Mill’s Constitution of Liberty: an alternative behavioural policy framework

Shaun P. Hargreaves Heap
Department of Political Economy, King’s College, London, UK
Email: s.hargreavesheap@kcl.ac.uk

(Received 31 January 2023; accepted 31 January 2023)

Abstract
This paper offers an interpretation of the behavioural evidence that casts doubt over whether people always have preferences or, if they do, that they are stable. If people do not have preferences in this sense, then the usual policy evaluation standard of preference satisfaction cannot be used in these cases. The paper then develops, to fill this gap, a Millian policy framework where policy is judged by whether it advances what Mill understands by individual liberty in On Liberty. This yields many recognisable policies but offers a different evaluative standard for them. It also yields some policy innovations: e.g. a basic income and a flat tax.

Keywords: Liberty; Mill; flat tax; basic income

Introduction
In this paper, I consider what types of policies might follow from accepting Mill’s (1859, 1989) famous argument in On Liberty for the ultimate value of individual freedom. I call this policy framework Mill’s Constitution of Liberty. Crucially for my argument, Mill’s understanding of individual freedom has two elements. One is that an individual should be able to do anything so long as it does not harm others. The other is that the reason for valuing individual freedom in this sense is that it enables a person to become an individual; and this gives all persons, as individuals, an equal entitlement to liberty. These two elements are responsible for the policy framework that I identify with Mill’s Constitution of Liberty. Many of the policies in this framework will be familiar in outline even if the reasons for their adoption (and hence their evaluation) become somewhat different. There is, though, one that marks a clear departure from the status quo: an income tax system with a larger tax-free allowance (equivalent in size to an affordable basic income) and thereafter a flat tax.

I claim that Mill’s policy framework is ‘behavioural’ because I propose the framework as one way of responding to a significant challenge set by many of the...
behavioural insights from experiments in psychology and economics for the usual approach to policy formation. In this traditional approach, policy is constructed to achieve desirable outcomes. In the dominant ‘economic’ approach to policy evaluation, for example, the putative desirable outcome is preference satisfaction and the tool that is used to establish whether a policy could produce a potential Pareto improvement (i.e. a potential improvement in some people’s preference satisfaction without undermining that of others) is cost–benefit analysis. Likewise, in the more recent ‘nudging’ policy innovations, the point of a ‘nudge’ is to achieve a particular outcome like more pensions contributions or less consumption of salty, sugary and fatty foodstuffs and these outcomes are also often justified because they promote preference satisfaction (see Sunstein & Thaler, 2008). The particular challenge to this usual, consequential policy approach, and to which Mill’s Constitution of Liberty is one response, comes from the behavioural evidence that casts doubt over whether people really have a full set of stable preferences that can be used in this way for evaluating policy.

In On Liberty (and N.B. unlike in his book, Utilitarianism), Mill does not subscribe to the idea that people have settled preferences in this sense – such that their satisfaction could be used to evaluate policy. Instead, Mill’s picture of people has them in a flux. They are engaging in ‘experiments in living’ (we might call it ‘learning by doing’ today), reflecting on these new experiences and testing the conclusions that are drawn in discussion with others. This dynamic process is how people become autonomous individuals for Mill: that is, how they become an individual in the meaningful sense that they are in a position to feel they are authors of who they are. This process is what enables them to think of themselves potentially and truly as an ‘individual’. Mill’s particular claim in On Liberty is that a society with rules that embody the principle of individual liberty enables its citizens to become individuals in this sense.

Mill’s approach to policy might, therefore, be categorised as procedural because it focuses on the procedural desirability of the rules in contrast with the consequentialism of the more usual approach. This may, however, be a moot distinction since the achievement of ‘individuality’ in Mill’s framework could be cast as an outcome, making his evaluative criterion ultimately as consequentialist as the more usual approach. Nevertheless, it is clear that there is an important difference in these approaches because policies in Mill’s framework are judged according to whether they advance individual liberty and not whether they achieve some specific outcome like more pension contributions or an array of outcomes that yield a positive net-benefit. One cannot know what preferences one might come to have, or indeed value in Mill’s view. This is inevitably an open question. But, one can say something about the process generating the outcomes with the result that the process can be judged even though the specific outcomes themselves cannot. In particular, for Mill, the question is: do the rules that guide and enable our actions advance individual liberty?

