

## *A Benthamite model of democracy and its problems*

### **Abstract**

In this chapter, I discuss a Benthamite argument for representative democracy (which closely resembles an argument by James Mill). I derive a simplified model of representative democracy from the argument and examine its merits and problems. I then suggest an improved model, or at least a number of conditions which such an improved model should satisfy. The upshot is that such an improved model must follow the principle of proportionality, which advocates that votes should be assigned in proportion to stakes.

### **1 Introduction**

This paper starts with a brief recapitulation of Jeremy Bentham's argument for representative democracy. It then suggests a simplified Benthamite model of democracy which suits his utilitarian purpose. The paper then proceeds to examine some problems with this proposal and to suggest possible solutions.

### **2 Bentham's argument recapitulated**

Bentham's main line of argument for representative democracy rests on three central claims, concerning the proper and actual ends of government, and their reconciliation. The first one we can call the *Proper End Claim*:

The right and proper end of government in every political community is the greatest happiness of all the individuals of which it is composed. Say in other words, the greatest happiness of the greatest number.<sup>1</sup>

This claim is derived from the more general *Principle of Utility*: that the end of all human action ought to be the maximisation of the sum total of happiness.<sup>2</sup> The corresponding interest for this end in individuals Bentham dubs the 'universal interest'.

Bentham's line of argument continues with what we can call the *Actual End Claim*:

The actual end of government is in every political community the greatest happiness of those, whether one or many, by whom the powers of government are exercised.<sup>3</sup>

This claim is derived from Bentham's general psychological theory of motivation according to the *Principle of Self-Preference*: that generally, the end of all human action is the

---

<sup>1</sup> Bentham (1989[1822]c:232).

<sup>2</sup> Bentham's "greatest happiness of the greatest number" claim should be interpreted to advocate the maximisation of the *sum total* of happiness (cf. Bentham 1983[1829]:309-10, Burns 2005:57-8, Goldworth 1969, Rosen 1983:201-3, Schofield 2006:28-50). The same interpretation applies to the proper end of government.

<sup>3</sup> Bentham (1989[1822]c:232).

maximisation of happiness for the individual in question.<sup>4</sup> To put it in terms of interest: generally, what motivates individual action is individual interest. Note that for Bentham happiness may be self- or other-regarding: in the latter case, it is derived from other individuals' happiness, or lack thereof (due to sympathetic or antipathetic interests), whereas in the former case, it is independent of other's happiness.

This principle does not imply that individuals necessarily or even generally are *deliberately motivated* by some such interest, i.e. that their motivation to act is caused by conscious deliberation according to the principle. Rather, it states that an action in general can be *explained* by reference to some such interest. Such an explanation ascribes motivational force to the relevant interest itself, not to conscious deliberation about it. Crucially, this descriptive principle implies that we are not to *expect* any individual to act against her individual interest. Thus, if we should found our social engineering on an expectation of “the predominance of social regard in the breast of any public man”, Bentham claims, the “practical consequences [would be] by far the most pernicious, productive of evil in the greatest quantity”.<sup>5</sup>

At this stage of the argument, the two claims allow for a discrepancy between the two ends. Given Bentham's assumptions, it can be shown that such discrepancy is not only possible, but also highly probable.<sup>6</sup> This bestows moral urgency on the quest of finding some solution for reconciling the two ends, or the two underlying corresponding interests. The gist of Bentham's argument is what we can call his *Reconciliation Claim*:

If [...] the conduct of no man can at any moment reasonably be expected to be determined by any interest that at the same moment stands in opposition to that which in his conception is his own individual interest, [it] follows that for causing it to take any direction in which it will be subservient to the universal interest, the nature of man, the nature of the case, affords no other

---

<sup>4</sup> Bentham (1989[1822]c:233).

<sup>5</sup> Bentham (1989[1822]a:15). Bentham actually points out *two* erroneous opinions here: one is the assumption that interest from self-regard is not “generally predominant”; the other that interest from social regard (e.g. an interest for others' happiness) is “altogether absent”. Thus, while self-regarding interest must be generally expected, other-regarding interest cannot be ruled out. Still, says Bentham, social engineering must always proceed from suspecting the worst in people (ibid.:14-5).

<sup>6</sup> The argument can only be indicated here: it involves Bentham's assumptions that (i) direct democracy is practically impossible (and thus, the set of individuals constituting government in the real world is necessarily non-identical to the set of individuals constituting the community), (ii) generally, since we must expect it to be every individual's motivating interest to maximise her own happiness, we must suspect everyone to – as far as possible – act so as to sacrifice everyone else's happiness to her own, and (iii) being a member of a powerful group like government grants many such possibilities. Claim (i) allows for the possibility of a discrepancy between actual and proper end of government, while (ii) and (iii) account for the high probability of such discrepancy.

method than that which consists in the *bringing the particular interest of the rulers into accordance with the universal interest*.<sup>7</sup>

The upshot of Bentham's argument is that the morally necessary reconciliation can be effectively set up under a specific form of government: representative democracy. The argument is often framed in the negative, in terms of excluding other forms of government, stating the reasons why monarchy or aristocracy (or any "mixed forms" of government) could not be expected to provide the same results. However, it is possible and highly enlightening to reconstruct a positive Benthamite argument for representative democracy.<sup>8</sup>

The general idea of this argument can be stated as a chain of sufficient conditions for achieving an identification of the functionaries' individual interests with the universal interest. Both conditions rest on Bentham's theory of motivation.

