

# Adaptive Preferences, Normative Individualism, and Individual Liberty

*Comment*

by

Christoph Vanberg\*

## 1 Introduction

*Welfare economics* is concerned with making judgements of the form “Policy  $z$  is better than policy  $x$ ”. The implicit purpose of doing so is to provide advice to someone (a policy maker) in a position to *choose* between these alternatives. Carl Christian von Weizscker (henceforth CCvW) defines *normative individualism* as an approach to this problem which demands that all such judgements be based exclusively on the “degree to which *wants* of citizens are satisfied,” where these “wants” are reflected in *individual preferences* with respect to the options under consideration.

A large literature in Social Choice Theory discusses the difficulties that arise in welfare economic analysis when individual interests differ, i.e. in the face of *interpersonal* preference heterogeneity. If Mr. 1 prefers policy  $x$  to policy  $z$  while Mr. 2 prefers  $z$  to  $x$ , which option should the welfare economist judge to be ‘*socially*’ better? Traditional Welfare Economics solves this problem by *aggregating* individual preferences to yield a *social preference* over the objects of choice.<sup>1</sup>

CCvW is concerned with a different challenge to the Welfare Economic enterprise. The complication that he addresses involves *intrapersonal* preference heterogeneity, and can arise even if the policies being considered concern only a single individual. Despite this difference, the problem is somewhat similar to that which arises in the presence of interpersonal preference heterogeneity: If  $x$  is preferred to  $z$  when  $x$  obtains, and  $z$  is preferred to  $x$  when  $z$  obtains, which option should the welfare economist judge to be ‘*truly*’ better? CCvW argues that this questions can be answered in a way that is consistent with normative individualism. His analysis involves a positive theory of

---

\*Department of Economics, University of Heidelberg.

<sup>1</sup>James Buchanan (1954) has argued that this solution is not in fact compatible with normative individualism, since unless all agree, any “social” ranking over the objects of choice will be inconsistent with the preferences of at least one individual. Buchanan’s own solution to this dilemma is to reject the very notion of deriving a single “social preference” over the set of alternatives. Instead he investigates the problem of evaluating alternative *rules and procedures* through which (collective) choices between alternatives are made. The central legitimizing principle in this *constitutional economics* perspective is the *voluntary consent* of the individuals involved.

preference change and a specific (but partially implicit) interpretation of normative individualism.

My comment is organized as follows. The next section discusses CCvW's positive theory of preference change. After offering some comments on its central assumptions, I turn to the normative implications of the theory in Section 3.

## 2 *The Positive Theory of Endogenous Preferences*

The central assumption in CCvW's positive theory of endogenous preferences is that of *adaptive preferences*. The main implication of this assumption is the existence of an exogenous '*limit preference*' with the property that alternative  $z$  is preferred to alternative  $x$  if and only if there exists a finite *improving sequence* leading from  $x$  to  $z$ . The other important element of CCvW's positive theory is the *Improvement Axiom*, which states that improving sequences are preferred to stationary paths. I will discuss each of these assumptions in turn.

### 2.1 *Adaptive preferences*

The assumption of adaptive preferences essentially says that individuals value goods (and perhaps more broadly, behaviors) more highly if they have chosen them in the past. CCvW offers two arguments in support of the empirical hypothesis that preferences are adaptive. The first is that adaptive preferences are descriptive of habit formation. The second is that they explain a widespread predictability and stability of behavior and institutions.

The term "habit formation" refers to the development of an increased tendency, on the part of an individual, to exhibit behaviors that he or she has engaged in in the past. In CCvW's theory, a person who consumes certain goods at time  $t$  will value those goods more highly (relative to goods not consumed at time  $t$ ) in period  $t + 1$ . This assumption seems psychologically plausible in many contexts, as for example when a person prefers the type of food that she is accustomed to. In many contexts, however, the reinforcement of past behaviors is likely to depend on (external or internal) rewards generated thereby. Although such learning mechanisms may be compatible with CCvW's formal theory, his analysis makes no mention of such rewards, as distinct from the endogenously changing preferences themselves. A more explicit theory of habit formation through preference change might treat the 'preference system' as a learning mechanisms. Within such a theory, I believe that a distinction would need to be made between the agent's *internal preferences* (or "utility function") with respect to the objects of choice, and the (external and internal) *rewards* that behaviors governed by those preferences generate.<sup>2</sup>

---

<sup>2</sup>Various authors have investigated the importance of this distinction in strategic contexts. In such contexts, attaching high 'internal' utility to actions that are 'objectively' harmful can constitute a beneficial commitment device. See, for example, Frank (1988) and Güth and Yari (1992).

