

The Merits of Merit Wants

Richard Sturn

Abstract Merit wants are a multi-faceted concept cutting through a complex array of problems associated with different levels of analysis. They are considered in this paper as a shorthand notion for concerns that are respectable and important, assuming a broadly individualist conception of welfare. So why are merit wants not a firmly established part of modern normative economics, given that simplifying, but still meaningful notions are suitable as conceptual starting point for a research program? In this paper I try to link the answer to this question with making explicit three levels of problems (limits of reason, higher order preferences, collective choice) which may be useful to locate and scrutinize various interpretations of and approaches to merit wants.

Keywords Bounded rationality • Communal preferences • Higher order preferences • Merit wants/goods • Social choice

1 Introduction

Merit wants are a multi-faceted concept cutting through a complex array of problems. Those problems are associated with different levels of analysis, which have in common that they are somehow related to the limits of standard concepts of consumer sovereignty. They either refer to choice situations where there are good reasons to conceptualize human agents not as consumers (but as individuals expressing their values, or as citizens), or they relate to situations where they are conceptualized as consumers, but apparently lack sovereignty in the appropriate sense. Merit wants can be seen as a shorthand notion for various respectable and important concerns; those concerns are particularly important in the context of theoretical frameworks taking human agency seriously as a foundational ingredient of evaluation and explanation of social states. Merit wants are not a normatively

R. Sturn (✉)

Institute of Public Economics, University of Graz, Universitaetsstrasse 15, 8010 Graz, Austria
e-mail: richard.sturn@uni-graz.at

empty box, as they were occasionally called [20]. Or so I will argue in this paper.

But why are merit wants not a firmly established part of modern normative economics, given that simplifying, but still meaningful notions are in principle suitable as a conceptual core of research programs? Is the ambiguous status of this concept perhaps caused by the excessive complexity lurking in the background of the variety of interpretations? Does under-conceptualized complexity prevent crucial issues from being made explicit at the proper level of analysis? If that is the case (and I will argue that it is the case to a certain extent), the kind of tension aptly summarized by Kaushik Basu [6, p. 220] should come as no surprise: “that some goods should be treated as merit goods most of us agree, but fail to express why we do so, with any rigour”. Referring to the dubious status of merit goods, Nick Baigent [4, p. 301] observes that they “have always been postulated in an ad hoc way without any real justification.”

In this paper I try to link the answer to the kind of questions sketched by way of introduction with making explicit three levels of analysis: call them for short (1) “behavioral limits of reason”, (2) “higher order preferences”, and (3) “collective choice”. By means of (1)–(3), various interpretations of and approaches to merit wants can be brought into some kind of preliminary structure. The social choice aspects of merit wants has been addressed by Nick Baigent [4], suggesting a possibility of integration of merit wants in the framework of Arrowian Social Choice theory. Issues related to the second level of analysis (referring to the epistemological status of preferences in general and higher order preferences in particular) are discussed in another paper by Nick Baigent [5], entitled “Behind the veil of preferences”. Behavioral limits of reason are the core of Munro’s [23] discussion of merit wants, taking on board pertinent findings of experimental and psychological economics.

Discussion of the crucial issues underlying merit wants shows that it is misleading to think of the disputes regarding merit wants as exhibiting a divide between the clear water of individual sovereignty and the muddy mixtures of collectivism, paternalism, and authoritarian elitism (words which are often used when referring to the problematic status of merit wants). I am arguing that merit wants as discussed by Musgrave and others reveal a divide within the individualist camp, amongst economists and philosophers sticking to liberal principles. While conceptions of merit wants are certainly not without difficulties, steering clear of the problems which they are meant to address is neither equivalent to their solution, nor can it be justified in terms of “staying at the safe side.”

Moreover, in a perspective emphasizing social choice as a dimension of merit wants, their recent revival in the wake of behavioral economics and libertarian paternalism (cf. e.g. Thaler and Sunstein [38, 39] and Munro [23]) must be regarded as conceptually unfinished business as yet. This revival is focused on findings of behavioral economics that provide evidence for persistent “mistakes” in individual choice behavior. To be sure, the methods and findings of behavioral economics allow for important progress regarding the empirics of context-dependent behavior

and pertinent limits of reason: this is a step forward in Mill's [21, V.xi] program of a situation-specific diagnosis of classes of circumstances rendering consumer sovereignty problematic. I will come back to that in Sects. 3 and 6, where some advantages of a conceptually complete theory of merit wants are suggested.

The remainder of this paper is organized as follows. In the following section, it is argued that merit wants are not an empty box. In Sect. 3, three levels of merit wants discussions are sketched, suggesting that each of the levels connotes a crucial question. Those questions (and the answers to them) may be seen as a first step to make visible the way in which alternative approaches to answering those questions are interrelated across various levels of analysis. Using that three-level scheme as interpretive background, Sects. 4 and 5 provide a selective overview of the pertinent literature. Section 4 deals with the pre-history of merit wants. The motivation for that rough and cursory overview is to show how old and pervasive the awareness of problems underlying merit wants conceptions is, and how various thinkers struggled to come to grips with its problems and to reconcile them with individualist frameworks. Section 5 offers some brief remarks on the history of merit wants. This supports the view that the merit wants concept may pose problems not so much on the grounds of lack of normative substance, but rather because it is complexly intertwined with "too many" potentially relevant ideas. The final section expands on some issues primarily related to collective choice and to identity/higher order preferences, complemented by some suggestions of how the different levels of merit wants may be interrelated. By way of conclusion, I suggest a tentative answer to the question of why merit wants may be a useful concept despite all difficulties.

2 Why Merit Wants Are Not an Empty Box

I am going to argue that merit wants (or at least the underlying problems) should be taken seriously. Nonetheless, three difficulties with the "traditional" merit want literature ought to be taken into explicit account. Those difficulties may have contributed to the perception that merit wants are an "empty box" and to the fact that the pertinent literature was limited to a "steady trickle", as John Head [17] put it.