Firstly, for such identification to occur, it is sufficient that certain "securities", i.e. incentives, are set up and enacted which make promoting any individual, non-universal interest less appealing for each functionary than promoting the proper, universal interest. In other words, there must be a well-known and reliable incentive structure which assigns punishment to those functionaries who act improperly. Bentham claims that this can be achieved by holding them both legally and morally responsible: i.e. they face being put out of office or legally punished on the one hand, and being condemned by the "popular or moral sanction" at the hands of the informal "tribunal of public opinion" on the other hand.<sup>9</sup>

Secondly, in order to secure that a proper incentive structure is set up and enacted, it is sufficient to entrust such a set of individuals (one or several) with this assignment whose predominant individual interest or aggregate of interests coincides with the universal interest. There is indeed one, and only one set of individuals whose predominant interests necessarily

---

<sup>7</sup> Bentham (1989[1822]c:235), italics added.

<sup>8</sup> For Bentham's negative argument, see e.g. Bentham (1989[1822]d:151-226). Given some limited existence claim, e.g. 'the three forms of government and their mixed forms exhaust the class of possible governments', it is possible to argue comparatively for Bentham's uniqueness claim that representative democracy is "the only apt" form of government (ibid.:216). The following non-comparative reconstruction of a positive argument, however, is not dependent on any such controversial limited existence claim. Still, in order to show that representative democracy is morally justified by the principle of utility, one would have to combine the positive argument with some maximum claim, e.g. 'there is no form of government generates more utility than representative democracy'.

<sup>9</sup> Bentham (1989[1822]a:56). Bentham discusses four other securities – these are not incentives but rather constraints to the representatives' set of permitted action alternatives. For a comprehensive discussion of these "securities" and their relation to the rulers' "appropriate moral aptitude", see Bentham's (1989[1822]a:28-76).

coincide with the universal interest of the community in question: the entire community itself, i.e. the people. The reason for this coincidence being necessary is that, according to Bentham, “the universal interest is nothing else but the aggregate of all individual interests”.<sup>10</sup>

The economic notion of setting up and enacting an incentive structure can be translated into political terms if one replaces it with the notion of exercising political power, or more specifically in the given context: exercising constitutive or sovereign power.<sup>11</sup> This is indeed what Bentham wants to entrust the people with: “if an equal share of the constitutive power in question is in the hands of every member of the community in question, the aggregate of the several personal interests is itself the universal interest” (1989[1822]a:96).

Demanding an equal share of power for each member of the community reflects an early idea of Bentham: Everyone’s individual interest is to be counted equally in aggregating the universal interest. Even though Bentham is aware that individuals’ amount of happiness might differ, and thus (the intensity of) their interests, he suggests that, lacking a trustworthy method of measurement, one should retreat to the next best claim, viz. assuming everyone’s interests to count as morally equal:

Chacun a un droit égal à tout le bonheur dont sa nature est susceptible. Faute de pouvoir déterminer le degré relatif de bonheur dont différens individus sont susceptibles, il faut partir de la supposition que ce degré est le même pour tous. Cette supposition, si elle n’est pas exactement vraie, approchera au moins autant de la vérité que toute autre supposition générale que l’on pourroit mettre à sa place. (2002[1788]:68.)<sup>12</sup>

Of course, this assumption is not uncontroversial (I will discuss it in another paper).

### **3 A simplified Benthamite model of democracy**

The model of democracy which emerges from Bentham’s argument can be described as a decision procedure, bridging the gap between the empirical facts of human psychology,

---

<sup>10</sup> Bentham (1989[1822]b:133).

<sup>11</sup> Cf. Bentham (1989[1822]c:242): “1. To those whose interest composes the universal interest give or leave as much power as possible. 2. To those whose interest is not the universal interest but in its very nature adverse to the universal interest give as little power as possible. 3. Keep on foot, in the character of a power prepared when occasion calls, a power superior to their own, the power of those whose interest is the universal interest, in readiness to act upon them in the character of Judges and punish them with dishonor [sic] and loss of office in their character of legislators.”

<sup>12</sup> “Everyone has an equal claim to all the happiness to which his nature is susceptible. Lacking the ability to determine the relative degree of happiness to which different individuals are susceptible, one has to assume that this degree is the same for all. This assumption, even if not exactly true, approximates the truth at least as much as any other general assumption which one could put in its place.” (My translation. I follow Schofield’s suggestion of translating “droit” with “claim” instead of “right”; cf. Schofield (2006:83n).)

according to Bentham's *Principle of Self-Preference*, and his normative *Principle of Utility*, understood as the criterion of rightness.<sup>13</sup> Bentham himself proposes a much more detailed version of such a model which cannot be accounted for in the present context.<sup>14</sup> Instead, I will suggest a simplified model of representative democracy which underlies his own complex proposal, and which is sufficient for his argument to work.

The Benthamite model is set up on two distinct levels of collective decision-making with single-winner outcomes: The first level of the model is the election of one functionary for each of the community's equally large sub-communities by this sub-community's voters. According to the *Principle of Self-Preference*, we should expect that the most predominant interest for any random voter is to maximally enhance her own happiness. Hence a candidate who promises to promote this voter's interest profile on all public policy issues will receive this voter's vote. Yet, says Bentham, "certain points common to all excepted, the interests of all [the] voters are in a state of opposition to one another".<sup>15</sup> We may take this to mean that any voter's interest profile is more or less idiosyncratic. But this means that no sufficiently large group exists to elect a candidate who promotes any such idiosyncratic profile. Yet supposedly it is the candidate's predominant interest to win the election; hence he promises to "promote the interest of each in those points in which it is not opposed by any other interests".<sup>16</sup>

Bentham's solution here is ambiguous: Should the functionary be elected unanimously or by majority? The first option has the candidate simply *disregard* any of his voters' particular interests which are in a "state of opposition", and promise to promote only those interests which are shared unanimously. The second option has the candidate promise to promote a policy profile which, on top of these shared or unopposed interests, includes a selection of

---

<sup>13</sup> See Bales (1971) on decision making procedures and criterion of rightness.

