How representative is the House of Commons – and why is it so difficult to measure

Draft 1

PHILIP COWLEY and AN OTHER

When we talked about the representativeness of a parliament we used to mean its partisan representativeness: the extent to which the party composition inside parliament related to the proportion of votes cast for each political party. Yet increasingly any such discussion now refers to what is usually called descriptive representation: the extent to which those elected are representative of the wider social and economic make-up of the country. How many women are there? How many MPs from ethnic minorities? How many from a middle-class background?

Both these topics are the subject of vast academic literatures, but much of this work looks at various aspects of representativeness in isolation.[1] Studies of descriptive representation usually consider one variable by which those elected to parliament may or may not represent the broader population, with the choice of that variable usually depending on the interests of the researcher and the research agenda being followed. It is relatively rare to find attempts to measure the combined social representativeness of a parliament, by taking multiple variables into account. No work (to our knowledge) attempts to take a range of variables and measure their combined representativeness systematically and comparatively, either spatially or temporally. And again (at least to our knowledge) no one has attempted to combine a measure of both social and partisan representativeness to measure the overall representativeness of a parliament.

As will become clear, there are all sorts of methodological difficulties involved in doing so, and (based on the experience of writing this paper) it soon becomes clear why no one else appears to have tried it – or at least why no one seems to have succeeded. But the exercise still seems a valuable one, because it would allow us to answer important questions. Which parliaments are the most or the least representative? Have parliaments become more or less representative over time? These are fairly fundamental questions to which we should have answers – but we currently do not.[2]

In this paper, we use a worked example – the British House of Commons - to demonstrate some of the problems but also hopefully also the value in doing so. We focus on the Commons, both because we are both familiar with it but also because there is relatively good data available, both on its members and on the wider British population (although, as will be discussed, even here, these data are not without problems).

We start with the situation after the 2019 general election and use this to explore some of the methodological issues inherent in the exercise. We then attempt to extend the data back in time, around 30 years, which throws up even more methodological issues. We then add in measures of partisan representativeness. We show that, combining both descriptive and partisan representation together, the British House of Commons has become more representative over time.

The Commons in 2019

---

[1] This literature is so vast that we do not cite it here, except where a piece of work is especially relevant.
[2] There is an analytically separable, if in practice often conflated, debate about diversity in parliaments. Diversity is a distinct concept and one which this paper does not touch on.

We start with the House of Commons after the 2019 general election, as detailed in Table 1.

The first column of data reports the percentage of the House of Commons in six often-discussed categories: ethnicity (showing the % of BAME MPs), sex (% women), education (% non-graduates), age (% of MPs aged between 18 and 39), sexuality (% LGB), and disability (% disabled). The second column of data then gives the equivalent percentage for the UK adult population.

The fourth and fifth columns report two different measures of representation, comparing the Commons with the wider population. The fourth column gives the percentage point difference between the two: that is, % MPs minus % population. Negative scores therefore indicate a group is under-represented in the Commons relative to its size in the wider adult population, positive that it is over-represented. The fifth column shows the extent to which a group is represented proportionally: that is, the percentage of a group in parliament divided by its percentage in the wider population (multiplied by 100). A score of 100 would mean the group is present in the House of Commons in the same proportion as in the wider population; a score of 50 that it is there in half its proportionate amount, 200 double, and so on.

1. Group representation in the 2019 UK House of Commons

Group	Group size		Difference	
	Adult population	MPs	Percentage points	Proportion represented
BAME	12.3%	10.0%	-2.3	81.3%
Women	51.1%	33.8%	-17.3	66.1%
Non-graduates	73.3%	13.0%	-60.3	17.7%
Young	36.2%	20.6%	-15.6	56.9%
LGB	2.7%	6.9%	+4.2	255.6%
Disabled	22.8%	0.8%	-22.0	3.5%

Source: Data on population sex and age are taken from ONS population estimates (NOMIS); data on population disability, education, and ethnicity are taken from the Annual Population Survey 2019; data on sexuality are taken from ONS, *Sexual Orientation, UK, 2019*.

