Economic Notes by Banca Monte dei Paschi di Siena SpA, vol. 31, no. 1-2002, pp. 175–178

Reply to 'Social Cost and Groves Mechanisms'

GIACOMO BONANNO*

In my 1992 paper in *Economic Notes*, I argued that the traditional heuristic interpretation of taxes in the pivotal mechanism (in terms of the utility loss imposed by the taxed individual on the rest of society) is not correct, since it takes into account only the effect that the individual has on the decision concerning the project and disregards the effect that the same individual has on the taxes paid by the other members of society. Campbell criticizes my observation on two grounds:

1 [Bonanno's] analysis cannot be generalized to the case of positive cost because the allocation that [Bonanno] employs to compute social cost is not feasible in that case, and
2 [Bonanno's] definition is not institution free.

As Campbell himself stresses, 'the mathematics are not in dispute': what is being disputed is the validity of an interpretation. Unlike me, he defends the traditional interpretation.

I would like to reply to both charges. The first is somewhat surprising, since Campbell himself suggests a way in which my calculation of the externality imposed by the taxed individual can be generalized to the case of costly projects. I chose the simple case of a costless, fixed-size project merely to illustrate the logic of my objection in as simple a way as possible. The example I gave was also sufficient to point out the impossibility of designing a Groves mechanism that uses only taxes (thus ruling out positive transfers) and requires an individual to be taxed only if his participation imposes a 'consistently calculated' net cost on the rest of society.

The second objection, namely that the traditional interpretation is superior because it is 'institution free' I find hard to understand and view as untenable. To explain this, I need to review the logic of the observation I made in my 1992 paper. Although that paper focused on the pivotal mechanism, its logic – as Campbell points out – has wider application. Here I will use Vickrey's second price auction to illustrate my point and indeed make use of the example that the author himself gives. An object (e.g. a painting) is auctioned among three individuals. Individual 1 values the object at $1000, individual 2 values it

* Department of Economics, University of California, Davis, CA 95616 - USA. E-mail: gfbonanno@ucdavis.edu

at $500 and individual 3 at $499. Vickrey's auction asks the individuals to submit a bid. The hightest bidder wins the object and she pays not her own bid but the second highest bid. It is well known that in this game bidding one's own true valuation is a dominant strategy (for this to be true, it is essential that the sum of money paid by the winner not be redistributed in any way among the three individuals). If the players play their dominant strategies, individual 1 wins the object for a payment of $500. The traditional interpretation equates this payment to the utility loss imposed on the rest of society by individual 1's participation in the auction. Campbell suggests that 'if agent 1 had not participated, the asset would have gone to agent 2'. Is this a plausible claim? I would argue that it is and, indeed, it is the most plausible claim one can make. What we are trying to do is to evaluate a counterfactual statement whose antecedent is 'if agent 1 had not taken part in the auction'. The philosophy literature on counterfactuals (Lewis, 1973; Stalnaker, 1968) suggests that, to evaluate the truth of the counterfactual 'if A had been the case then B would have been true', one should check if B is true at the A-worlds (that is, possible worlds where A is true) that are *closest to the actual world*. In the actual world, a particular institution, namely the second-price auction, is used to allocate an object among three individuals. The closest possible world to this where individual 1 is not present is a world where the *same* institution (a second-price auction) is used allocate the object between individuals 2 and 3. Indeed, how else could one justify the claim that, without agent 1, the object would go to agent 2?[1] According to the traditional view, since agent 1's participation deprives agent 2 of an object he values at $500, the price agent 1 pays is exactly equal to the utility loss her participation imposes on the rest of society (obviously agent 3 would not get the object in either case and therefore experiences a zero utility loss). The objection I raised in my 1992 paper is that in that counterfactual world where individual 1 is not present, the object would be assigned to individual 2 *in exchange for a payment of $499*. Thus individual 2 would experience, in that counterfactual world, a net increase in utility equal to $1. Hence the utility loss imposed on the rest of society by agent 1's participation is $1 not $500![2] It seems to me that if we take the counterfactual

[1] Trying to unambiguously evaluate the counterfactual in an 'institution free' way seems to me impossible, if not meaningless.

[2] As I pointed out in my 1992 paper (p. 442):

> The traditional interpretation is correct if a new agent is added to the community, one who *has no interest at all in the project and who receives the total amount of the tax* (and, in order not to distort incentives, we also need to add the requirement that the members of society do not care about the welfare of this new agent). Call this new agent individual 0. Then the traditional view is correct, in the sense that the tax paid by an individual under the pivotal mechanism is equal to the externality the individual imposes on the rest of society, *if the latter is interpreted in a broader sense so as to include also individual 0*.

seriously, its logic is inescapable and, to reject it, it is not enough to state that one does not like it.[3]

I would like to conclude by addressing the issue of the correct computation of the true externality caused by individual *i* in the decision concerning a *costly* fixed-size project when the pivotal mechanism is used. Campbell suggests one way of generalizing the computation that I put forward. I would argue that this is not necessarily the correct way to proceed. Here the evaluation of the counterfactual requires a specification of how the cost of the project would be spread among the sub-society obtained by removing individual *i*. Unless a general cost-allocation *rule* is specified (e.g. equal shares) as part of the mechanism, one simply cannot evaluate the truth of the counterfactual.

[3] I don't know how else to interpret Campbell's statement 'We don't want to say that the cost to society of giving the asset to agent 1 is $1 = 500 - 499$. The social opportunity cost of awarding the painting to person 1 is clearly 500, because that is the maximum utility that could be produced by an alternative allocation'. The truth is that a second-price auction between agents 2 and 3 generates a net utility for agent 2 of $1 not $500.

REFERENCES

G. BONANNO (1992), "On the Interpretation of Taxes in the Pivotal Mechanism", *Economic Notes*, 21(3), pp. 438–43.

D. K. LEWIS (1973), *Counterfactuals*, Blackwell, Oxford.

R. C. STALNAKER (1968), "A Theory of Conditionals", in N. Rescher (ed.), *Studies in Logical Theory*, Blackwell, Oxford, pp. 98–112.