Next, consider the argument that the observed degree of predictability of human behavior, as well as the overall stability of institutions, constitute something of an “indirect proof” that preferences are adaptive in CCvW’s sense. This argument raises two questions. The first concerns the empirical status of the statement that human conduct is “predictable” and institutions are largely “stable”. Clearly, much of human conduct is *not* predictable, and likewise both formal and informal institutions are constantly changing. And indeed CCvW’s theory would seem consistent with such changes, at least in the “short run”. If change, rather than stability, is the rule, then the central prediction of CCvW’s theory would seem to be that change is “noncircular”. The second question concerns the specification of alternative hypotheses. As it stands, the “indirect proof” is by contradiction: If preferences were “anti-adaptive,” then human conduct would be completely unpredictable. To conclude that preferences must be adaptive implicitly assumes that human behavior is governed *either* by “adaptive” or by “anti-adaptive” preferences. However the argument does not establish that there cannot be other theories of behavior that do not satisfy the assumption of adaptive preferences but are nonetheless consistent with some degree of predictability in human conduct.

Despite these objections, adaptive preferences seem to capture important features of human psychology. As documented, among others, by Kahnemann et al (1991), the majority of people exhibit certain “biases” consistent with an overall tendency to value options chosen in the past more highly than others. And indeed everyday experience suggests that much of our behavior is governed by a fundamental desire to obtain more of what we are used to. Finally, as indicated above, more elaborate theories of habit formation involving reinforcement may well be consistent with the assumption of adaptive preferences.<sup>3</sup>

Before moving on, let me briefly restate what the assumption of adaptive preferences, if accepted, implies for the argument being constructed. The conclusions that follow are summarized in CCvW’s Theorems 1 and 2. Theorem 1 says that all ‘improving sequences’ are non-circular. Theorem 2 says that this non-circularity (along with further technical conditions) implies the convergence of preferences to a long-run preference. This “limit preference” has the interesting property of being an ‘indicator function’ that tells us, for each pair of options  $x$  and  $z$ , whether or not there exists an improving sequence leading from  $x$  to  $z$ .

## 2.2 The Improvement Axiom

The second important ingredient in CCvW’s positive theory is the Improvement Axiom. It states that an individual faced with a choice between a stationary consumption path  $\{x, x, x, \dots\}$  and an improving path  $\{x, y, z, \dots\}$  prefers the improving path. The normative implications of this axiom will be discussed in the next section. Here I will discuss its plausibility as an empirical statement. Before doing so, it also seems important to point out that the Improvement Axiom introduces an entirely new kind of “preference” into the analysis.

---

<sup>3</sup>In particular this would seem likely if the environment is sufficiently stable.

Outside of this axiom, CCvW’s agents are assumed to face choices between what he calls *consumption bundles*. However we might interpret this term, the formal analysis assumes that the alternatives being evaluated are associated with specific *points in time*. At each such point, both the individual and the policy maker are assumed to evaluate and / or choose among the in-period alternatives based on (‘momentary’) preference relations over those alternatives.<sup>4</sup> The Improvement Axiom deviates significantly from this setup. Suddenly, CCvW’s agent is faced with a (hypothetical?) choice between *sequences* of consumption bundles. His “preference” for the improving path is, both formally and substantively, something entirely distinct from a preference over in-period alternatives. This raises questions concerning the nature of the choices the analysis is meant to shed light on.<sup>5</sup>

Let us now consider the Improvement Axiom as an empirical hypothesis. By definition, an *improving path* has the property that the agent prefers each option chosen along the path to the one which precedes it. Thus it may at first appear innocuous to assume that she will prefer to embark on such a path if faced with such a choice. CCvW provides examples of persons acquiring desirable habits and tastes, such as learning to (enjoy to) play the piano. Indeed it seems plausible in such contexts that an individual would consent to an improving path. On the other hand, it is easy to come up with examples of decidedly *undesirable* habit formation, for example in the context of alcohol or tobacco consumption. While the assumption of adaptive preferences seems to apply well to those contexts, the Improvement Axiom appears rather less plausible.

### 2.3 Implications

If the assumptions entering CCvW’s positive theory are accepted, the following implications are relevant for the normative interpretation. First, preferences converge to a unique limit preference in the long run. Second, the limit preference has the property that  $z$  is preferred to  $x$  if and only if there exists a finite improving sequence leading from  $x$  to  $z$ . Third, the individual prefers an improving sequence to a stationary path, and would therefore consent to embarking on such a sequence. In the next section, I will take these results as given and discuss the normative conclusions that CCvW draws from them.