It may seem somewhat puzzling to cast this orientation to the rules as a behavioural approach that differs from the consequentialism of ‘nudging’ when ‘nudging’ has been explicitly developed as a behaviourally informed policy framework. The puzzle though only arises if the behavioural evidence is interpreted in a particular way. To
clear the ground of this puzzle, I take up this issue of how to interpret the behavioural evidence in the next section.

In section ‘Individual liberty’, I develop the characteristic features of Mill’s Constitution of Liberty and in section ‘A Millian progressive flat tax system’, I focus on the specific form of social insurance that is warranted by the no-harm principle and which explains the proposal for a relatively large tax-free allowance coupled to a relatively large flat tax on income. In fact, this proposal has a special claim to be Millian. Mill advances it in his *Principles of Political Economy*, although only partly for the reasons that I develop from his arguments in *On Liberty*.

**Interpreting the behavioural evidence**

There are two broad insights from experiments in psychology and economics. Both, for different reasons, can make problematic the assumption that people have stable individual preferences that motivate their actions. As a result, the usual approach to policy evaluation can lose its anchoring standard of preference satisfaction and this leaves a policy void. Mill’s Constitution of Liberty is one response to that void. In this section, I discuss how the behavioural evidence can have this effect.

One of these broad and potentially destabilising insights is that we frequently do not behave as rational choice theory would predict in non-interactive settings: that is, in manner consistent with having a preference ordering over outcomes and acting so as to best satisfy these preferences. One example of this is that people seem to be influenced by what should be entirely extraneous information. For instance, suppose subjects are offered a choice between a weekend in Rome and a weekend in Paris, both with all expenses paid. Some will choose one, some the other. It seems, though, that when subjects are also offered a third option of Paris, but here the subject has to pay for the breakfasts, more people opt for the all expenses paid weekend in Paris than when there are only the two all expenses paid options (see Ariely, 2008).

Another illustrative anomaly comes from experiments where people face two lotteries: the Pbet one has a large chance of a small win and the $bet alternative has a small chance of a big win. When asked how much they are willing to pay, most people place a higher value on the $bet, but, when given a straight choice between the two, most people select the Pbet. Thus, they appear to reverse their preferences depending how the same choice is presented.

One interpretation of such anomalies is that people have two systems of decision making (see Kahneman, 2003). System 1 uses a set of rules or heuristics for decision making. It is quick and requires little, if any, use of mental energy. System 2 involves a rational reflection over the options and a considered decision over which might best satisfy a person’s interests. It takes time and uses up scarce mental energy. Of necessity, people can only use system 2 to a limited extent (as people have limited mental energy or cognitive processing power) and so most decisions are made using system 1. This means relatively simple changes to the decision problem, that ought from the perspective of rational choice theory to have no effect, may, in fact, alter decisions because the change in the choice architecture has interacted with the system 1 decision rule that is being used.
This is the interpretation underpinning ‘nudging’ (see Sunstein and Thaler, 2008). People, in this interpretation, are in a deep sense rational in a rational choice sense, it is just that the ‘of necessity’ use of system 1 on frequent occasions may sometimes produce decision that are anomalous from the perspective of rational choice theory. The ‘nudge’ is designed to interact with the decision heuristic to set matters straight for the inner rational choice agent.

An alternative interpretation of the susceptibility of people’s decisions to the way the same decision problem is presented is that people do not have preferences as such; and the different ways of presenting the decision problem trigger different ways of thinking about which option to choose. In the case of the Paris extraneous information example, people cannot easily make up their mind between Paris and Rome when there are only the two all expenses paid options because they do not have well-defined preferences from which a conclusion over which best satisfies these preferences can be drawn. People may be able to see some of the merits (and drawbacks possibly) of both. Paris, after all, is well known for its romance, it was also home to various of the 20th century’s modernisms, etc.; Rome, on the other hand, speaks to our classical heritage, both in a landscape sense and because Rome launched many of our contemporary political sensibilities, it will also likely be warmer, etc. The likely experiences of the two cities are, however, in many respects incommensurable. How can one weigh romance against heritage of this sort?