First, in the literature we can find interesting and sometimes inspiring suggestions of how to deal with *various aspects of merit wants* (including various sorts of higher order preferences and community preferences, or the distinction between consumer/voter/reflective sovereignty suggested by Brennan and Lomasky [10]), but there is no unified theoretical foundation of merit wants which could serve as a focal point integrating pertinent discussions. Second, in the first decades after Musgrave [24], pertinent conceptualizations seemed difficult to integrate within the research programs of empirical economics. In certain respects, this obstacle is now overcome due to the progress made in behavioral economics. Third, merit wants were difficult to integrate in the framework of individualist welfare economics, a

problem that now may be alleviated by the progress of non-welfarist conceptions in normative economics.

Musgrave [24, p. 341] was well aware of those difficulties. In his first paper introducing merit wants, he suggested that an important aspect of merit wants could be accommodated in his Distribution Branch. Beyond that, he argues, this concept “introduces a new feature so far there is no place in our theoretical framework”. Despite various attempts of integration by Musgrave and others, the ambiguity concerning the question whether and how this can be done constrained the usefulness of the concept throughout the following decades.

Authors such as Charles McLure [20] concluded that merit wants are “a normatively empty box”, as the pertinent concepts either can be re-formulated in terms of individualistic market failure theory or else invoke unacceptable authoritarian or collectivistic premises. In the present paper, it is argued that merit wants neither are a normatively empty box, nor should they be considered as a concept of “essentially residual nature” [15], a view which is of course invited by the often-used “definition”, according to which merit wants apply to all cases where consumer sovereignty is a problematic assumption.

The real question is not so much whether that box has any specific content at all, but whether it is packed with too many different approaches, too many important and difficult issues, and perhaps too many levels of analysis. Yet in order to reject the empty box-interpretation, two kinds of arguments have to be dealt with. The first concerns the possible reconstruction of certain apparent cases of merit wants in terms of standard theory. The second is related to the claim that merit wants (insofar they cannot be reconstructed along the lines of standard theory) are at odds with standard conceptions of individualism, as they rely on heavily collectivistic views of society.

Let us start with the second argument. Do merit wants indeed rely on collectivistic views incompatible with the kind of individualism which is a foundational part of normative and positive economics? The answer is no: at least some of the approaches making use of merit arguments need to be taken seriously from a “broadly” individualist perspective. Put another way, unless we rule “broad” conceptions of normative individualism out of court from the very beginning and stick to narrow ones, we have good reasons to be interested in merit wants. What I call a “broad individualism” simply amounts to rejecting the equation of individualism with claims according to which (in cases not covered by traditional market failure¹) market-mediated individual choices are the only possible sources of normative authority. While “broad” individualism requires moving beyond narrow versions of consumer sovereignty, it does not entail philosophically more demanding conceptions of autonomy or positive freedom (cf. Berlin [8] for a crisp account of problems implied by the latter) or a rejection of welfarism. Notice in particular that “broad” individualism does not imply the kind of “broad” rationality (a notion suggested by

¹For an even narrower version of subjectivist individualism endorsed by Libertarian/Austrian Economics, the qualification added in parenthesis must be dropped (cf. e.g. [18, pp. 3–26]).

Munro [23, pp. 5–6]) which includes the theoretical option in favor of a reflective stance with regard to goals (addressing the question whether goals are reasonable in terms of “true interest”), along with instrumental rationality (viz. consistency either in choices or in preferences). While that kind of “broad” rationality no doubt prepares the ground for interesting interpretations (notably those hinging on higher order preferences and multiple selves), analyzing discrepancies between “desires” and “satisfactions” as emphasized by Pigou [27] does not presuppose a fully-fledged reflective stance with regard to goals, but may be discussed at the level “mistakes” of instrumental rationality. The distinction between desires (wantings) and satisfactions (likings) is rather common in discussions on merit wants (see e.g. Head [17, p. 232]) and libertarian paternalism. Similar issues have been discussed in-depth in the context of utilitarian ethics. Utilitarianism incorporates various axioms capturing its consequentialist, welfarist, and individualist nature. But even utilitarians who reject discriminating between different kinds of preference satisfaction (“pushpin is as good as poetry”) are not committed to the view that all individuals will (in any situation) understand their own interests best, and act accordingly.

If empirical evidence provided by behavioral economics shows some systematic violation of axioms capturing instrumental rationality, there is a *prima facie* reason to take merit policies into consideration. The case for such policies is made stronger if a definite cause triggering the violation can be identified and if the concomitant distortion can be addressed by certain policies. It is by no means clear that “broad rationality” needs to be invoked for this kind of argument. A less demanding notion of “broad individualism” may be sufficient, hinging only on the assumptions required for the diagnosis of “irrationality” (perhaps including implications for personal identity regarding the continuity of selves, but not necessarily a reflective self; see Sects. 3 and 6). Neither broad individualism nor broad rationality implies the imposition of collectivist values.²

I now move on to the other argument supporting the empty box view, claiming that the “interesting cases” of merit wants are generally reducible to cases of standard market failure. To be sure, there are some enlightening attempts showing that various apparent cases of merit wants are well captured by meaningful extensions of public goods or externalities. Head [17, pp. 240–7] includes an overview of such attempts: in one type of cases, shortcomings in individual decision making can be explained in terms of lack of information; lack of information (which can be assumed to be a public good) may be caused by its undersupply in the private-provision equilibrium. Hence public information policies (already stressed by Mill [21, V.xi]) are a remedy justified by standard public good-arguments.

²Broad rationality including reflective preferences can be taken into consideration in a manner which *takes individualist concerns seriously* and is combined with a clearly critical (but not a priori dismissive) stance regarding the normative status of reflective preferences. An example in case is the discussion of reflective preferences, p-preferences and m-preferences by Brennan and Lomasky [10], located in a merit wants context.

(Such a kind of argument could be probably extended to cases where framing effects are shown to distort individual choices: establishing a less distortive frame could be construed as a public good.) Other reconstructions include “psychic” externalities³ or preference externalities, which inter alia were used to explain merit aspects of redistributive policies. Furthermore, one may argue that merit wants associated with what Musgrave (e.g. [25]) calls community preferences are often related to certain historically emerging solutions of public good problems (cf. [17, p. 246]). For instance, national defense is a public good in abstracto, but particular military arrangements and expenses can perhaps only be explained in terms of community preferences, reflecting the values, traditions and culture of some particular community.

To sum up: some kinds of prima facie merit-based public policies (providing useful information or better “framing”) can be explained or justified as giving effect to the first-order preferences of individuals, who in absence of those policies are caught in a social dilemma. Institutional remedies suitable to overcome social dilemma situations (such as public good problems) often play an important role in policies designed to address merit wants-problems.