### 3 Welfare economics and normative individualism

CCvW’S main *normative* conclusion is that the “exogeneity of the [limit preference] is the platform by which we are then able to do welfare economics.” The precise meaning of this conclusion remains rather implicit, however. I will argue that a possible interpretation of CCvW’s theory is that  $z$  is to be judged ‘*truly better*’ than  $x$  whenever the limit

---

<sup>4</sup>Even the “welfare economist” is assumed to be seeking to construct such a ranking over the space of consumption bundles.

<sup>5</sup>If the individual (or policy maker) can choose between *sequences* of alternatives, it is not clear why the welfare economist should seek to provide a single ranking over the alternatives themselves.

preference ranks  $z$  better than  $x$ . This interpretation of the theory, however, could be used to justify coercion, as when the individual is forced to consume  $z$  even if she (‘erronously’) prefers  $x$ . This ‘paternalistic’ interpretation of CCvW’s arguments appears to be inconsistent with a normative individualism that respects individual liberty and freedom of choice.

To understand what it might mean to “do welfare economics” using the limit preference, we may consider a passage at the beginning of the article. There, CCvW describes a situation involving three options  $x$ ,  $y$ , and  $z$ , with the property that  $x$  is initially preferred to  $z$ , but  $\{x, y, z\}$  is an improving sequence. He then argues that “in a sense, although an immediate jump from  $x$  to  $z$  is not an improvement,  $z$  may be considered ‘better’ than  $x$ .”<sup>6</sup> This passage would seem to suggest that the *basic principle* underlying CCvW’s normative approach is that the welfare economist can judge the end point of an improving sequence to be “better than” its beginning, even if the individual himself does not share that judgement in the beginning. The *positive* theory of adaptive preferences establishes that the application of this principle yields a unique (exogenous) and transitive ordering over the set of alternatives.<sup>7</sup>

What would follow from this? If the limit preference is ‘correct,’ then it would seem that the individual is ‘wrong’ about his true interests under some circumstances. Although he would consent to gradual change along an improving path towards that which is truly better for him, he would refuse to ‘jump’ immediately to the end of that path. If the Welfare Economist uses the individual’s ‘true’ preferences for policy evaluation, he may therefore recommend *imposing* the jump - forcing the individual to choose the option that is ‘truly’ better, as when a parent forces her child to play the piano.

In his larger companion paper, CCvW (2013) identifies normative individualism with a concern for “freedom and liberty,” declaring that “we understand liberty to be a situation in which an agent has the choice (...) [and] his choice is justified (...) simply [because] it is his choice.” This would seem to suggest that CCvW wishes to endorse what Viktor Vanberg (2005) has referred to “choice individualism”. In contrast to “utility individualism”, which demands that welfare judgements reflect the extent to which different alternatives serve the (‘true’) interests of the individual, choice individualism

---

<sup>6</sup>It seems important to point out that this sentence involves judgements being made from two very different perspectives. The words ‘not an improvement’ refer to the individual’s judgement concerning the jump from  $x$  to  $z$ . The words “better than,” in contrast, refer to a judgement made by some other entity - presumably the “welfare economist”.

<sup>7</sup>It appears that the Improvement Axiom is meant to provide a foundation for this basic principle. As I have pointed out above, this axiom makes a statement about the individual’s preferences over alternative *sequences* of alternatives. In concluding that this axiom allows the “onlooking welfare economist (...) to evaluate the consumption bundles,” CCvW seems to be moving from individual preferences over *sequences* of alternatives to welfare judgements concerning the alternatives themselves. This suggests that an implicit axiom is being employed according to which the end point of an *individually preferred sequence* can be judged by the welfare economist to be “better than” its beginning, even if the individual himself does not share that judgement in the beginning

demands that the individual himself should be allowed to choose. Which of two alternatives,  $x$  or  $z$ , is ‘truly’ better is irrelevant from this perspective, and so it is unnecessary for there to exist a single ‘correct’ ranking of alternatives. If this is the normative perspective that CCvW favors, it would seem that it can be maintained irrespective of whether preferences are adaptive in his sense.

### *References*

BUCHANAN, JAMES (1954). “Social Choice, Democracy, and Free Markets,” *Journal of Political Economy*, 64 (2), 114-123.

FRANK, ROBERT (1988). *Passions within Reason - The Strategic Role of the Emotions*. Norton.

GÜETH, WERNER AND MENAHEM YARI (1992). “Explaining reciprocal behavior in simple strategic games: An evolutionary approach,” In: *Explaining process and change: Approaches to evolutionary economics*, Witt U. (ed). Michigan University Press, 23 - 34.

VANBERG, VIKTOR (2005). “Market and state: the perspective of constitutional political economy,” *Journal of Institutional Economics* 1, 23-49.

VON WEIZSÄCKER, CARL CHRISTIAN (2013). “Freedom, Wealth, and Adaptive Preferences,” *mimeo*, Max-Planck-Institute for Research on Collective Goods.