One psychological explanation of decision making that answers this question imagines that people build a balance sheet of ‘reasons’ for each action and, when one achieves a suitable threshold number of reasons, it gets chosen (see Stewart, 2009). This can make the decision sensitive to the way that the options are presented because the mode of presentation can trigger different reasons to enter into this balance sheet. For example, put in the third Paris option that is dominated by the all expenses Paris weekend and it seems perhaps that there is one more reason for choosing the Paris all expenses paid weekend – it is clearly better now than one of the other options, the third one where you have to pay for breakfasts in Paris. Of course, it does not make super-good sense to be swayed in this way, but it is surely intelligible.

The difference in interpretation of the evidence is not so much over whether people use decision heuristics, since Stewart’s (2009) decision by sampling is a decision heuristic. Rather the difference is whether to assume there is also system 2. The presumption of a background system 2 appeals in this respect, when system 1 is cast as an unthinking application of a decision heuristic, because people believe that they have the capacity for reason-directed action. So, there has to be some space in the model of decision making for reason and a presumption of system 2 provides it. However, the cognitive models of psychologists like Stewart (2009) avoid the need to build-in system 2 in the background for this purpose because they endow the decision heuristics themselves with the attributes of reason, albeit understood more loosely than, say, the rational choice sense of reason. We use the kind of balance sheet enumeration of ‘reasons’ precisely because we have not got some underlying set of preferences to consult (see also Simon, 1978). This is a model of decision making where the decision indeed reveals the preference but not because there are underlying preferences, antecedent to the decision. It is the decision that creates the apparent preference and not the other way round. The decision itself is best
understood through the cognitive processes whereby we process information and act upon it; and if there is a measure of consistency across decision tasks, then it is because other cognitive processes, like those of cognitive dissonance removal, reinforce that preference once it has been conjured into existence by the original decision.

In support of an interpretation that dispenses with system 2 in this way, there is a much older tradition in economics that puts matters this way round without the explicit cognitive psychology: the Austrian tradition and the work of Hayek in particular (see Hayek, 1960). So, it is not without deeper roots in economics. For Hayek, it was simply self-evident that we did not have *ab initio* preferences for all the extraordinary new goods and services that the market has furnished. It was also self-evident to him that we had a category of what we would now call ‘social preferences’ (the social norms and customs in a society) that did not belong to a single individual the way ordinary preferences might and which changed over time.

This Hayekian connection is a useful cue for the second broad set of behavioural insights from experiments that I want to mention, but it is also helpful for the general argument of this paper in another way. Hayek identifies the value of social customs and norms (i.e. social preferences) through the way that they coordinate behaviour and so help overcome what he sees as the developing ‘knowledge problem’ in affluent societies. For him, the knowledge problem arises because the extension of the division of labour responsible for growth increases the complexity of the economy. No single mind can entertain how the various parts of this complex whole fit together and how they will evolve. It only works as a system because we rely on institutions like the market (and social customs) to coordinate our behaviour when all we can do is act on the very limited information of the world that we can possibly ever have. This ‘knowledge problem’ provides a different reason for why policy cannot always (and increasingly cannot be) evaluated by the outcomes that it produces. We simply cannot know what the outcomes will be. This is the unintended consequences message that comes with the complexity of the economy. The problem with policy evaluations that are oriented to their outcomes, therefore, is not just that we may not have the basis to value outcomes using the metric of preference satisfaction, we may not be able in a complex economy to identify many of the outcomes of a policy.

The second broad set of behavioural insights relate to the existence of social preferences. In many interactive settings, we behave unselfishly and this behaviour is subsumed within the rational choice model by allowing that some of our preferences may be ‘social’ – that is, oriented in one way or another to the interests of others. The particular insight that I want to draw on for the argument of this paper is that these social preferences often seem in experiments to be endogenous rather than exogenous. In particular, there are many experiments that illustrate the ‘crowding out’ of social preferences when a particular decision problem is shifted to a market rather than non-market setting or when a decision problem is given an explicit financial frame (see Bowles & Polania-Reyes, 2012, for a survey). In other words, we often have social preferences, but they are not stable.