Successful exercises in clarification of that kind certainly are most welcome. Merit wants were often invoked in an ad hoc fashion, or in ways offering no resources to take adequate care of elitist-authoritarian implications. While the range of implications of the above-sketched reconstructions for the status of merit wants is far from obvious, they are in any case preferable to ad hoc-reasoning. But even if meaningful reconstructions are possible, some of the examples seem to suggest that merit wants-reasoning still has a role to play. Take the case of distribution as an example. As Musgrave (e.g. [25]) aptly observes, voluntary giving as well as redistribution through a voting procedure sometimes takes a specific “paternalistic” form, the donors favoring in-kind transfers instead of cash transfers. Now it is certainly possible to address this issue in terms of preference externalities. (I derive utility from seeing the recipients obtaining in-kind transfers such as educational vouchers, while I derive zero or even negative utility from seeing them getting money, which *they* would prefer to the in-kind transfers, and which I believe they are likely to spend on alcohol.) While this kind of reasoning may *explain* certain patterns and modes of redistribution,⁴ important *normative* questions remain open. Should we accept whatever pattern emerges as a case of “Pareto optimal

³Andel [1, p. 635] credits Musgrave’s graduate student Charles Tiebout for having been involved in suggesting an analysis of merit goods based on “psychic” externalities [40, p. 414].

⁴The contingent empirical fact that many people *factually endorse* an in-kind transfer policy similar to one which is, say, inspired by Rawls’s [28] conception of basic goods or Sen’s (e.g. [33, pp. 86–89]) capabilities and functionings could be taken as evidence that pertinent evaluative standards are supported by actual moral sentiments. That kind of support may be considered as important in the context of a theory of “justice as the first virtue of social institutions”, as Rawls [28, p. 3] puts it.

redistribution”⁵? If yes, does this kind of consumer sovereignty also apply to other preference externalities, which came to be known as *nosy preferences* with regard to other people’s way of life? John Stuart Mill [21, V.xi] categorically rejects policies which would give effect to such nosy preferences, even though (in the same section) he strongly advocates merit policies in the sphere of education. His pertinent arguments include the following: “Those who most need to be made wiser and better, usually desire it least, and if they desired it, would be incapable of finding the way to it by their own lights” [21, p. 868]. It seems obvious that making a distinction between mere nosiness (which carries with it, as Mill recognized, destructive potentials for a liberal society) and defensible “merit externalities” requires a step beyond the framework of consumer sovereignty, as it cannot be reconstructed in the standard externality framework. The involved issues cannot be meaningfully discussed without reference to what I call the second and third level of merit wants, including issues of higher order preferences and social choice. It may turn out that in-kind transfer policies are *justified* under certain assumptions which need to be discussed at those levels of argument.

3 Three Levels of Merit Wants Analysis

In the previous section, two kinds of arguments were scrutinized which support the position that merit wants are an empty box. It was argued that merit wants are not an empty box under the premise of broadly individualist frameworks of analysis and evaluation. This remains true if we take into account the explanatory potential of standard market failure theory in order to shed light on various cases of purported merit wants. Put another way, merit wants are not a normatively empty box, given that our starting point is the comparably cautious broadening of individualism described above, even if we take on board possibilities of reconstructing parts of their presumed domain in terms of public goods and externalities.

The concept of merit wants may not be an empty box, but is it useful? It appears as a catch-all conception summarizing aspects of human agency which cannot be addressed in frameworks confined to exogenous first-order preferences or choices revealed in a market context. How should we deal with the embarrass de richesses which seems to be on offer when considering the pertinent literature aiming at some sort of individualistic foundation for merit goods? The fragmented discussions include various aspects of

- robust and systematic “mistakes” in the choice behavior of many individuals,
- custodial choices,
- frame/context dependent and endogenous preferences,

⁵See Andel [1, p. 636] for a useful summary of the turn of the merit wants discussion towards the issue of so-called Pareto optimal redistribution in the late 1960s and early 1970s.

- “psychic” externalities, other-regarding and expressive preferences,
- concepts of multiple selves and higher order preferences,
- communal preferences and social choice,
- ethical values and conceptions of well-being and distributive justice invoking specific goods.

I suggest dividing the various discussions associated to merit wants into three broad categories:

1. “Limits of reason: This level primarily concerns the empirical identification of the limits of consumer sovereignty occasioned either by bounded rationality or by decision heuristics which are inadequate with regard to a given choice context;
2. “Higher order preferences”: This includes normative issues, dealing with the quest for an adequate framework locating normative authority when “*broad*” rationality is taken into consideration, including a reflective stance regarding goals;
3. “Collective choice”: The social choice level, making explicit the ways in which choices giving effect to merit wants are related to the dimension of social choice, including distinctions such as the one between tastes and values in Arrowian Social Choice.

As indicated above, each of the three levels connotes at least one issue of major importance:

1. “Classes of people vs. types of choice situations”.
 2. “Unavoidable paternalism vs. higher order preferences”.
 3. “Political choice as if it were market choice vs. social choice”.
1. *Classes of people vs. types of choice situations.* In a social order based on enlightenment values of individualism and rationalism, the question needs to be addressed: When can we rely on people properly using their reason, and when not? A traditional answer puts the focus on classes of people: minors and people with clinical symptoms of mental disorder are subject to custodial choices and excluded from political decision making and freedom of contract. But the class-specific view of mental maturity was extended to adult women (married women in particular) and indigenous people. Both were widely (either implicitly or explicitly) regarded as unfit to use their reason in important domains of modern social life. Hence in much of the liberal era, the problems of the imperfections of human agency were dealt with in terms of diagnostic categories with clearly discriminatory connotations. By contrast, a strand of reasoning from Mill to modern behavioural economics approaches this problem in terms of “difficult” choice situations where humans tend to go wrong, without invoking distinctions of race, class and gender.
 2. *Unavoidable paternalism vs. higher order preferences.* Sunstein and Thaler argue that paternalism is inevitable (see D’Amico ([15, Sect. 1], for a crisp summary): real-live choice does not take place in an undistorted state of nature, but is often subject to *circumstances* that influence choices in a way that fails to be

welfare-promoting: framing effects and status-quo biases are responsible for the effects of circumstances such as default rules, anchors, and the way in which the menu of choice is presented (neither of which should play a role according to standard choice theory). Put another way, some kind of “choice architecture” is in place anyway, whether produced by purposive private action (e.g. manipulative marketing) or as a spontaneous by-product of the evolution of markets and institutions. So why not improve this choice architecture in a welfare-enhancing way using the tools of libertarian (or soft) paternalism?