<sup>14</sup> For a short summary of the institutions of Bentham's democratic government, see Rosen (1983:130-167). Cf. Schofield (2006) for a comprehensive account of the development of Bentham's political thinking.

<sup>15</sup> Bentham (1989[1822]b:135). If public policy decisions at the second level are about resource distribution in a wide sense, such opposition can be explained in terms of everyone deriving their happiness from incompatible distribution patterns. This may e.g. be realised in a community of extreme egoists (where everyone's predominant interest is to attain a maximum of resources), but also a community of extreme altruists (where everyone's predominant interest is to provide everyone but themselves with a maximum of resources), or a community of voters who each derive their happiness from these and other (e.g. egalitarian), jointly incompatible distribution patterns. On this account, however, a community of e.g. perfect egalitarians (whose predominant interest is to provide everyone with an equal share of resources) would *not* display this opposition.

<sup>16</sup> Bentham (1989[1822]b:135).

interests which are held by some majority among his voters but which are opposed by others. We will get back to this problem below.

On the second level of the model, the elected functionaries decide on policies which, presumably, affect the entire community. On this level, one sub-community's aggregate interest, as represented by their functionary, does not necessarily match the other sub-communities' respective aggregate interests. However, each functionary can be expected to vote according to his promises, since, if the incentive structure of "securities" is set up properly, this will be the only course of action promoting his own individual interest (staying in office and paving the way to succeed even in the next election). But moreover, the functionaries' behaviour will be without "evil consequence"<sup>17</sup>: each functionary finds himself either as part of a majority of functionaries voting for the same interest, or as opposed by some other majority voting for another interest – either way, since the majority wins, the universal interest of the community is promoted, as required by the principle of utility.<sup>18</sup>

The above exposition of the model directs attention to another of Bentham's assumptions. In order to make the model work, in the sense of making it produce the right outcomes, the voters must be assumed to generally know their own interests better than any other. This is what we can call the general *Claim to Self-Knowledge*:

With a benefit of a certain degree of experience it may be delivered in the character of a general proposition [that] every man is a better judge of what is conducive to his own well-being than any other man can be.<sup>19</sup>

The assumption is not uncontroversial, but we will accept it here for the sake of the argument.

One of the crucial features of this model is the mechanistic role of the functionary. The functionary constitutes an aggregating channel for the (according to the *Claim to Self-Knowledge*) generally correct judgments of his voters; and the proper functioning of this channel is effectively secured by a combination of the predominant interest of the functionary and the predominant interests of his voters (as predicted by the *Principle of Self-Preference*). Another way to describe how this proper functioning is secured is by employing the notion of a duty, as suggested by Bentham.

---

<sup>17</sup> Bentham (1989[1822]b:135).

<sup>18</sup> Bentham (1989[1822]b:135-7).

<sup>19</sup> Bentham (1983[1814-31]:131).

First and foremost, the functionary's duty *towards his voters* requires that he votes for the aggregate of their interests, as perceived by them, in public policy decisions on the higher level. Yet, according to Bentham, the functionary also has a duty *towards the public* which requires that he should enlighten his voters about their proper interests by giving speeches to this effect.<sup>20</sup> The former duty is effective regardless of whether the voters are right or wrong about their interests. The latter duty is added since the voters occasionally may be wrong after all; i.e. the result on the lower level may be incorrect, and such incorrectness would spill over to the higher level – possibly, though not necessarily, to the detriment of the universal interest.

Despite the problems that infest the Benthamite model, as we will see below, it has some recommending features. Firstly, the model makes democracy pragmatically feasible since it requires collective decision-making about a limited number of policy profiles, as presented by each candidate, instead of inconceivable multitudes of single issue policies. Secondly, in making the social choice depend on the voters' preferences, each of which, as we assumed, correlates for each voter with the preferred outcomes' utility, it is designed to promote social utility, as required by the utilitarian normative framework under consideration (given Bentham's definition of the sum total of utility as the aggregate of individual utility).

#### **4 The problems**

There are some problems surrounding the Benthamite model which deserve closer attention.

##### ***4.1 A functionary's duty to the public***

The notion of a functionary's duty towards the public might still seem odd. Firstly there are cases where it would be detrimental to the universal interest if the functionary would fulfil such a duty.

Assume, e.g., that some community is composed of just three sub-communities,  $c_1$ ,  $c_2$ , and  $c_3$ . As a matter of fact, their aggregate (majority or unanimous) interests on some public policy issue are  $a$ ,  $b$ , and  $b$  respectively; thus, as a matter of fact, the universal interest is  $b$ . But the members of  $c_1$  and  $c_2$  falsely believe their interests to be  $b$  and  $a$  respectively; solely the members of  $c_3$  (in the aggregate) have correct beliefs about their interest. If the three

---

<sup>20</sup> Bentham (1983[1830]:43).

functionaries simply vote according to their sub-communities' perceived interests, thereby fulfilling their duty towards the voters, *b* will result and – fortuitously – the universal interest will be promoted. If, on the other hand, *c*<sub>1</sub>'s functionary acts according to his duty to the public and, by delivering speeches, convinces his voters that their interest actually is *a*, whereas *c*<sub>2</sub>'s functionary fails to do the equivalent, the functionaries' final vote will promote *a*, to the detriment of the universal interest. Can there nonetheless be a morally justified duty for *c*<sub>1</sub>'s functionary to deliver his speech in favour of *a*?

One might be tempted to answer this question in the affirmative, for the following reason: for any such case, if *all* the functionaries did their duty (and succeeded in enlightening their voters as to their proper interest), the universal interest would be promoted after all. Yet it is doubtful if this kind of rule-utilitarian defence of the functionary's duty is what Benthamites (should) have in mind, since it faces the usual dilemma: Either the defence leads us straight back to an act-utilitarian conclusion (why should it not, in the above case, advocate the amended "rule" that all functionaries should do their duties except in cases where this doesn't maximise utility?). Or it implies that we in fact abandon the goal of maximising utility.