In summary, I do not want to suggest that we never have preferences and that they cannot be stable in ways that, in principle, could be used to evaluate some outcomes of policy. I simply want to suggest that there are reasons, particularly from
experiments in decision making, for supposing that we often do not have stable and antecedent preferences that explain our decisions. Thus, we cannot always evaluate policy by how well outcomes satisfy those preferences. There is room to dispute how many outcomes are affected in this way, but the general point of the observation is reinforced by the Hayekian argument over the difficulty of identifying the full set of outcomes of any policy in a complex economy. The shared implication of these points is that we need to find some other way of evaluating policy than preference satisfaction.

**Individual liberty**

One response to this challenge is to judge policies by whether they have desirable procedural characteristics. Do they advance, for example, the principle of individual liberty? This is the possibility that I pursue in this section by developing what might be the character of policies that would advance Mill’s version of individual liberty. There are two important features of his account of individual liberty for this purpose. One is the no-harm principle. Mill did not advance a free-for-all version of individual liberty: people should be free to do anything so long as it does not harm others. The other is that everyone, by virtue of being an individual, has the same claim to individual liberty because liberty is what enables an individual to acquire their individuality.

This last feature has two important consequences for policy. The first is the obvious one that policies should treat people equally. This property is sometimes referred as the rule of law. Second, liberty is important for Mill because it allows experiments in living which are the material for critical reflection and discussion; and it is this process of reflection and discussion, guided by the canons of free speech, that enable the growth of an individual’s sense of autonomy. However, while liberty is a necessary condition, it is unlikely to be a sufficient one for the generation of individual autonomy. Individuals will need some critical capacities and maybe some material ones too in order to be able to reflect on experience and engage in discussion with others. Policy should therefore ensure that people have these capabilities – for example through the provision of public education and perhaps something like a basic income. Of course, there will be room for policy discussion over what the appropriate level of these capabilities might be, but the policy framework should provide for such capabilities and the ensuing discussion might in many respects be different from the current debates around the education and welfare systems. For example, an education system may be judged less by how well it serves the economy (e.g. by producing workplace specific skills and by ranking students through public credentials) and more by how it equips everyone with general purpose cognitive capacities.

The no-harm principle has three important implications for a policy framework. First, people need to know what aspect of the outcomes associated with their actions are harms if they are to be guided by the no-harm principle in their decision making. For this purpose, the definition of a harm must be public and shared because a purely individually subjective definition of a harm will be open to strategic manipulation by a person to advance their own interests over those of others. In practice, the courts and the government supply these definitions and so, rather like the education and
welfare system requirements, this Millian requirement for institutions that generate policy is a recognisable feature of our contemporary policy landscape. Once publicly defined in this way, harms have, in effect, a ‘price’ put on them and people can as a result *ex ante* take the force of the no-harm principle into account when making their decisions.

Second, policy has to be alert and adapt to the dynamism of an economy and society. Otherwise, inherited policies may in new circumstances cause the generation of widespread harms unwittingly and so work against the no-harm principle. For example, for much of human history the generation of greenhouse gas emissions in any time period has been below the threshold level that can be absorbed by the atmosphere in that time period without changing the concentration of these gases. As it is the concentration of these gases that affects our climate, policy need not attend to these emissions until this threshold is reached. But once it is reached, policy innovation is required to establish property rights in the atmosphere or some equivalent form of control over its use so as to avoid the otherwise wholesale generation of harms and derogation of liberty that comes through climate change. Likewise, our inherited policy regarding the absence of rights in aspects of our personal data may have to change in the future as a result of the digital technological revolution. This is because our personal data may potentially, under current arrangements that allow its free harvesting and algorithmic manipulation by Facebook, Twitter and others, be used in ways that cause wholesale harms.

The point of such illustrations is that no set of property rights (or policies) is sacrosanct and a Millian policy framework should encode their provisionality in this sense. Zero-based budgeting is one example of such a presumption. Unlike normal budget setting, where the historical funding levels are the reference point for discussing what are bound, if this approach is followed in all cases, to prove small increases/decreases, zero-based budgeting makes no such concession to the historical funding level. Its reference point is zero. The case for any spending, rather than small departures from the *status quo*, has to be made afresh. Another example are sunset clauses. They likewise build-in provisionality by giving an expiry date for any programme and its funding.