I do not argue that this kind of argument (as far as it goes) is off the mark. But Sugden’s [36, p. 229] objection to the conception of normative economics implicit in the above-sketched perspective should be taken seriously: according to Sugden, the libertarian paternalist “presupposes a planner with the responsibility to collate information about individuals’ preferences and then, guided by that information, to promote the overall social good.” The core of Sugden’s objection is the “single, neutral point of view” which must be assumed to guide the design of a benevolent planner engaged in the implementation of libertarian paternalist schemes. Sugden criticizes the underlying neutrality assumption, arguing that the planner always must rely on (somehow imposed) non-neutral normative judgments when determining whether individual choices are distorted or not. The diagnosis of distortions triggers his intervention and the provision of a “superior” choice architecture. The answer to Sugden’s objection is not a watertight theory of the benevolent libertarian-paternalist planner. A more promising way of dealing with such objections is to be expected from challenging the linkage between merit wants and paternalism (whether libertarian, coercive, or otherwise) in a qualified way. Challenging the concept of paternalism in the context of the present problems will be accompanied by a richer concept of agency, including “broad” rationality as mentioned above. This can be achieved if a *credible account of multiple selves and higher order preferences* is provided—as a basis for models of self-correction and self-commitment, of dealing with one’s own imperfections, and perhaps even of character-planning and sociability.⁶

3. *Market choice vs. social choice.* In his later writings on the subject, Musgrave (e.g. [25]) considers community preferences as the core of merit wants. There are good reasons for this emphasis. A first approach bringing to the fore what is at stake here is to be found in Kenneth Arrow’s [2] distinction between “tastes” and “values” (referring to aspects of social states such as the overall distribution of income and wealth) in the context of Social Choice. In a perspective making explicit the institutional dimension, social choice conceptualized as value-based (or expressive) political choice (distinct from the taste-based choices typical for market allocations) can be considered in its relation to institutions which enable individuals to give effect to their attempts towards self-correction and self-

⁶In a programmatic article, Sen [31] discusses a whole range of issues pertinent to this level of merit want analysis, including the possible role of higher order preferences.

commitment, to their higher order preferences, or to their communal preferences. (See Brennan and Lomasky [10], for an analysis of markets and collective choice procedures as different environments with respect to the way in which the preferences of “acting” selves or “reflective” selves may find their articulation.) The controversial issue at this level is this: Can we really make a strong case for genuinely *social* choice in the above-sketched sense? The alternatives involved here can be grasped by considering positions such as the one defended by Robert Sugden. Sugden [36, p. 243] argues that the market implies “the privileging of the *acting self*—the self as buyer, seller and consumer, rather than the self as the maker of plans or as the source of reflective judgments about the well-being of the continuing person. Or more accurately, the market privileges the preferences of acting selves. Sugden defends a view which tends to identify the agent with her “acting self”. This goes along with a vindication of markets and *reduced forms* of social choice which mimic markets in that they “privilege” the tastes of the “acting selves”.

4 On the Pre-history of Merit Wants: Aristotle, Locke, John Stuart Mill

Merit wants more often than not were invoked in an ad-hoc fashion, but not all discussions of merit wants lack theoretical foundations. Moreover, some of those theoretical foundations make it clear that “paternalism” is a rather simplistic and perhaps misleading way to express the problems of the merit wants-concept. Summarizing those problems by means of notions like paternalism or collectivism appears even less convincing when we consider the pre-history of merit wants and related developments which are sometimes (but not always) explicitly linked to merit wants.

Merit wants are based on claims that individual tastes are insufficient as a basis of evaluation and/or explanation in certain classes of cases. The pre-history of merit wants indicates the enduring relevance of some of the pertinent concerns. It moreover provides some illustrations of how those concerns are expressed in a variety of theoretical frameworks.

Aristotle is remarkable for the comprehensive treatment of almost the whole range of issues relevant in the context of merit wants. First, according to Aristotle, pursuing the appropriate course of action is executively difficult: Aristotle devotes considerable attention to weakness of will (*akrasia*). Second, it is epistemically difficult: to determine the right course of action requires knowledge and judgmental powers presupposing complex learning processes. Those learning processes include practice and experience, culminating in the development of virtues which are conducive to true happiness. Third, the communal dimension looms large, articulated by the conception of humans as political animals; political life has an important place in Aristotle’s universe of values. All in all, a view of human agency with its

certain imperfections and possible achievements emerges which seems congenial to the concerns coming to the fore in the discussions on merit wants. Aristotle's view of agency and sociability indeed was a source of inspiration for modern currents of thought for which ideas coming close to merit wants are of key importance: think of some versions of communitarian thought, of Martha Nussbaum's capabilities approach, or of Carl Menger, whose integration of communal wants in a staunchly individualist approach may be related to his Aristotelian background. A further aspect is perhaps even more important for modern discussions: the philosophical critique of Aristotelian virtue ethics (taking issue with (a) the conception of virtue connoting a kind of epistemological privilege and (b) the related universe of values which implies a thick concept of the good) may be relevant for establishing a defensible concept of merit wants in a society characterized by cultural pluralism, where thick concepts of the good are problematic. This critique hence exhibits some of the possible pitfalls of merit wants policies. It provides some background to the refinements of some common lines of critique focusing the "paternalistic" implications of merit wants: defensible merit want-policies must not be vulnerable to the objection that they end up with the attempt of one sub-cultural group educating the other(s).

The development of liberal thought from John Locke to John Stuart Mill should not be understood as an attenuation of individualism, but as a progressive shift of focus. Let us begin with Locke [19]. After emphasizing that only Adam was created with full powers of agency, and that it is difficult to determine whether actual persons are sufficiently endowed with agency-related powers, Locke's [19, §52–61] considerations are oriented towards the question, Which classes of people can and which cannot be relied upon as capable of properly using their reason? Locke's focus is on minors and people with clinical symptoms of mental disorder. As pointed out above, in much of the liberal era, the problems of the imperfections of human agency were dealt with in a discriminatory fashion, excluding other classes of individuals as well.