Our previous classification of the Benthamite model as a decision procedure can throw some light on the notion of duties here. To the extent that we can speak of *moral* duties, they must be derived from the moral criterion of rightness, i.e. the principle of utility. The duties of a functionary towards his voters and towards the public, however, are not directly derived from the principle of utility, but rather from the Benthamite model. Therefore, they are *institutional* duties, i.e. duties derived from the set of institutions (the decision procedure) which are justified by the principle utility. It then suffices to argue that implementing – on top of the functionary's duty towards his voters – a duty towards the public has better *overall* consequences than not implementing it. Such an argument might look as follows.

As stated above, a functionary's duty towards his voters is to vote according to their (aggregate) interest, as they perceive it. The duty is derived from the Benthamite model, whose ability to maximize overall utility relies on the correctness of representation of interests. The crucial assumption here is that, generally, the voters are the best judges of their own interests. The functionary's duty towards the community, on the other hand, consists in giving speeches to the effect of enlightening his voters as to their proper (aggregate) interest. It is derived from the Benthamite model in an analogous way, while accommodating the

additional assumption that the voters may at times be mistaken regarding their interests, in which case there is need for some corrective mechanism within the system.

This means that even in a case such as described above, where one functionary's failure to fulfil his duty to the public ( $c_2$ 's functionary fails to enlighten his voters about their real interest), this doesn't absolve some other functionary of his duty to the public (so  $c_1$ 's functionary should still enlighten his voters). In this particular case it leads to a non-optimal outcome. But given that the model is set up properly, regarding its duty-enforcing incentives, such failures are generally not to be expected, and hence the overall consequences from this model which includes such a duty towards the public are better than the overall consequences from this model without such a duty.<sup>21</sup> Thus the model with its two functionary duties is vindicated as a decision method in the light of the principle of utility.

However, the functionary's duty towards the public still seems odd given Bentham's theory on human motivation. According to Bentham, we generally cannot rely on any sort of sense of duty to actually motivate individuals against their interests. Bentham even suggests that we cannot truthfully speak of any imperative as a duty, insofar as compliance would go against the individual's interest: "it never is, to any practical purpose, a man's duty to do that which it is his interest not to do" (1983[1814-31]:121). This extreme view reveals Bentham's rather mechanistic psychology according to which individual interest determines what we are motivated to and hence psychologically can do.<sup>22</sup> Since, for Bentham, duty presupposes ability in this psychological sense, what one has a duty to do must necessarily coincide with what is in one's interest to do. Duties may therefore require some sort of incentive structures.<sup>23</sup>

So how does this play out for the functionary's duty towards the public in the above example? If there is such a duty, which demands the functionary to deliver speeches and is justified by

---

<sup>21</sup> One may wonder whether e.g. the following duty towards the public would work even better, by taking care of the above case with a functionary who fails to do his duty: "When conducive to overall utility, give speeches to the effect of enlightening your voters as to their proper aggregate interest, when not, don't!" However, there would then be problems with the epistemic feasibility: the proposal hinges on the assumption that the functionary (at least generally) is able to estimate overall utility for the entire community.

<sup>22</sup> Thanks to Roger Crisp for pressing this point.

<sup>23</sup> Cf. even Hart's interpretation of Bentham's notion of a duty: to say that someone has a duty to act in a certain way is just to say that failure to act in such a way will probably reduce her happiness, due to different kinds of sanctions attached to such failure. (Hart 1982:131-2). In other words: duties can be constituted by incentive structures, such as the one manifested in the model.

its beneficial impact on the universal interest, there must be some incentive structure to the effect of identifying his individual interest with his duty. The set of individuals who is most likely to maintain such an incentive structure is the set whose aggregate interests constitute the universal interest, i.e. the entire community. Thus, the functionary would not only be responsible towards (that is, punishable and rewardable by) both his voters, regarding his voting behaviour (since he has a duty towards them), but also towards the entire community, regarding his speech behaviour. Such double responsibility could create dilemmas: what should happen to a functionary who always votes according to his voters' perceived interests while never giving any required speeches and who thus is to be rewarded, i.e. kept in office (and possibly re-elected), by his voters – but punished, i.e. put out of office, by the entire community?

As it turns out, this seemingly double responsibility is simply a misconception. To be sure, the duty to the public is designed to ensure the promotion of the universal interest. But this doesn't mean that the incentive structure which establishes the duty necessarily has to be enforced by the entire community. In fact, in cases where the functionary's voters are wrong about their own interests, it is also in their interest to be enlightened by their functionary; and thus learning (retrospectively) that he failed at his duty would motivate them to punish him by dislocation or refused re-election. The incentive structure for the functionary's duty towards the public may therefore be set up and enforced by his own voters. The double responsibility vanishes.

Of course, now we get a new problem: following his duty towards the public promotes his voters' interests (as well as the universal interest) only if the functionary has correct beliefs in this matter. This seems to constitute shaky ground for such a duty. Maybe one could argue from the assumptions that (i) generally, the voters have incorrect beliefs as to their interest if and only if they lack certain information and (ii) a functionary's speeches (aim to) provide his voters with what he perceives to be the right kind of information. Then, if the functionary is right and his voters wrong about their interests, his speeches compensate for their lack of information, which would lead to correct beliefs. If, on the other hand, the functionary is wrong and his voters right, there is no harm in some (ineffective) speeches. Of course, if both are wrong, there is nothing much that can be done.

#### ***4.2 Unanimous vs. majoritarian elections***

Another problem concerns the question whether a functionary should be elected by a simple majority or unanimously, or in other words: whether, on the second level, he should vote according to the aggregate interest of either the majority or the whole of his voters. Bentham is ambiguous on this point – but the upshot of my argument is that both ideas are flawed.