A small comment is worth making at this point. These illustrative arguments for policy adaptation seem to turn on what looks like the standard policy argument for policy intervention when there are public goods or market failure. Correct. Furthermore, the case for sunset clauses and zero-based budgeting is also usually made on the same grounds that ‘times change and so should policy’. The argument is different, however, in one important respect. The criterion for a policy intervention is the presence of unnecessary wholesale harms, where the definition of harm is public, shared and determined through the institutions of government and the courts. In comparison, in the standard policy argument in relation to market failure, a policy intervention is triggered when the policy can generate a potential Pareto improvement in terms of people’s subjective experience of preference satisfaction. The Millian proposal has a different (potential) Pareto principle: policy should remove harms where in principle the removal of one person’s experience of harms does not increase the harms experienced by another.
Third, while the courts and government can establish what are harms, people will not be able to take account *ex ante* of all these prospective harms in their decision making for the simple reason of Hayek’s ‘knowledge problem’. People cannot anticipate all the outcomes of their actions and some of the unintended outcomes will entail harms. This is what happens in a complex economy/society and the problem seems likely to get bigger as economies become more complex. A commitment to the no-harm principle means that there must be policies that address this possibility of unintended harms. By its nature, this problem cannot be solved through policies that try to influence *ex ante* decisions by taking these harms into account. The policy will instead have to operate *ex post* to compensate people for harms they suffer. These harms are not anticipable by the authors of the actions that have caused them. Indeed, such is the complexity of life that these harms may never, even *ex post*, be capable of being traced to their precipitating actions. What is needed instead are *ex post* policies of social insurance that cover and compensate those who experience harms in general.

Again, there are recognisable policies for social insurance in the contemporary policy landscape: policies of public health care offer insurance for citizens against health harms; and unemployment benefits offer similar insurance against the financial harm that comes from becoming unemployed. In a similar fashion, the income tax system gives those who experience falls in their income an insulating cut in the tax that they pay. The difference that comes from Mill is in their justification and therefore in the nature of the policy debate around their amendment and change. I develop this point in more detail with the tax system in the next section.

**A Millian progressive flat tax system**

The usual policy discussion of income tax systems, taking as given the level of public expenditure, focuses on the incentives to work that are created by the system and the trade-off this may create between efficiency and equity when the tax (and benefit) system is also used to promote equity. The trade-off arises because progressivity is typically engineered through higher marginal tax rates for the better-off and this will discourage work and most highly so among those who are the most productive. Another aspect of this trade-off arises when welfare benefits (introduced for reasons of equity) are conditional on income as the effective marginal tax rate now also depends on the rate at which benefits are withdrawn as income rises.

From the perspective of Mill’s Constitution of Liberty, such a discussion overlooks the key *desiderata* of the income tax system: its social insurance properties. Indeed, from this vantage point, the typical tax system with higher marginal tax rates for the better-off perversely (and undesirably) gives more pay-outs for negative income shocks to those who plausibly least need it – the already well-off, who are, after all, better equipped for self-insurance. A constant marginal tax rate would treat all people equally in the sense that they receive the same insulation from income shocks of the same magnitude. This would also be attractive on the grounds of ‘equal treatment’ from the ‘rule of law’ Millian policy requirement. Of course, equal treatment need not necessarily require that everyone is literally treated equally in this way because the concept of equal treatment can be sensitive to relevant sources of individual
difference. The virtue of a flat tax in this respect, though, is that it requires no further discussion or agreement over what the relevant sources of individual difference might be. Absent such discussion and agreement, the flat tax is attractive because it, relatively speaking, uncontestably treats people equally.

Of course, the drawback of a flat tax is that it undermines the progressivity of the income tax system. However, progressivity does not depend exclusively on an increasing marginal tax rate. It can also be achieved by increasing/creating a tax-free personal income tax allowance. Mill (1848, Book V, Chapters 2 and 3), incidentally, makes just such a proposal for a tax free level of income to cover ‘necessities’ and a flat tax thereafter on income that is productively earned and used to purchase ‘luxuries’.