Adam Smith's [34] theory of human agency presents the whole range of issues underlying merit wants in a specific way: from systematic biases such as overconfidence causing mistakes of instrumental rationality to the complex learning processes, including the virtues of the statesman as discussed in a part added to the 6th edition of the *Theory of Moral Sentiments* (1790, VI), where a non-technocratic idea of leadership in a polycentric modern society is suggested.

For sake of brevity, I skip Hegel in my historical sketch. Hegel's importance in the context of merit goods comes to the fore in W. Ver Eecke's [42, pp. 63–65] discussion of merit characteristics of various dimensions of social policy.

If we take Locke's briefly sketched discussion of agency-related powers as point of reference, Mill's [21] progressive shift introduces a new focus. While in some passages the older tendency to identify classes of agents with insufficient agency-related powers is shining through, Mill's innovative focus clearly is on types of situations (in which eventually any individual tends to act in ways which are not conducive to her welfare). Hence it is not without reason that Mill is one of the favorite references of authors dealing with merit wants (cf. e.g. [6] or [1]). The tools

of modern Behavioral Economics are suitable to pursue this agenda and to transform it into systematic research strategies.

Mill's [21, V.xi] *Principles* provide a subtle discussion of the limits of *laissez faire*. After mentioning protection against force and fraud as core functions of the state, Mill considers the regulation of external effects. Interestingly, he includes effects based on "the moral influence of example" in this discussion, remarking that such psychic externalities are not a proper foundation for public intervention: the prude who suffers during sleepless nights as he cannot get the images of the corruptive lifestyle of the lewd out of his mind, cannot hope for public regulation, unless that lifestyle causes tangible externalities as well (see the discussion on the tension between libertarianism⁷ and the Pareto principle discussed by Sen [32, pp. 285–326]).

Protection against force, fraud and regulation of tangible externalities are seen as part of the basic rules of a System of Natural Liberty (to use Adam Smith's phrase). Apart from that, Mill discusses five further types of situations where some kind of public intervention may be justified. The onus of argument always falls on those who are in favor of the intervention, as Mill emphasizes in a famous passage. Interestingly, Mill first discusses some types of situations which are associated with "overruling" the judgment of individuals who, for various reasons, cannot be assumed to be the best judges of their interests (i.e. merit wants cases). Undeveloped judgmental powers, lack of opportunity to learn from experience and choices entailing far-reaching or/and irreversible commitments are cases in point discussed by Mill. Along those lines, Mill [22] also argues in favor of limits to contract freedom in the case of slave contracts, while undeveloped judgmental powers are invoked in the case of educational choices by Mill [21, p. 868]: "Those who most need to be made wiser and better, usually desire it least, and if they desired it, would be incapable of finding the way to it by their own lights."

The fourth type of situation, as Mill explicitly states, is associated with giving effect to the actual preferences of individuals: Mill says that the task of public governance in that class of cases is "not to overrule the judgment of individuals respecting their own interest, but to give effect to that judgment; they being unable to give effect to it except by concert, which concert again cannot be effectual until it receives validity and sanction from the law." This clearly is a specific formulation of the social dilemma structure that may emerge in cases of public good provision. As a fifth class, Mill discusses other-regarding choices in a distributive context together with aspects of empowerment, including policies with first-order effects on future generations. The latter are related to merit wants in a particularly complex way, insofar we are dealing with the preferences and opportunities of individuals not yet born.

The programmatic association of behavioral economics with libertarian paternalism (including the design of a tool-box of instruments allowing for better targeting of policies) is not merely a terminological provocation whose oxymoronic flavor provides genuine food for thought. It offers the perspective of policies that are better

⁷The status of minimal libertarianism is analyzed by Baigent [3].

targeted to the diagnosed problems as compared to traditional patterns of public intervention. “Better targeted” should not be understood in a merely pragmatic or technocratic sense. It refers to improved diagnosis of classes of choice situations in which merit policies are commendable. Declaring certain classes of individuals as unfit to choose on their own lights (motivating class-specific merit policies) is rendered obsolete to some extent. In this sense, the association of behavioral economics and libertarian paternalism is a program very much in the spirit of Mill.

Of all currents of thought in which issues related to merit wants played a role, the German language public finance literature is the most direct source of inspiration for Musgrave’s conceptual development. Musgrave became conversant with enlightened versions of this tradition during his studies in Heidelberg, and was reminded of it by the contributions of another German-Jewish emigrant to the U.S., Gerhard Colm [13, 14], economic advisor to President Truman. In the German language public finance literature (which included Austrian and Swedish authors) a conceptual differentiation emerged in the course of the nineteenth and early twentieth century: disambiguating what was originally summarized by the notion of “collective needs” in two different classes of phenomena: (1) individual needs or wants which are accommodated by goods which are in different degrees non-rival and non-excludable, preparing the ground for Wicksell’s and Lindahl’s pertinent contributions and the literature on public goods; and (2) needs or wants which presuppose the existence of some form of community, even though they are felt by individuals (communal wants). A doctoral dissertation by Margit Cassel [12] comes close to a complete overview of the whole range of problems and pertinent concepts. Before Cassel [12], a discussion including Austrian, Italian and Scandinavian authors (notably Sax, Mazzola, Wicksell, and Margit Cassel’s father Gustav) had successfully transformed the concepts inherited from German public finance within a marginal utility framework, including the terminological transformation eventually leading to the notion of “public goods” instead of social wants.

By way of conclusion of Sect. 4, a remark on Charles Taylor’s [37, pp. 127–45] concept of “irreducibly social goods” is in order. Taylor’s discussion is helpful in getting clear about the differences between public goods and the kind of merit wants which are directly related to a genuinely communal level. Taylor’s arguments can be summarized as follows: Merit wants of this kind are accommodated by irreducibly public goods, i.e. by goods or services for which there are no private substitutes. The goods considered here are categorically lacking private substitutes (i.e. private substitutes are not merely unattractive for empirically contingent reasons, such as currently available technologies and relative prices). A fair distribution may be seen as an irreducibly public good in that sense, but further kinds of goods may correspond to that definition: goods which are related to goals/wants that make no sense in absence of a pre-existing community with its culture and its values. Think for instance of the goals guiding Britain’s defense effort during the Battle of Britain, which can be addressed as national goals (this terminology is used by Colm [13, 14]). Hence there appears to exist a close connection between communal preferences (as discussed in the German Public Finance tradition) and Taylor’s irreducibly social goods.