In broad terms, the table should be of little surprise to anyone who follows these things at all closely. The House of Commons has seen a well-documented rise in recent years in the number of women MPs, as well as those from ethnic minority backgrounds, but both groups are still under-represented numerically. Like many parliaments, it contains a lot of university graduates and exhibits considerable middle-aged spread in its composition. It has been previously noted for having a high number of LGB members, while the representation of disabled people has long been known to be very poor.

Yet it is relatively rare to see these data collected and presented together like this, along with comparisons to the wider population. Just by listing the differences next to one another in this way, it becomes clear where the larger disparities are and the extent to which some groups are over- or under-represented more than others.

The two measures of representation tap into different things and can produce different findings in cases where the groups are of different sizes. For example, for Britain's BAME population, the point difference might seem small at just over two percentage points, but this represents only 80% of its

size relative to the wider population. A two-point difference for women (by contrast) would equate to 96% representation.

As a result, there are some slight differences in the rank ordering of the six groups, depending on whether we are interested in the point difference or the proportional difference. But non-graduates and the disabled vie for being the least well represented – depending on whether we are measuring the difference by percentage point or proportionally. Next come women and the young, again with the precise order depending on which measure is utilised. The second-best represented group, albeit still under-represented, are BAME, while the best represented group are LGB – the only group of the six to be over-represented in the Commons compared to their numbers in the wider population.

Drude Dahlerup famously devised a categorisation for analysing the percentage of women in an institution: a parliament's composition was 'uniform' if women constituted just 0-15% of members; it was 'tilted' if women made up between 15 and 40%; while it was 'balanced' if 40% or more. We can adapt these benchmarks and labels, to take into account that not all the groups studied here constitute roughly 50% of the population. We say a group is 'excluded' if they are represented at between 0-30% proportionately. Just as Dahlerup's uniform parliament was not literally uniform, such a group is not literally excluded, but their numbers are so small that they may as well be. A group is 'under-represented' between 30 and 80% - that is, it is present in the institution but at levels clearly below proportionality. Between 80% until they reach 120%, a group is 'represented'. Again, just as Dahlerup's balanced parliament was not literally balanced, the group may not have reached perfect proportionality but it is now represented in large enough numbers that this may not matter (and similarly, it may even have gone over 100% by a similar amount). Above 120% we can say that a group is 'over-represented'. On this basis, in the House of Commons in 2019, non-graduates and the disabled constitute excluded groups; women and the young are under-represented; BAME people are represented; while LGB people are over-represented.

Various methodological issues should already be apparent.

First, it is never clear in the discussion on this topic whether a parliament's population should match the *total* population of a country or its total *adult* population. Or perhaps even the *adult working age* population. It is possible to make a plausible case for all three of these comparisons. If Parliament is to be a "mirror of the nation", and to claim to represent everyone, perhaps we should compare with the total population. But MPs legally cannot be aged under 18, and so to include children in the comparator can skew any findings. Similarly, although there is no compulsory retirement age for British MPs, they mostly still do stand down at some point and we should not reasonably expect large numbers of the very elderly to be present in parliament. For some of the characteristics in which we are interested, these three figures will be near enough identical, but for others they will not be. There are issues with age (obviously), ethnicity (where the white population is more elderly than the non-white population) and perhaps most starkly of all, disability. The 2019-2020 Family Resource Survey, for example, reports disability figures of 46% for those of state pensioner age, 19% for working age adults, and just 8% for children. These differences do not matter much when making general points about under- or -over representation – the disabled are clearly under-represented in parliament, whichever metric we use – but they will matter if the goal is to construct more detailed measures of the level of representation. Here we use adult population where possible, but we accept this is moot.

Second, there are issues of data collection and reliability, which are non-trivial even when it comes to a relatively well-documented institution such as the House of Commons or a wider population like that in the UK, where good census data exists (and as discussed further below, these problems

become even greater when it comes to trying to produce comparable data across time). These problems can apply to census data – it is, for example, surprisingly difficult to get census data solely of the adult population, for example – and some of the measures are based on samples or are otherwise estimates. But there are also issues with data on MPs, which is not always as robust as you might expect. Not all data sources agree with one another. These issues are perhaps most obvious when it comes to measuring sexuality and disability. Because so many disabilities are invisible or "hidden", one study of the involvement of the disabled in political life argued that it was "almost impossible to capture the precise numbers of disabled politicians".[3] The figure of five disabled MPs for 2019 (that is, under 1%) is almost certainly an under-estimate, but it is the only figure currently available. Data on sexuality is similarly tricky. For the wider population we use the ONS figure from 2019 of 2.7% LGB of those aged over 16.[4] For MPs, we use those MPs "out" after an election. This too will almost certainly be an under-estimate of the actual number of LGB MPs – indeed, we know it is, because some MPs come out later in their careers.[5] But it does have the merit of comparing those who are out at Westminster with those who are willing to be out to a survey company, which is broadly a like-for-like comparison. We focus here on LGB, and not LGBT (or any wider definition) simply because there is no reliable data for the wider population.

