Routledge
Taylor & Francis Group

Journal of Economic Methodology

ISSN: 1350-178X (Print) 1469-9427 (Online) Journal homepage: http://www.tandfonline.com/loi/rjec20

Intentions in invisible-hand accounts

Aki Lehtinen

To cite this article: Aki Lehtinen (2009) Intentions in invisible-hand accounts, Journal of Economic Methodology, 16:4, 409-416, DOI: 10.1080/13501780903337354

To link to this article: http://dx.doi.org/10.1080/13501780903337354

Journal of Economic Methodology,
Vol. 16, No. 4, December 2009, 409–416

Intentions in invisible-hand accounts

Aki Lehtinen*

Department of Social and Moral Philosophy, University of Helsinki, PO Box 9, 00014 University of Helsinki, Finland

N. Emrah Aydinonat's account of the invisible-hand is analysed. One of the conditions for unintended social consequences is it requires that individuals' intentions are exclusively directed at the individual level. This condition is weakened in order to accommodate cases in which individuals may also aim at consequences at the social level but the model clearly depicts the invisible hand. Lehtinen's model of counterbalancing strategic votes is proposed as an example that satisfies Aydinonat's conditions, if they are modified as suggested.

Keywords: invisible hand; intentions; mechanisms; counterbalancing; strategic voting

Introduction

N. Emrah Aydinonat presents an informative account of invisible-hand explanations in his recent book (Aydinonat 2008). The purpose of this comment is to propose a slight modification in order to accommodate some cases that his account, by virtue of being too strong, rules out as invisible-hand models.

I propose that Lehtinen's (2006, 2007a, 2007b, 2008) model of *counterbalancing strategic votes* should be taken to provide an invisible-hand account. The models presented in the different papers are different in their details, but I will refer to them collectively as the 'counterbalancing model' because they are all based on the same counterbalancing mechanism. The consequences depicted in this model violate only one of Aydinonat's conditions, and I believe it is better to slightly reformulate his account than to conclude that the model does not count as an invisible-hand account. Given that it is at least a potential invisible-hand model, and has not previously been discussed in the literature, analysing its properties may have some independent relevance to students of economic and social-scientific methodology because models that truly provide invisible-hand accounts are rare.

Aydinonat's account

Some scholars have claimed that it is the particular models that are interesting rather than the details of how to define invisible-hand accounts (Nozick 1994; Hull 1997). However, Aydinonat deserves credit for having formulated a full account that can be analysed item by item. Ultimately the purpose of analysing the definition of such accounts is to illustrate *why* they are interesting, and I propose my modification with this goal in mind. I think Aydinonat also deserves credit for not including various items that have been discussed

*Email: Aki.lehtinen@helsinki.fi

ISSN 1350-178X print/ISSN 1469-9427 online

DOI: 10.1080/13501780903337354
http://www.informaworld.com

in the literature in his definition. It seems right not to require that the unintended consequences are beneficial because the structure of models of the 'invisible backhand' (Brennan and Pettit 1993) seems to be similar to that of models of the invisible-hand. Similarly, although he discusses the end-state and the process interpretations of invisible-hand accounts, he does not require the unintended consequences to arise in a process that takes time.[1]

Various writers have emphasised the need to specify explicitly the relevant mechanism in invisible-hand explanations (see e.g. Ullman-Margalit 1978; Ylikoski 1995). While these authors may be taken to impose conditions for *satisfactory* explanations, my proposal brings mechanisms into the definition of invisible-hand *models*.

Aydinonat sums up his account (2008, pp. 23–24, 89) as follows. An *unintended social consequence* has the following characteristics:

(1) The consequence is located at the social level (or, it is a social consequence).
(2) The consequence was not intended by any individual.
(3) It is mediated through a multiplicity of individuals.
(4) Individual intentions are directed at the individual level.

Two more conditions are needed for unintended social consequences to be *invisible-hand consequences*. In order to exclude cases in which a single individual's actions have social consequences he adds the following condition:

(5) The action of one individual is not sufficient to produce the unintended (social) consequence.

Finally, invisible-hand consequences also satisfy the following two conditions:

(6) Individuals do not pursue the same end collectively (that is, collective intentionality is excluded).
(7) The actions resulting from one single individual's intention cannot affect the social level directly, or in isolation.[2]

My criticism concerns condition (4). According to Aydinonat, this condition 'excludes unintended social consequences that were brought about by actions of individuals who were intending to bring about social consequences' (p. 24). The problem with this condition is that it is too strong: it rules out cases in which we are inclined to say that a model provides an invisible-hand account even though it violates the condition.