5 A Brief Retrospective on Merit Wants

Since Richard Musgrave coined the notion of merit wants⁸ in 1956, a modest, but steady trickle of articles explicitly dealt with their conceptual clarification. Musgrave's pertinent contributions throughout the following decades show some shifts of emphasis (recapitulated by [1]). A succinct summary of conceptual issues is to be found in Musgrave [25].

All in all, in the more than 50 years after the introduction of merit wants by Richard Musgrave, its development exhibited considerable vicissitudes. It was not an unambiguous success story. Musgrave's original problem was posed at the explanatory level: a considerable share of public sector activity is difficult to explain solely on the basis of standard micro-based market failure theory. Initially, Musgrave's focus was public provision of rival and excludable goods and services: merit wants were seen as a possibility to explain the provision of such goods. Later he made clear that the merit dimension is independent of the degree of rivalry and excludability. More generally, the way in which merit wants were located in his overall theoretical framework of market failure seems to have changed over time (cf. [1]). In particular, his emphasis shifted to the issue of "leadership in a democratic society" and communal wants as the core of merit wants (see [25]), a perspective which is also pursued in a number of articles in the 1990s where his roots in the German language Public Finance tradition are better visible than in earlier work.

While Musgrave's discussions of merit wants are full of perceptive remarks throughout the various shifts of emphasis, he developed no unified theory of merit wants, let alone a coherent integration within the modern micro-based theory of the public sector and modern normative economics. Despite various attempts to define them in a more specific way, the shifts of emphasis may have contributed to the view (explicit or implicit) that they are best understood as a residual category, covering all the cases where consumer sovereignty fails to be a convincing concept either as basis of normative authority or of explanation. Such a residual view provides few resources against invoking merit want-arguments in an ad-hoc fashion. Hence as a potential guide to matters of public policy, it is notoriously associated with the dangers of elitist/authoritarian imposition of values, insofar there is no justifiable answer to questions such as: Whose values ought to matter? Which values ought to matter, and how? Having in mind suchlike objections, Musgrave himself was rather cautious regarding the scope and usefulness of the merit wants-concept.

⁸A referee suggested to comment on the terminological ambiguity of merit wants vs. merit goods. In keeping with the German-language Public Finance tradition, in his early contributions Musgrave used the notion of merit wants as well as social wants (for public goods). As he was eager to get rid of terminological heritage which might hinder the development of a unified micro-based theory of market failure, he soon adopted the now common terminology. This terminological shift included merit goods, even though the latter are mostly explained in terms of properties of individual values or preferences, whereas rivalry and excludability are properties of goods rather than of wants.

Yet there exists respectable work which contributed to the conceptual clarification of merit wants. In parts of it, the roots of the merit wants concepts (owing much to German Public Finance and Gerhard Colm in particular) is made explicit. Geoff Brennan and Loren Lomasky [10] were among the first who contributed to the conceptual development of the communal wants dimension of merit wants beyond Colm [13, 14]. They explicitly referred to Colm's distinction between homo oeconomicus and homo politicus (and Colm's corresponding distinction between market and politico-fiscal processes), developing those ideas in the context of a democratic institutional perspective with an explanatory agenda. Nick Baigent's [4] attempt to integrate merit wants within Arrowian Social Choice theory can be seen as the counterpart in the sphere of normative economics. Colm emphasized the conceptualization of the judgments of homo politicus as referring to overarching national goals (e.g. defense or education policies) in a specific way. Arrow's individual "values" (complementing the tastes of consumer theory) refer to aspects of social states other than individually consumed quantities of goods. While the analytical difference between the two kinds of distinction would require some further discussion, their main thrust is somewhat similar.

Brennan and Lomasky [10] introduced distinctions such as the one between m-preferences (typically related to outcomes in a market environment), reflective preferences (which are not immediately outcome-oriented in the way m-preferences are) and p-preferences, which are "intimately related" to reflective preferences as political choice processes typically allow for the expression of preferences with small expected costs. Kaushik Basu's [6] distinction between actual choice and retrospective choice is based on a different dimension of multiple selves, namely between present and future selves. This distinction is important in particular in combination with circumstances such as choices related to learning processes and important irreversible consequences, discussed already by Mill [21, 22]. Mill is also a source of inspiration in Basu's [7] more recent writings on the limits of the "principle of free contract".

John Head [17] provides the most comprehensive survey on the literature dealing with conceptual issues.⁹ This survey does not yet reflect the perspectives of reinvigoration of that concept made available by advances in behavioral economics. The latter are explicitly combined with conceptual work on merit wants in Alistair Munro's monograph [23] entitled "Bounded Rationality and Public Policy."

A different level of merit goods is emphasized by W. Ver Eecke [41, 42], whose discussion is framed by a constitutional perspective based on a voluntary exchange conception of public goods provision. According to Ver Eecke, the specific difference of merit goods is that they cannot be based on voluntary exchange, as their provision implies losers. Ver Eecke starts with a rather widely shared justification of markets as social arrangements giving effect to individual preferences. Merit

⁹Other authors using the concept of merit wants were not so much interested in conceptual issues but in implications not least for the economics of taxation. See e.g. Pazner [26], Besley [9], Schroyen [29], Capéau and Ooghe [11].

wants in turn are justified as a specific class of possibility conditions either (along “Kantian” lines) for the functioning of market economies or (along “Hegelian” lines) for preserving a sufficient degree of freedom for all individuals in a market society, notably by means of social policies in the face of economic hardship. Both kinds of conditions are characterized by the impossibility of a policy design which would render them beneficial for everybody. The implementation of pertinent policies is accompanied by losses of some individuals.

The issue of distribution has ever been present in Musgrave’s discussions on merit wants. In Ver Eecke’s writings, it is put to the center of stage. His idea of distribution-sensitive “possibility conditions” is interesting. Yet two questions need further discussion in order to prevent overstressing the concept: (1) How does all this relate to the well-known limitations of a Wicksell-Lindahl voluntary exchange approach which can be discussed in a conventional public good-framework? (2) How does it relate to analyses of strategic structures representing social dilemma situations in which certain socially desirable “solutions” are available, but are not symmetric in terms of individual advantages?

