explored: the commitment, at personal cost to utility maximisation, to the production of certain "moral" goods. While we have been rather conceptual in our definition of these "moral" goods we might consider these to

be goods consistent with any behaviour which honours obligation at some individual cost. As Oswald (1982) puts it in discussing behaviour that transcends the selfish, "More generally, of course, people help others, die for their country, care about their children, worry about their neighbours' health, return dropped wallets, and feel for those hurt and hungry on the other side of the world. It is hard to believe that human beings are wholly selfish." (op. cite, p2). Consistent with these examples Sen (1977) notes the assumption of selfishness is used generally in economics for analytical convenience and is not based on empirical observation. Sen, himself introduced the term commitment to differentiate between sympathy, which he defines as synonymous with externality, and behaviour which "involves choosing an action that yields a lower expected welfare than an alternative available action" (op.cite., p328). Of course, in this sense, commitment is closely related to moral choice and obligation, hence our labelling, but we wish the term to encompass a wider range of behaviour. By commitment we merely mean those actions that are consistent with a wider definition of preferences than those solely confined to self-pursuit. In this way we do not rule out moral obligations, but neither do we exclude commitment to the provision of certain types of goods (such as public goods or goods encompassing some basic necessary attribute like health care), or class or institutional loyalties as defined by Akerlof (1983) where self-sacrifice may be involved. In the model above such expression of preferences is merely compatible with the individual incurring a (shadow) cost of commitment. As such the model attempts to provide general formal framework within which to consider a basic deontological moral stance within a utility maximising model. Bordignon (1990) adopted a similar, but less general approach.

We have developed the approach taken by both Margolis (1982) and Oswald (1982) to outline how notions of commitment may be incorporated within a traditional utility maximisation model, thereby highlighting how moral goods may be optimally provided through individual actions. While deriving a number of known outcomes, the model builds on this earlier work by explicitly incorporating contribution to the moral good as a cost to the individual, while retaining individual commitment as a means of showing how individual behaviour may return a societal first best level of production in the moral

good. That said, the model is incomplete in a number of ways and future efforts will concentrate a fuller specification of the compatibility of individual and societal equilibrium, incorporate uncertainty, and be more explicit about the trade-offs the individual is making across goods with different characteristics, in particular be more explicit concerning the motivation of commitment in the support for specific commodities such as health care.

...

5 REFERENCES

Akerlof, G., 1983, "Loyalty filters", *American Economic Review*, 73, 54-63

Akerlof, G. and Kranton, R., 2000, "Economics and identity", *Quarterly Journal of Economics*, 105, 715-753

Akerlof, G. and Kranton, R., 2005, "Identity and the economics of organisations", *Journal of Economic Perspectives*, 19, 9-32

Andreoni, J., Harbaugh, W., and Vesterlund, L., 2005, "Altruism in Experiments" in the *New Palgrave Dictionary of Economics*, (2nd edition), edited by Steven N. Durlauf and Lawrence E. Blume, (Palgrave Macmillan, USA)

Berdignon, M., 1990, "Was Kant right? Voluntary provision of public goods under the principle of unconditional commitment", *Economic Notes*, 3, 342-370

Collard, D., 1983, "On the economics of philanthropy: A comment", *Economic Journal*, 93, 637-8

Margolis, H., 1982, *Selfishness, Altruism and Rationality*, (CUP, Cambridge Mass.)

Oswald, A., 1982, "An approach to the economics of unselfishness", mimeo St John's College, Oxford

Pareto, V., 1916, *Trattato di Sociologia Generale* (published in English as *The Mind and Society* (1935))

Sen, A., 1977, "Rational fools: a critique of the behavioural foundations of economic theory", *Philosophy and Public Affairs*, 6, 317-344

Sugden, R., 1982, "On the economics of philanthropy", *Economic Journal*, 92, 341-350

Sugden, R., 1983, "On the economics of philanthropy: A reply", *Economic Journal*, 93, 639

Sugden, R., 1984, "Reciprocity: the supply of public goods through voluntary contributions", *Economic Journal*, 94, 772-787

Tobin, J., 1970, "On limiting the domain of inequality", *Journal of Law and Economics*, 13, 263-277