I base my argument on an observation that Aydinonat (2008, pp. 18–19) himself makes: individuals may have intentions directed at the individual and the social level at the same time. He provides the example of a person who is capable, without fear of being punished, of changing the social security system in a way that is beneficial to him. My interpretation of this idea is that a person's intentions in a given situation can be *described* in two different ways, one showing the individual's actions as being directed at the individual level, and the other at the social level. There are two ways of interpreting the existence of several descriptions, one that concerns a scholar who is concerned to describe invisible-hand consequences, and one that concerns how the individuals in some real-world situation perceive their situation. I will be employing both the scholar-relative and the subject-relative notion in what follows, and I hope this will not cause confusion. The basic point in my proposal is that condition (4) should allow for cases in which individuals intend to bring about social consequences as long as they are not the same consequences as those that the other conditions concern. One should only require

intentions that are directly involved in the mechanism that brings about the social-level unintended consequence to be described as being directed at the individual level.

The counterbalancing model

In order to prepare the ground for my proposal, allow me to review the main characteristics of the counterbalancing model. Here is Anthony Downs' description of the reasoning that may lead to strategic voting:

> Assume that there are three parties: Right, Center, and Left. Voter X prefers Right to Center and Center to Left, but he believes that Right has the least chance of winning. If he greatly prefers Right to Center and is almost indifferent between Center and Left, he is less likely to switch his vote from Right to Center than if he slightly prefers Right to Center but abhors Left. (Downs 1957, p. 49)

Aydinonat would call this reasoning the 'individual mechanism', and he claims that influential invisible-hand models such as Schelling's checkerboard model contain a plausible individual mechanism (see also Aydinonat 2007). Most of us probably recognise this kind of reasoning, and it seems plausible. What about the idea that strategic voting has beneficial consequences, however? Are voting results better, in terms of the voting individuals' preferences on the aggregate level, when some voters engage in strategic voting than when everybody votes sincerely? The counterbalancing model suggests the surprising result that the aggregate-level consequences are better when people vote strategically than when they do not. The beneficiality of strategic voting must appear paradoxical because it is defined as voting for a candidate that one does not consider the best under most voting rules. Even though individuals vote for alternatives that are not the best in terms of their preference orderings, the best outcome in terms of the social ordering emerges. The result thus has the air of a paradox that is often associated with invisible-hand accounts (e.g. Ullman-Margalit 1978).

In the following I will show that this model satisfies all other conditions except condition (4). The main result is that strategic voting is likely to be beneficial in the sense that the frequency of simulated voting games in which the best candidate is selected is higher if voters engage in strategic manipulation than if everyone votes sincerely. The best candidate is defined as the 'utilitarian winner', i.e. the candidate for which the aggregate sum of utility is the highest. Such a candidate is thus defined on the basis of the very same preference profile that underlies voting decisions.

The counterbalancing model distinguishes between two different behavioural assumptions: sincere voting behaviour (SV behaviour) and expected utility-maximising behaviour (EU behaviour). Under the former voters always vote sincerely, while under the latter they cast a sincere or a strategic vote, depending on their incentives. *Utilitarian efficiency* is defined as the percentage of voting games in which the utilitarian winner is selected. EU behaviour is *welfare-increasing* if the utilitarian efficiency is higher than under SV behaviour. If this is the case, one could also say that strategic voting is welfare-increasing. The main result may now be formulated more succinctly: strategic voting is welfare-increasing under a wide variety of parameter values of the model, and under a wide variety of voting rules.

Let us now apply Aydinonat's conditions to the counterbalancing model. The *invisible-hand consequence* is that the utilitarian winner is selected rather than some other candidate, and this is due to strategic voting. Note that, in addition to depicting invisible-hand consequences, models may have what I call *invisible-hand model-results*. The invisible-hand model-result is that strategic voting is welfare-increasing. The invisible-hand model-result thus generalises the invisible-hand consequences here.

It seems natural to say that invisible-hand models primarily explain results rather than consequences, but I see no reason why they could not explain both. In the following, I show that invisible-hand consequences that are being depicted in the counterbalancing model satisfy Aydinonat's conditions (1) to (3) and (5) to (7).