Let us start with the first suggestion: functionary election by majority.<sup>24</sup> Now, even if we assume that each functionary is elected by an equally populated sub-community, different functionaries might still be elected by quite unequally large majorities within each of these sub-communities. Thus, we can construct cases where a majority of functionaries, voting for their sub-communities' aggregate interest (according to majority decision), vote for an alternative which is in the interest of less than half of the whole community members, i.e. not the universal interest.

This can be shown with a simple case: Assume there are three sub-communities, consisting of three voters each, who first are to elect three functionaries. The functionaries are then to vote on a public policy with two options: *a* or *b*. In sub-community  $c_1$ , the three voters' interests are *a*, *a*, and *b*, respectively; the same holds true of  $c_2$ . In  $c_3$ , all three voters' interest is *b*. Thus, the majorities of the  $c_1$ - and  $c_2$ -voters will each elect a functionary who will vote for *a*, whereas  $c_3$ -voters unanimously elect a functionary who will vote for *b*. When these three functionaries vote, in turn, the outcome by majority will be *a*. But this means that an option is chosen which is in the interest of a minority of all voters: only four have interest *a*, whereas five have interest *b*. The problem is that significant minorities may add up with majorities over communities, which may lead to non-optimal outcomes within the Benthamite model.<sup>25</sup>

This problem does not arise if we propose unanimous election of functionaries. For equally populated sub-communities, whenever a majority of functionaries, representing the unanimous interest of their voters, votes for an option, this option will be in the interest of a majority of all voters within the entire community.

---

<sup>24</sup> There may be different suggestions for the required majority level – here we can simply assume simple majority, requiring more than half the voters' votes, as default position

<sup>25</sup> This problem can easily be modified to apply to similar models which require a qualified majority for the election of a functionary. A simple case-generating formula is the following: for all  $m, n, x, y$  such that  $x$  denotes the number of communities with a preference for A to the degree of  $m$ , where  $m$  is above the required majority threshold, and  $y$  denotes the number of communities with a preference for B to the degree of  $n$ , if  $x > y$  and  $2xm + y < 2yn + x$ , a simple majority rule on the higher level will pick a non-optimific outcome. (Note that simple majority requires  $m > 0.5$ , not  $m = 0.5$ .)

However, this unanimity-proposal faces a dilemma: Either the proposal does or it does not allow heterogeneous interests (such as in  $c_1$  and  $c_2$  above) within any given sub-community. If it does allow this, then either some voters are likely to veto any decision, or a unanimously elected functionary will promote an incomplete policy profile which contains only unanimously agreed upon options – if any. This must be assumed due to the assumption that voters vote according to their (generally correctly) perceived interests, which means that they would not vote for any proposal which goes against their interests. Hence, only an incomplete proposal could be unanimously supported. But this means that the functionary's vote on any other issue is undetermined. However this problem is solved (whether, e.g., the functionary may abstain on these issues or flip a coin), the proposal fails to aggregate interests in some cases.

If, on the other hand, the proposal does not allow for heterogeneous interests within sub-communities, it will not work for any case where there actually isn't homogeneity in interest within each sub-community. Again, the assumption is that voters vote according to their perceived interests, so they cannot “compromise” in order to reach consensus. This means that the proposal again fails to aggregate interests in some cases.

So neither majority nor unanimity representation works for Bentham's purposes. One easy solution to this problem is the suggestion that we should stick with unanimity representation, but that the sub-communities boundaries should be flexible: they should be “redrawn” prior to each new election in order to include equal populations of individuals with homogeneous interests. Another solution is the suggestion that we should stick with majority representation, but that the functionary should “split” his vote and cast its parts in proportion to the relative amounts of aggregate interests within his sub-community. But these are just complicated ways of stating a simple alternative model which disregards sub-communities and has the entire community's voters cast their vote among a variety of candidates who promise to promote different voters' interests, and which assigns votes to any functionary in proportion to the number of his voters. Thus, the suggestion is that the Benthamite model's constituency representation should be abandoned in favour of *proportional interest representation*.

#### ***4.3 Motivation and ability assumptions***

The Benthamite model has each representative vote on a multitude of public policy decisions during his term of office. This has the advantage of reducing for each voter the total amount of decisions. But now we may question whether (1) each functionary candidate is able and motivated to design the "right" kind of policy profile, and whether (2) each voter is able and motivated to discern the actual content and value (in terms of her interests) of each candidate's policy profile. Let us start by looking at the second problem (2).

(2.a) The voters' *motivation* to discern content and value of policy profiles must be related to our motivational assumptions. Since the voter's end of maximising her own utility most likely presupposes knowledge of the available alternatives, we can infer the required motivational disposition. Yet, on the other hand, it seems obvious that finding and voting for the best policy profile is not the only means to utility maximisation for each voter, but rather one among many others. In other words, the *overall* benefits of attaining knowledge of the alternatives must be related to the *overall* utility reducing costs of doing so, in terms of resources (time, effort, money, and so on).

The problem becomes particularly acute when we consider the problem of the rationality of voting: the larger the community of voters, the less probable that any particular voter's vote should be decisive. So unless there are only a few voters, or the stakes in a decision are extremely high for some voter, her expected utility from voting is quite low, while the costs of getting informed and actually turning up at the ballot are rather high. So there may be greater or less chancy benefits in other activities for each voter. [There should be a separate chapter on the voter paradox...]

The upshot of this problem is that in order to make the Benthamite model work properly in the face of this problem, we would need to introduce some more incentives. These should either ensure that the costs of getting informed and voting are extremely low – or alternatively, that the costs of not getting informed and not voting are quite high. A pragmatic solution would be to make voting legally required and punish abstention, e.g. by fining abstainers, and to make information about the options easily accessible, e.g. by educating voters and spreading information for free.