6 Conclusion: Towards a Conceptually Complete Theory of Merit Wants?

Taking stock of some contributions to the “steady trickle”, approaches seeking to provide theoretical clarification fall into three distinctive classes: (1) the diagnostic level of choice frames, contexts, heuristics and bounded rationality, (2) various versions of higher order preferences, and (3) various ways of associating merit wants with the logic of collective choice.

An integrated (conceptually complete) theory of merit wants across the three levels is complex and difficult. But is it attractive as a research agenda? To be sure, those three levels were present from the beginning in the writings of Richard Musgrave and especially Gerhard Colm [13, 14], who was an important source of inspiration for Musgrave. Taken together, Colm’s and Musgrave’s pertinent writings suggest that the analysis of institutionalized mechanisms (presumably) giving effect to merit wants conceptually and practically will involve all three levels. But this is not more than a hint that a conceptually complete theory of merit wants could be a plausible research agenda. Hence by way of conclusion, I try to provide some additional hints regarding the attractiveness of such an agenda.

The present revival of merit wants is primarily triggered by developments in behavioral economics. They by and large entail a focus on level (1), accompanied by some discussions at the level of (2). Munro [23, p. 6] draws a map of merit wants-aspects composed of four regions related to various diagnoses of limits of reason; i.e., the map *prima facie* refers almost exclusively to (1):

- (i) Defective telescopic faculties;
- (ii) Defective information processing;

- (iii) Frame dependent preferences; and
- (iv) Preference misalignment.

Given the progress of behavioral economics, (i), (ii) and (iii) may seem to be fairly well-explored territory. Preference misalignment (iv) is located in the core area of Munro's map and can be seen as still partly uncharted territory. In Munro's exposition, (iv) seems to be the residual region of merit wants; it is in keeping with the residual character which is sometimes attributed to merit wants as a whole.

Behavioral economics brought about unambiguous progress on the diagnostic level; there can be little doubt that its tools and findings will keep a firm place in the economist's toolbox. Yet upon closer inspection some loose ends become apparent which invite further thought. Here are three examples illustrating that claim. First, the problems related to (i) ("defective telescopic faculties") may raise the question of whether or not issues of a temporal sequence of selves are involved, where later selves may regret some irreversible decisions of the former selves (e.g. to forego educational opportunities, preferring the pleasures of consumption). Basu [6] reconstructs merit wants along those lines. Corresponding merit want policies may be understood in terms of giving effect to the preferences of later selves. This sounds plausible, but there is a problem: according to such policies, the preferences of later selves eventually are made to trump those of the earlier selves. (Why) Can this be legitimate? While there are arguments supporting public policies acting as advocates of later selves (Basu aptly draws on Mill [21, V.xi]), the issue is not so simply resolved. Mill's arguments include irreversibility and systematic lack of experience-based imagination of some aspects of future states. Under certain circumstances (indicated by Mill), the present self cannot be expected to be an effective advocate of the later self.

Notice though that such arguments favor some conceptions of personal identity and exclude others. Consider, for instance, the foundational discussion on "libertarian paternalism" as a program motivated by certain cases of limits of reason. Robert Sugden's (e.g. [35, 36]) suggestion of an interpretation of consumer sovereignty which does not hinge upon coherent preferences, along with the related idea of "privileging the acting self" mentioned in the above, not only amounts to a rejection of libertarian paternalism, but also of other kinds of merit wants policies motivated by the arguments just sketched.

It can be argued that this provides a motivation for a conceptually complete account of merit wants as a research agenda. Sugden's explicit critique of libertarian paternalism (and the implicit critique of a larger class of merit policies) entails a specific conception of personal identity as a sequence of time-slice selves (for a summary, see D'Amico [15, Sect. 2]). At a collective choice-level it entails a specific conception of contractarianism treating political choice *as if it were market choice*. Hence considering discussions on merit wants which first seem to be confined to issues of bounded rationality and the corresponding toolbox of libertarian paternalism at level (1), a significant divide pertinent to levels (2) and (3) becomes visible. Considering those discussions (for an overview see [15]), it becomes obvious that any kind of systematic suggestion regarding the evaluative

implications of behavioral findings (such as Sugden's interpretation of consumer sovereignty which does not rely on coherent preferences) explicitly or implicitly hinges on certain answers to the crucial questions to be posed at levels 2 and 3. Ultimately, the divide concerns the status of collectives and collective choice, and it is related to different conceptions of personal identity and agency. The nature and the implications of the divide cannot be properly addressed when the third level, the level of social choice is not taken into explicit consideration. Hence we have to move to level 2 and to level 3 for a proper discussion of the underlying problems, unless we subscribe to a kind of naïve technocratic conception of policy which is difficult to defend, given a pluralistic and polycentric society.

Integrating levels 2 and 3 also seem to make sense in cases such as nosy preferences. Behavioral economists now do experiments where nosy preferences play a role: people deviate from what they are expected to do in the standard model, for reasons that could be circumscribed with "nosy preferences" (see [30] for a brief summary of findings). Yet the normative implications of such findings are unclear, motivating a more encompassing theorization including levels 2 and 3. Even if "nosy preferences" can be shown to be efficiency-enhancing in some cases (e.g. when people punish free-riders) and to inefficient situations in others, the normative status of the "nosy preferences" as such needs to be clarified, as well as the procedural legitimacy of mechanisms either giving effect to those nosy preferences or laundering them.

Last but not least, the scope of frame dependence may be sufficiently wide as to make integration of the three levels sometimes indispensable: the market and the political forum may be regarded as two different decision contexts—and as two different behavioral frames. In an essay entitled "The market and the forum", Jon Elster [16] explains some possible backgrounds of such contextual differences (but cf. also the paper by Brennan and Lomasky [10], quoted above). Communal wants with their cultural connotations invoke a specifically far-reaching form of frame-dependence, insofar they are presupposing the existence of the specific frame of a given community (cf. Cassel [12, §§16–22]).