(1) The consequence is surely located on the social level because the utilitarian winner is defined on the basis of the whole electorate and is not even describable on the individual level.

(2) Even though strategic voters surely intend to bring it about that the utilitarian winner is selected rather than the candidate that would have resulted from sincere voting, the invisible-hand consequence is not intended by any voter because they do not perceive the vote as being between the 'utilitarian winner' and 'some other candidate', but rather between, say, candidate 'X' and candidate 'Y'. Voters may well intend to maximise the chances that candidate X is selected while they are making their strategic-voting decision. Such intentions do not compromise the unintended status of the consequence even though they may be taken to be directed at the social level. The reason for this is that intentional states are *intensional*: the truth-values of statements concerning intentional states depend on the description of those states. In the famous example Lois Lane loves Superman but not Clark Kent. Similarly, even if X happens to be the utilitarian winner and the outcome would have been Y under SV behaviour but turns out to be X under EU behaviour, the status of the invisible-hand consequence is not compromised because voters gave their vote to candidate X under the 'X' description rather than under the 'utilitarian winner' description.

    The invisible-hand model-result is self-evidently not intended by any individual voter. No individual voter intends to bring about higher utilitarian efficiency. In fact, this result could not have been intended by any individual because the literature on voting before this model was put forward usually assumed that strategic voting was harmful. Individual voters did not even know that their actions might have beneficial consequences, not just for themselves but also for the whole society, and the scholars seemed to reinforce the idea that strategic voting was morally condemnable.[3]

(3) The consequence is mediated through a multiplicity of individuals because the very description of voting requires reference to several.

(5) Clearly the consequence cannot be brought about by just one individual.

(6) Voters do not pursue the same end collectively.

(7) One individual's actions are not sufficient to determine the consequence.

A modification

Let us now look at condition (4), that individual intentions are directed at the individual level. There are at least three different kinds of intentions involved in voting (see also Mäki 1991). When people engage in strategic voting their expected utility calculus involves only their own utilities and beliefs. In this sense, their intentions are directed at the individual level. On the other hand, the very purpose of voting is to bring about changes on the social level, to change or retain the current policies, and in this sense their intentions are also directed at the social level. Voters may also intend to maximise the chances that a satisfactory policy or candidate be selected rather than some other policy or candidate, and this intention may also be taken to be directed at the social level. Strategic voting may be taken to be one possible means for attaining this end. Voters' intentions may thus be described as being directed

at both the individual and the aggregate level. They are individual-level-directed under one description and aggregate-level-directed under another.

It could be argued that there is no need to modify condition (4) because the counterbalancing model already satisfies it. Even though we know that voters may wish to bring about changes on the social level, these aims are not, in fact, part of the description of the counterbalancing model.[4] However, given that intentions directed at the social level constitute part of the situation voters are facing, such intentions should not be ruled out in the definition of unintended social consequences even though they are not part of the description of the model that accounts for such consequences. Aydinonat takes invisible-hand consequences as explananda for invisible-hand explanations. If we wish to be faithful to Aydinonat's viewpoint, although unintended consequences should be *explained* through mechanisms, they should be formulated as part of an account of *invisible-hand models* rather than as part of the definition of invisible-hand consequences. We should thus not say that the counterbalancing *model* already satisfies condition (4) due to the fact that it does not describe intentions that are directed at the social level[5] because that would imply that the *definition* of unintended social consequences would depend on the properties of explanatory models, i.e. whether or not they describe certain kinds of intentions. Whether such descriptions are contained in a model is a matter of satisfying conditions for invisible-hand *models* rather than for invisible-hand *consequences*.

Given that the purpose of condition (4) is to exclude cases in which individuals intend to bring about social consequences, we must require that intentions directed at the social level are not responsible for the unintended social consequences. On the basis of these considerations, I propose to modify condition (4) as follows:

(4′) At least some individual intentions are directed at the individual level. If individuals also have intentions directed at the social level they are not responsible for the *same* social-level consequences that concern conditions (1) to (3), and (5) to (7).

This condition is deliberately designed to merely define unintended social consequences without explaining them.