(2.b) The voters' *ability* to discern content and value of policy profiles should be related to our epistemic assumptions. The *Principle of Self-Knowledge* assumes each voter to generally be a

better judge of the means to her well-being than anyone else. If this is taken to mean that the voter knows the content of any given policy profile, as well as the policy profile's value for her – at least better than anyone else – the assumption becomes quite implausible.

Let us then return to the first problem: whether each functionary candidate is able and motivated to design the "right" kind of policy profile. (1.a) What motivation is there for each functionary candidate to design the "right" kind of policy profile? First of all, what motivates each functionary candidate (once he has decided to run) is to be elected, i.e. to maximise votes. This he can achieve by designing a policy profile which attracts a maximum number (at least the required majority) of his voters; thus what he is after is a maximally broad compromise solution. Yet, the broader the compromise, the less attractive it will be to most voters. In other words, there is a tension between the functionary candidate's desire to attract as many voters as possible, and his desire to attract each individual voter as strongly as possible.

It has been suggested that the rational strategy for any candidate, ensuring maximum voter attraction, is then one of double ambiguity. (i) The candidate strives to spread his policies over a broad spectrum of possible policy preferences. Thus, the candidate's net position (the mean of his policies) is rendered ambiguous, since the candidate's net position from the perspective of each voter depends on how the latter weights each of the candidate's policies in the light of her own interests. (ii) The candidate strives to render each single policy ambiguous. This double strategy is rational for the candidate in the sense that reducing any of these ambiguities would cost him voter support.<sup>26</sup>

Unfortunately, the candidates' strategy increases uncertainty for the voters as to what policy profiles they actually vote on. This means that the policies which the elected functionary will promote may actually not represent (a majority of) his voters. Thus, our assumptions about the candidates' motivation reinforce the epistemic and motivational problems for the voters, as discussed above.

#### ***4.4 Minority coalitions***

---

<sup>26</sup> Cf. Downs, Anthony (1957) *An Economic Theory of Democracy*, New York: Harper & Row (ch.8).

Moreover, the feature of policy profiles can give rise to the problem of minority coalitions. Imagine the following case of *Shifting Majorities*. A community is made up of three voters, A, B, and C. They have the following preference profile on issues x, y, and z, with "+" and "-" indicating their policy preference on the issue in question: A ( $x^+$ ,  $y^-$ ,  $z^+$ ), B ( $x^+$ ,  $y^+$ ,  $z^-$ ), C ( $x^-$ ,  $y^+$ ,  $z^+$ ).

We can assume that any one of the voters may decide to run for representative. Running for representative gives them the benefit of promoting a preferred profile, yet it comes at a certain cost, so we can assume that they will only run if they have a chance of being elected. At first sight, since they are self-maximisers, all of them would run on a policy profile which matches their own profile perfectly. However, then there would be three candidates who each get only one vote – their own. So none of them would have a chance to get elected. Then any of them would win by adjusting such an idiosyncratic policy profile to attract a larger number of voters. Let's suppose that A promises to promote ( $x^+$ ,  $y^+$ ,  $z^+$ ) instead. Then, if A gets elected, each of the three voters gets two (out of three) of their preferred options. But suppose that B promises to promote ( $x^-$ ,  $y^+$ ,  $z^-$ ). This is a worse profile for A (who would then get zero out of three preferred options), but equally good for B and C. So the majority of B and C might as well vote for B. And so on. So there is no given winner here. (In terms of utility, if the stakes for all voters in every decision are equal, then profile ( $x^+$ ,  $y^+$ ,  $z^+$ ) would be optimal.)

However, imagine in addition that the three voters do not only differ in preference profile, but also in preference intensity on certain policies; this is the case of *A Passionate Minority*. Thus, B cares strongly about  $z^-$  – much more strongly than she cares about both  $x^+$  and  $y^+$ . Similarly, C cares strongly about  $x^-$  – much more strongly than she cares about both  $y^+$  and  $z^+$ . Thus, a candidate who would promote policy profile ( $x^-, y^+, z^-$ ) would receive B's and C's vote (disregarding A's possible preference intensities).<sup>27</sup>

In such a case, generally speaking, more than half the voters are in a minority regarding some policies. Furthermore, their preferences regarding these policies are stronger than the preferences they hold for policies which they happen to share with the majority. Thus, a

---

<sup>27</sup> Note that, at this point, we assume certainty for voters with regard to other voters' preferences.

candidate who promotes a selection of the policies which are intensely preferred by different minorities could easily outrun a candidate who supports the majority's views on all policies.<sup>28</sup>

#### *4.5 The case for equality of suffrage*

A further problem for the Benthamite model springs from the fact that it fails, in effect, in being sensitive to the voters' intensity of preferences. On the assumption that an individual's intensity of preference for some policy is correlated with this policy's amount of utility for the individual (which may be argued to be implicit in our *Claim to Self-Knowledge*), a model which takes the preferences as its base of aggregation but fails to take into account their intensity may lead to suboptimal outcomes in certain cases. (Bentham's assumption of equal intensity of preference is made due to the recognition of measurement difficulties, which might seem to be an unsatisfactory basis for the normative claim.)

One may object to this on the grounds of the previous argument: did the *Passionate Minority* case not show that the model does take account of intensity of preference in the aggregation? After all, the passionate minority was shown to outweigh the majority.

But surely, this is off track: we must distinguish between models which are sensitive to the voters' intensity of preferences in just some way, and those which are sensitive to the voters' intensity of preferences in the way which is required for its proper function.