I give two illustrations for the UK of how the existing broad pattern of progressivity could be maintained by moving to a flat tax when combined with a larger tax-free personal allowance. For the purposes of making this comparison, I give the current marginal and average rates of the UK tax system at the time of writing in columns 2 and 3 of Table 1. In columns 4 and 5, I change the tax free allowance from the current £12.5 k to £15.5 k and introduce a flat tax of 45%. In columns 6 and 7, I have a £15 k tax free allowance and a flat tax thereafter of 42%.

Both flat tax illustrations are similarly (or even more, perhaps) progressive than the current system: e.g. both tax those earning £20 k less and those earning above the average income of £30 k more. There is one exception to this progressivity: those earning over £200 k pay less tax. With these features, the total tax take is unlikely to change materially with either flat tax. I chose these particular illustrations partly for this reason, but also because a tax free allowance of £15.5 k or £15 k would be the equivalent for those earning more than this amount to a tax saving of £7 k or £6.3 k. This is important because another way of interpreting this illustrative alternative tax system is that there is a flat tax of 45%–42% and a universal basic income of £7 k or £6.3 k instead of the tax-free allowance. Consider, for example, someone earning £20 k: with the £15.5 k tax free and a 0.45 flat tax, they pay c.£2 k tax and have take home income of £18 k. With the alternative basic income of £7 k and a flat tax of 0.45 on all the earned £20 k, they pay £9 k tax but receive £7 k as basic income and so their net payment to the Treasury is again £2 k and their take home income is again £18 k. The virtue of the £7 k–£6.3 k tax saving (or basic income) figure in this context is that the cost of a universal basic income of this order of magnitude would be roughly the same as the cost of all current welfare payments including pensions and the current tax free allowance in the UK. In other words, this level for a basic income would be affordable as a substitute for those programmes.

Thus, an alternative and more radical interpretation of the illustrations in Table 1 is that they represent a universal and unconditional basic income of £7 k–£6.3 k together with a flat tax on all earned income of 45%–42%. One virtue of this alternative interpretation is that it connects to the earlier suggestion that Mill’s requirement of some minimal capabilities might be satisfied in part through a basic income. The other virtue of construing the tax free allowance in these illustrative calculations as coming from a basic income that substitutes for the current panoply of welfare programmes is that the basic income avoids the disincentive effects to work,

---

1 Mill allows for an increasing tax rate on unearned income, like gifts.
also mentioned earlier. The point is that basic income is unconditional and so there is none of disincentive to work effect that arises under the current system where there is benefit withdrawal as income rises.

### Conclusion

We frequently do not know all the outcomes arising from a policy and we may not be able to evaluate those outcomes we can anticipate. Mill’s Constitution of Liberty is a response to the challenge of how to engage in policy evaluation in such circumstances. It offers a different rationale (and hence evaluative criteria) for many of the central policies in our current framework and it potentially suggests a more conspicuous change with respect to our current tax and welfare system: a basic income plus a flat tax.

### References


---

**Table 1.** Current and illustrative UK tax rates (source for current: the Martin Lewis Money Saving site).

<table>
<thead>
<tr>
<th>Income</th>
<th>Current marginal tax rate</th>
<th>Current average tax rate</th>
<th>New marginal rate % (£15.5 k tax free)</th>
<th>New average tax rate % (£15.5 k tax free)</th>
<th>New marginal rate % (£15 k tax free)</th>
<th>New average tax rate % (£15 k tax free)</th>
</tr>
</thead>
<tbody>
<tr>
<td>20 k</td>
<td>33.3</td>
<td>12.5</td>
<td>45</td>
<td>10.1</td>
<td>42</td>
<td>10.5</td>
</tr>
<tr>
<td>30 k</td>
<td>33.3</td>
<td>19.6</td>
<td>45</td>
<td>21.8</td>
<td>42</td>
<td>21</td>
</tr>
<tr>
<td>55 k</td>
<td>43.3</td>
<td>26.7</td>
<td>45</td>
<td>32.3</td>
<td>42</td>
<td>30.5</td>
</tr>
<tr>
<td>100 k</td>
<td>43.3</td>
<td>34.1</td>
<td>45</td>
<td>38</td>
<td>42</td>
<td>36</td>
</tr>
<tr>
<td>200 k</td>
<td>48.3</td>
<td>42.5</td>
<td>45</td>
<td>42</td>
<td>42</td>
<td>39</td>
</tr>
</tbody>
</table>