Invisible-hand models

Even though it is perfectly possible to define invisible-hand consequences without any effort at explaining them, ultimately they are recognised by way of inspecting the properties of models that account for them. An account of *invisible-hand models* should thus contain a condition that provides this explanatory part, and I will now formulate such a condition. Whether a model qualifies as an invisible-hand account depends on the individual intentions that are necessary for the mechanism that brings about the aggregate-level results. These intentions must be directed at the individual level. The agents may have intentions directed at the social level in an invisible-hand account if the mechanism that brings about the invisible-hand result could function as and be described in terms of intentions that are directed *only* at the individual level.

Consider the counterbalancing model as an example. Thus far I have stated nothing about why strategic voting is welfare-increasing. The model simply specifies an expected utility model, and compares the voting outcomes in terms of the underlying preferences. The *mechanism* of *counterbalancing strategic votes* needs to be invoked in order to explain the outcome.

Preference intensities are a matter of how *much* an individual prefers one candidate to another. Downs expresses the idea that preference intensities for the choice alternatives and beliefs concerning their chances of winning are crucial. The counterbalancing

mechanism explains what kind of aggregate-level consequences strategic voting is likely to have. In an electorate with a large number of voters who make their decisions based on incomplete information, some individuals typically have an incentive to vote strategically for a candidate, say X, but at the same time others have an incentive to strategically desert this very same candidate X by voting for some other alternative Y. The conditions for voting strategically in terms of the structure of voters' preferences and their information are different under different voting rules (see Lehtinen's papers), but they always indicate that, *ceteris paribus*, an alternative that has a large sum of utility is likely to gain many and lose few strategic votes. Therefore, many strategic votes for the utilitarian winner are likely to be counterbalanced by few strategic votes against. This explains why utilitarian winners are more likely to be selected if voters engage in strategic voting.

Counterbalancing is a 'transformational' mechanism because it shows how 'a number of individuals, through their actions and interactions, generate macro-level outcomes' (Hedström and Swedberg 1998, p. 22).[6] This kind of mechanism shows how individual actions are transformed into a collective outcome. The need for transformational mechanisms in invisible-hand accounts seems highly natural. If there is to be a model that specifies how consequences emerge on the social level in such a way that they are the result of many individuals' actions, it seems clear that it should contain at least one transformational mechanism.[7]

Aydinonat proposes a set of criteria for good invisible-hand explanations (2008, p. 90). While many of these criteria are based on articulating mechanisms, he does not make requirements concerning the role of intentions in those mechanisms. I find this unsatisfactory because this role is precisely what provides us with a way of identifying invisible-hand models. *Invisible-hand accounts* should always identify a transformational mechanism, and I propose the following identification condition:

(I) The functioning of the transformational mechanism that brings about the unintended social consequences depends only on intentions that are directed at the individual level.

Note that condition (I) is applicable only if the notion of invisible-hand consequences has already been defined, and it cannot be used for defining such consequences.

This condition allows individuals to have intentions that are directed at the social level as long as the transformational mechanism depends only on intentions that are directed at the individual level. If condition (I) is applied to the counterbalancing model, it means the following. Individual intentions that are necessary for the transformational mechanism are defined by their expected-utility calculations, which determine whether they vote strategically or sincerely. These expected-utility calculations are necessary for the functioning of the counterbalancing mechanism. It constitutes the transformational mechanism that allows for explaining the invisible-hand consequence.

When voters maximise expected utility, they aim at obtaining the best possible outcome in terms of their own preferences. They do have intentions directed at the social level, but not in a way that would compromise the status of the counterbalancing model as an invisible-hand account. They do not aim at maximising the chances that the utilitarian winner is selected (even though, as explained above, their behaviour may amount to acting in such a way), and they do not aim at making strategic voting welfare-increasing.

Although voters' preferences could be described as being *about* social-level policies, when they make their decisions concerning whether or not to vote strategically, their intentions could be described as being directed at the individual level because describing

the reasons that lead some to vote strategically and some to vote sincerely does not require any reference to the social level. These decisions are determined by voters' preferences and beliefs, and not by the reasons *why* they have such preferences. In other words, in order to describe the counterbalancing mechanism it is not necessary to know what kind of social policies people would like to see in place. The invisible-hand consequence would come about even if people had different social-level intentions, i.e. if they wanted to have different policies in place, and even if they did not want to make any social changes at all, say, if they voted out of habit.