To see that the Benthamite model does accomplish the first but not the latter task, let us consider the case of *Multiple Passionate Minorities*: again, there are three voters and their preference profiles A ( $x^+, y^-, z^+$ ), B ( $x^+, y^+, z^-$ ), C ( $x^-, y^+, z^+$ ). Again, we know that B cares strongly about  $z^-$  and that C cares strongly about  $x^-$ , which makes it rational for them to vote for policy profile ( $x^-, y^+, z^-$ ). But imagine now in addition that A cares strongly about  $y^-$  – much more strongly than she cares about both  $x^+$  and  $z^+$ , but even much more strongly than B or C care about  $z^-$  and  $x^-$  respectively. Thus, even A could form a minority coalition – with B for ( $x^+, y^-, z^-$ ), or with C for ( $x^-, y^-, z^+$ ). Yet, given any rational course of action for A, B, and C, the optimific outcome ( $x^-, y^-, z^-$ ) will not be effected. Even though the model can take intensity of preference into account, in virtue of there being passionate coalitions, it cannot

---

<sup>28</sup> Cf. Downs (ch.4). "[...] majority rule prevails in government policy formation only when there is a consensus of intensities [i.e. majority agreement on which issues are most important] as well as a consensus of views [i.e. majority agreement on which policy is to be preferred on each issue]." (Downs:67.) CHECK: James (1981:50) who criticises Bentham's argument for allowing a "tyranny of a number of minority interest groups".

take differing levels of preference intensity into account – and the latter is a prerequisite for optimific outcomes.<sup>29</sup>

Economic theory and its market models move from the assumption of rationality in individuals, commonly understood in terms of self-interest and efficiency, via the instrument of money, towards the end of wealth-maximisation. The model of democracy under discussion differs slightly: employing the same assumption for individuals, it moves via the instrument of *votes* towards the end of *utility-maximisation*. In order for such models to work properly in this respect, its instruments must ensure the proper relation between input and output. Thus, what is required for our model of democracy is that the model's instrument, i.e. the vote, should carry all information in the input which is relevant for determining the right output. The information required, apart from the *content* of the individual's interest or preference, is the *intensity* of her interest or preference, which, as we assume, correlates to her stakes in the decision, in terms of utility.

Ultimately, it seems that these difficulties call for a requirement of *proportional distribution of votes* according to stakes in the election of representatives. This means that representatives would not represent a number of voters but rather a number of stakes. The question is how this ever could be realised.

#### **4.6 Policy profiles**

Since single issue representation is impracticable, some form of policy profile representation is required. So the next question is: how should this be arranged?

As we assumed, voters are rational self-maximisers. Each votes, on single decisions, for whatever option maximises her utility. As has been established in the discussion on strategic voting, if there is a series of such decisions, the voter would vote for that combination of options which *in sum* maximises her utility. So for some random voter and a given series of decisions, there is an option profile, or policy profile, which in sum maximises her utility.

---

<sup>29</sup> In order to be more specific, and to explicitly connect policy preferences to outcome utilities, we can assign utility numbers for each voter's preferred policy outcome: e.g.  $(Ay^-)=10$ ,  $(Bz^-)=5$ ,  $(Cx^-)=5$ ; all the other utilities are 1 for the voter getting her way on the policy in question, and 0 for her not getting her way. The potentially winning policy profiles  $(x^-, y^+, z^-)$ ,  $(x^+, y^-, z^-)$  or  $(x^-, y^-, z^+)$  would then be worth 12 (0 for A; 6 for B and C each), 17 (11 for A; 6 for B; 0 for C) and 17 (11 for A; 0 for B; 6 for C) utility units respectively. In comparison, the optimum candidate's policy profile  $(x^-, y^-, z^-)$  would be worth 20 utility units (10 for A; 5 for B and C each). Yet, given these numbers, we can see clearly why any one of the former three but not the latter would be elected.

This profile may not map on to others' profiles. In fact, the more (mutually independent) decisions there are, the higher the probability that such profiles will be idiosyncratic.

Since it is any candidate's predominant interest to be elected, which requires a sufficient number of votes, there will most likely be no candidate who would promote our random voter's idiosyncratic preference profile. The same reasoning precludes even our random voter from becoming a candidate representing only herself. However, she now has two alternatives: either (a) to vote for a candidate who promotes some compromise solution which matches her preference profile as closely as possible while still attracting a sufficient number of voters, or (b) to become a candidate whose compromise solution fulfils these two conditions. Since running as candidate will be costly, voters will prefer (a), as long as their losses from "misrepresentation" are smaller than the costs of running. Thus, what emerges are "compromise" profiles which stand a chance of winning.<sup>30</sup> Depending on the structure of the community's profiles, as well as on the specific voting method, a number of candidates can then be expected to emerge among the voters.<sup>31</sup>

What has been explicitly introduced now is a three-level model of collective decision-making: on level (i) each voter decides whether to run as a candidate representative or not; on level (ii) each voter decides whether to vote for some candidate representative(s) or not (abstention), and on level (iii) each elected representative decides whether to vote for some public policy alternative(s) or not (abstention) (cf. Besley & Coate 1997). Keeping the Benthamite sanctions for representatives in place, we thus arrive at an improved model of interest representation.

Given the high costs of attaining knowledge on complex policy profiles, it has been argued that it may be rational to support candidates on an ideology basis. Thus, ideologies may work as cost efficient short cuts in an otherwise too costly process of decision making on level (ii).<sup>32</sup> On the other hand, ideologies are ambiguous and thus may increase the level of voter uncertainty further.

---

<sup>30</sup> I presuppose here that the incentive structure for representatives is set up properly so as to prevent them from promoting their own preference profiles once endowed with the power of office.

<sup>31</sup> Cf. Downs (ch.8).

<sup>32</sup> Cf. Downs (ch.7): "[Voting] on ideological competency, not on specific issues [...] is rational in two situations (1) having informed himself reasonably well, the voter cannot distinguish between parties on an issue basis, but can on an ideology basis; or (2) he votes by means of ideologies in order to save himself the cost of becoming informed about specific issues." (p.99.)