To sum up: Important progress at the diagnostic level notwithstanding, and acknowledging the multifaceted aspects of pertinent diagnosis of the limits of reason (see Munro [23, pp. 5–6]), discussions concerned with the integration of recent behavioral findings within normative economics indicate that steering clear of levels 2 and 3 leaves open crucial questions regarding the status of merit wants and related concepts. A conceptually complete theory of merit wants encompassing the three levels is no doubt complex. But it is an agenda which is foreshadowed in pertinent work (particularly by Musgrave and Colm), and which could add considerable leverage to the advances made in behavioral economics.

In the years after Musgrave [24], merit wants served as a reminder that public goods and externalities may not accommodate all explanatory challenges with regard to non-market goods and public interference. Colm and Musgrave framed the problem more or less in this way. Moreover, it served as a reminder that the concepts of New Welfare Economics may not solve all evaluative problems.

Research programs such as behavioural economics and non-welfarist Social Choice have changed the situation. Merit wants could now serve as a heuristic device, or as a focal concept encompassing and combining the three above-sketches levels of analysis. If it is at all useful today, it is not useful as a minor chapel off the main naves of public economics, normative economics and Social Choice. It is not useful as a concept suitable to discuss a residuum of some abstruse cases. If it has a function, its scope is more encompassing and foundational.

Acknowledgements The paper has benefitted from comments by two anonymous referees and Maxime Desmarais-Tremblay.

References

1. Andel N (1984) Zum Konzept der meritorischen Güter. *Finanzarchiv* 42:630–648
2. Arrow K (1963) *Social choice and individual values*, 2nd edn. Yale University Press, New Haven
3. Baigent N (1981) Decompositions of minimal libertarianism. *Econ Lett* 7:29–32
4. Baigent N (1981) Social choice and merit goods. *Econ Lett* 7:301–305
5. Baigent N (1995) Behind the veil of preference. *Jpn Econ Rev* 46:88–101
6. Basu K (1975/1976) Retrospective choice and merit goods. *Finanzarchiv* 34:220–225
7. Basu K (2011) *Beyond the invisible hand: groundwork for a new economics*. Princeton University Press, Princeton
8. Berlin I (1969) *Four essays on liberty*. Oxford University Press, Oxford
9. Besley T (1988) A simple model for merit good arguments. *J Public Econ* 35:371–384
10. Brennan G, Lomasky L (1983) Institutional aspects of ‘merit goods’ analysis. *Finanzarchiv* 41:183–206
11. Capéau B, Ooghe E (2003) Merit goods and phantom agents. *Econ Bull* 8:1–5
12. Cassel M (1924) *Die Gemeinwirtschaft oder Die Gründe einer öffentlichen Haushaltung*. Inaugural-Dissertation. A. Deichertsche Verlagsbuchhandlung Dr. Werner Scholl, Leipzig
13. Colm G (1955) *Essays in public finance and fiscal policy*. Oxford University Press, New York
14. Colm G (1965) National goals analysis and marginal utility economics. Some non-technical comments on a highly technical topic. *Finanzarchiv* 24:209–224
15. D’Amico D (2009) Merit goods, paternalism, and responsibility. Mimeo, University of Pavia, Pavia
16. Elster J (1989) The market and the forum. In: Elster J, Hylland A (eds) *Foundations of social choice theory*. Cambridge University Press, Cambridge, pp 103–132
17. Head J (1991) Merit wants: analysis and taxonomy. In: Eden L (ed) *Retrospectives on public finance*. Duke University Press, Durham, pp 229–252
18. Hoppe H-H (1993) *The economics and ethics of private property*. Kluwer, Boston
19. Locke J (1690) *Two treatises of government*. Awnsham Churchill, London
20. McLure C (1968) Merit wants: a normatively empty box. *Finanzarchiv* 27:474–483
21. Mill JS (1848) *Principles of political economy*. Longman, London
22. Mill JS (1859) *On liberty*. Longman, London
23. Munro A (2009) Bounded rationality and public policy. A perspective from behavioural economics. *The economics of non-market goods and resources*, vol 12. Springer, Heidelberg
24. Musgrave R (1956/1957) A multiple theory of budget determination. *Finanzarchiv* 17:333–343
25. Musgrave R (1987) Merit goods. In: Eatwell J, Milgate M, Newman P (eds) *The new palgrave: a dictionary of economics*. Macmillan, London, pp 452–453
26. Pazner EA (1972) Merit wants and the theory of taxation. *Public Finance* 27:460–472

27. Pigou AC (1932) *Economics of welfare*, 4th edn. Macmillan, London
28. Rawls J (1971) *A theory of justice*. Harvard University Press, Cambridge
29. Schroyen F (2003) An alternative way to model merit good arguments. Paper, Department of Economics, Norwegian School of Economics & Business Administration, Bergen
30. Schulz JF, Thoeni C (2013) Paternalismus, Rationalität, systematische Fehler, nudges: Befunde der experimentellen Ökonomik. *Jahrbuch für normative institutionelle Grundfragen der Ökonomik* 12:63–82
31. Sen A (1977) Rational fools: a critique of the behavioural foundations of economic theory. *Philos Public Aff* 6:317–344
32. Sen A (1982) *Choice, welfare and measurement*. Blackwell, Oxford
33. Sen A (2002) *Rationality and freedom*. The Belknap Press of Harvard University Press, Cambridge
34. Smith A (1759, 1790, [1976]) *The theory of moral sentiments*. In: Raphael DD, Macfie AL (eds) *Glasgow edition of the works and correspondence of Adam Smith*, vol 1. Clarendon Press, Oxford
35. Sugden R (2004) The opportunity criterion. Consumer sovereignty without the assumption of coherent preferences. *Am Econ Rev* 94:1014–1033
36. Sugden R (2008) Why incoherent preferences do not justify paternalism. *Const Polit Econ* 19:226–248
37. Taylor C (1995) *Philosophical arguments*. Harvard University Press, Cambridge
38. Thaler R, Sunstein C (2003) Libertarian paternalism is not an oxymoron. *Univ Chicago Law Rev* 70:1159–1202
39. Thaler R, Sunstein C (2008) *Nudge*. Yale University Press, New Haven
40. Tiebout CM, Houston DB (1962) Metropolitan finance reconsidered. *Rev Econ Stat* 44:412–417
41. Ver Eecke W (2007) *An anthology regarding merit goods*. Purdue University Press, West Lafayette
42. Ver Eecke W (2013) *Ethical reflections on the financial crisis 2007/2008*. Springer briefs in economics. Springer, Berlin