Conclusion

It is appropriate to weaken condition (4) into (4′) because the counterbalancing model shows that not all intentions need to be directed at the individual level in an invisible-hand account. The consequences depicted in the counterbalancing model satisfy all of Aydinonat's conditions if condition (4) is modified along the lines suggested here. To account for the invisible-hand consequences, it is appropriate to add a condition that shows how unintended social-level outcomes are generated from intentions directed at the individual level. Such a condition must specify that the explanatory mechanism depends only on intentions directed at the individual level. Aydinonat's account of invisible-hand models, if appropriately modified, is coherent and illuminating.

Acknowledgements

I am grateful to Emrah Aydinonat, Uskali Mäki, Jaakko Kuorikoski, Cecilia Therman and two anonymous referees for their helpful comments on the manuscript. The usual disclaimer applies.

Notes

1. Counterbalancing is a mechanism that works in an instant, or at least the possibility of its functioning does not depend on there being a temporal dimension to it. The counterbalancing model thus fits Aydinonat's account better than earlier ones in terms of some conditions.
2. This last condition is not mentioned in the version on pages 23–25. I have my doubts about whether it is redundant given that condition (5) seems to be stating something similar, but I will not discuss this matter here.
3. I agree, however, with Aydinonat (2008, pp. 14, 25) that unforeseen consequences are not always unintended, and vice versa.
4. This point derives from a referee's comment, but it was not formulated in exactly these terms. I took the liberty of giving it my own interpretation.
5. Another referee suggested that aiming to maximise the chances that a particular candidate is selected is to be counted as an intention directed at the social level, so that the model in fact also contains such intentions, and condition (4) is violated.
6. Hedström and Swedberg's typology also contains 'situational' mechanisms and 'action-formation' mechanisms. The former specifies how macro-level events or conditions affect the individual, and the latter shows how a specific combination of individual desires, beliefs and action opportunities generate a specific action. The latter thus corresponds roughly to Aydinonat's 'individual' mechanisms.
7. This is not to deny that other kinds of mechanisms may also be part of the model. We have already seen that the counterbalancing model also contains an individual mechanism.

References

Aydinonat, N.E. (2007), 'Models, Conjectures and Exploration: An Analysis of Schelling's Checkerboard Model of Residential Segregation', *Journal of Economic Methodology*, 14(4), 429–454.

——— (2008), *The Invisible Hand in Economics: How Economists Explain Unintended Social Consequences*, London: Routledge.
Brennan, G., and Pettit, P. (1993), 'Hands Invisible and Intangible', *Synthese*, 94, 191–225.
Downs, A. (1957), *An Economic Theory of Democracy*, New York: Harper.
Hedström, P., and Swedberg, R. (1998), 'Social Mechanisms: An Introductory Essay', in *Social Mechanisms: An Analytical Approach to Social Theory* (1st ed.), eds. P. Hedström and R. Swedberg, Cambridge: Cambridge University Press, pp. 1–30.
Hull, D.L. (1997), 'What's Wrong With Invisible-Hand Explanations?', *Philosophy of Science*, 64(4), 117–126.
Lehtinen, A. (2006), 'The Welfare Consequences of Strategic Voting in Plurality and Runoff Rules', University of Helsinki, Department of Social and Moral Philosophy, *http://www.valt.helsinki.fi/staff/alehtine/pandr.pdf*.
——— (2007a), 'The Borda Rule Is Also Intended for Dishonest Men', *Public Choice*, 133(1–2), 73–90.
——— (2007b), 'The Welfare Consequences of Strategic Voting in Two Commonly Used Parliamentary Agendas', *Theory and Decision*, 63(1), 1–40.
——— (2008), 'The Welfare Consequences of Strategic Behaviour Under Approval and Plurality Voting', *European Journal of Political Economy*, 24, 688–704.
Mäki, U. (1991), 'Practical Syllogism, Entrepreneurship and the Invisible Hand: A Critique of the Analytic Hermeneutics of G.H. Von Wright', in *Economics and Hermeneutics*, ed. D. Lavoie, London: Routledge, pp. 149–176.
Nozick, R. (1994), 'Invisible-Hand Explanations', *The American Economic Review, Papers and Proceedings*, 84(2), 314–318.
Ullman-Margalit, E. (1978), 'Invisible-Hand Explanations', *Synthese*, 39, 263–291.
Ylikoski, P. (1995), 'The Invisible Hand and Science', *Science Studies*, 8(2), 32–43.