What motivation is there for each representative candidate to design the "right" kind of ideology profile? As assumed above, what motivates each candidate (once he has decided to run) is to be elected, i.e. to maximise votes. This he can achieve by designing an ideology profile which attracts a maximum number (at least the required majority) of the voters; thus what he is after is a maximally broad compromise solution. Such breadth can once again be efficiently achieved by appeal to ambiguity. Thus, an ideology profile which, e.g. with regard to certain resource distribution, appeals to "justice" is more likely to attract a large number of voters than any profile advocating a more specific distributive rule. On the other hand, since candidates are forced to distinguish their ideology profile from their competitors', there is some pressure towards less ambiguous commitments.<sup>33</sup>

Once again, there is a tension between the candidate's desire to attract as many voters as possible, and his desire to attract each individual voter as strongly as possible. Thus, even on this model, there is a tendency towards ambiguity, viz. on candidates' ideology contents. This means that the candidates' strategy increases uncertainty for the voters as to what policy profiles they actually can expect from their elected representative, and thus, that the policies which the representative will promote may not represent a (a majority of) the voters.

## References

- Bales, R. Eugene (1971) *Act-Utilitarianism: Account of Right-Making Characteristics or Decision-Making Procedure?* American Philosophical Quarterly, Vol. 8, No. 3 (Jul., 1971), pp. 257-265.
- Bentham, Jeremy (1962[1817]) "Plan of Parliamentary Reform: Introduction", in John Bowring (ed.) *The Works of Jeremy Bentham*, Vol.iii. New York: Russell & Russell.
- Bentham, Jeremy (1970[1782]) *Of Laws in General*, H.L.A. Hart (ed.), London: Athlone Press.
- Bentham, Jeremy (1970[1789/1823]) *An Introduction to the Principles of Morals and Legislation*. J.H. Burns & H.L.A. Hart (eds.), London: Athlone Press.

---

<sup>33</sup> Of course, in our three-level model, this uncertainty enters not only on level (ii), on which voters must choose between the different alternatives, but also on level (i), on which they must choose whether to run as a candidate or not. Given that there is maximum competition on level (i) – if costs of running are kept low – this may, however, be a lesser problem.

- Bentham, Jeremy (1983[1815/17]) "A Table of the Springs of Action", in: Amnon Goldworth (ed.) *Deontology together with A Table of the Springs of Action and Article on Utilitarianism*, Oxford: Clarendon Press.
- Bentham, Jeremy (1983[1829]) "Article on Utilitarianism: Long Version", in: Amnon Goldworth (ed.) *Deontology together with A Table of the Springs of Action and Article on Utilitarianism*, Oxford: Clarendon Press.
- Bentham, Jeremy (1983[1814-31]) "Deontology: or Morality Made Easy", in: Amnon Goldworth (ed.) *Deontology together with A Table of the Springs of Action and Article on Utilitarianism*, Oxford: Clarendon Press.
- Bentham, Jeremy (1983[1830]) "Constitutional Code" in Frederick Rosen & J.H. Burns (eds.) *Constitutional Code, Vol. I*, Oxford: Clarendon Press.
- Bentham, Jeremy (1989[1822]a) "Economy as applied to Office", in: Philip Schofield (ed.) *First Principles Preparatory to Constitutional Code*, Oxford: Clarendon Press.
- Bentham, Jeremy (1989[1822]b) "Identification of Interests", in: Philip Schofield (ed.) *First Principles Preparatory to Constitutional Code*, Oxford: Clarendon Press.
- Bentham, Jeremy (1989[1822]c) "Constitutional Code Rationale", in: Philip Schofield (ed.) *First Principles Preparatory to Constitutional Code*, Oxford: Clarendon Press.
- Bentham, Jeremy (1989[1822]d) "Supreme Operative", in: Philip Schofield (ed.) *First Principles Preparatory to Constitutional Code*, Oxford: Clarendon Press.
- Bentham, Jeremy (2002[1788]) "Considérations d'un Anglois sur la Composition des États-Généraux", in Philip Schofield et al. (eds.) *Rights, Representation, and Reform: Nonsense Upon Stilts and Other Writings on the French Revolution*, Oxford: Clarendon Press.
- Broome, John (1991) *Weighing Goods. Equality, Uncertainty and Time*, Oxford: Basil Blackwell.
- Burns, J.H. (2005) "Happiness and Utility: Jeremy Bentham's Equation", in *Utilitas*, Vol.17, No.1, March 2005, pp.46-61.
- Dinwiddy, J.R. (1975) "Bentham's Transition to Political Radicalism, 1809-10", in: *Journal of the History of Ideas*, Vol.36, No.4, October-December 1975, pp.683-700.
- Goldworth, Amnon (1969) "The Meaning of Bentham's Greatest Happiness Principle", in: *Journal of the History of Philosophy* Vol.7, No.3, July 1969, pp.315-321.
- Goldworth, Amnon (1972) "Bentham's Concept of Pleasure: Its Relation to Fictitious Terms", in: *Ethics* Vol.82, No.4, July 1972, pp.334-343.
- Hart, H.L.A. (1982) "Legal Duty and Obligation", in: *Essays on Bentham. Jurisprudence and Political Theory*. Oxford: Clarendon Press.

James, Michael (1981) "Public Interest and Majority Rule in Bentham's Democratic Theory", in *Political Theory*, Vol.9, No.1, February 1981, pp.49-64.

Rosen, Frederick (1983) *Jeremy Bentham and Representative Democracy. A Study of the "Constitutional Code"*, Oxford: Clarendon Press.

Schofield, Philip (2006) *Utility and Democracy. The Political Thought of Jeremy Bentham*, Oxford: Oxford University Press.

Sprigge, T.L.S. (1999) "The Relation between Jeremy Bentham's Psychological, and his Ethical, Hedonism", in *Utilitas*, Vol.11, No.3, November 1999, pp.296-319.