Third, there are issues of coding and categorisation. For example, the table uses an overall BAME category, which is how ethnicity in the Commons is usually reported.[6] Such broad categories can sometimes mask as much as they illuminate, hiding differences between different ethnic groups. Similarly, it uses an overall LGB figure, rather than providing three separate pieces of data. When discussing age, we focus on the 18–40-year-old category, when different measures might yield different figures.[7] This is anyway a fairly broad definition of youth (albeit not broad enough to include one of the authors anymore), but it is one used by the only existing comparative study of age representation in parliaments.[8] For disability, we use the overall figures for all those with "substantial and long-term" disabilities; the figure for what are sometimes called major impairments is about half this, and there must be a suspicion that the figure for MPs mostly covers those with major impairments. Plus, there is no attempt here to measure or discuss intersectionality – the extent to which the variables interact with each other – or partisan differences, the extent to which these results might differ by political party.

Then, fourth, more fundamentally, there is the question of which variables to include at all. These six variables obviously do not cover the full range of ways in which parliament may or may represent the nation. Discussion of this subject normally focuses on characteristics which are politically significant, but this still leaves plenty of other characteristics unexamined here. We could, for example, include social class, schooling, religion, or regional background.[9] Moreover, even of those

> **Commented [PC1]:** Having done it, I do wonder if we need to include social class and schooling.
>
> Schooling probably wont be too difficult (% state?).
>
> But class will be a swine. I've never seen any decent data which has coded MPs into the various categories which we use for the population. It *might* be possible to do % working class…?

---

[3] L. Evans and S. Reher, 'Barriers to elected office for disabled people', Government Equalities Office, 2021.
[4] A figure for those over 18 was not available.
[5] This poses an especial problem when MPs who were not out in previous time periods come out later in their careers. Should they be counted as LGB in earlier parliaments? Here we err on the side of caution and include only those out at the time.
[6] However unappealing the terminology, in reality this measure is actually measuring non-white MPs; such figures exclude Irish or Jewish MPs, for example.
[7] There must also be a concern that focussing on the young obscures what is happening at the other end of the age range – those aged over 70, for example. Perhaps a better measure might be the middle aged -40-60, say, who really are the most over-represented group.
[8] Sundström and Stockemer (2020).
[9] A heretical footnote: the most significant ways in which parliaments are unrepresentative of their wider population, of course, is that they are full of people from political parties and/or who are hugely interested in

who believe in the merits of descriptive representation, not everyone would agree that over-representation of graduates is necessarily a bad thing; they might think this is something to be encouraged.[10] Fairly obviously, we may reach different conclusions depending which variables we include.

Many of these issues are resolvable, at least in theory. It should be possible to construct different measures for various sub-groups – such as different ethnic background or sexuality – although self-evidently with any such increases in specificity come additional data collection problems. The same applies to attempts to show intersectionality or to measure partisan differences. We can obviously include different variables, at least where the relevant data can be found. We could also generate different benchmarks or utilise different data.

That said, much of this is unlikely to make a huge difference to the overall point being made in the table. Take, for example, the figures for disability. We might decide to halve the figure for the wider population, to focus only on those with major impairment; we might simultaneously double the figure for the number of MPs, arguing that there were more disabled MPs than just five – and yet *still* the disabled would be one of the most under-represented groups in the table. Or, similarly, we might double the figure for LGB people in the wider population, believing the ONS estimate to be too low. Yet LGB MPs would *still* be the best represented group of the six examined here. In other words, while there may be a whole range of issues with the precise figures detailed in the table, we are fairly confident that they present a broadly accurate portrayal of descriptive representation in the House of Commons in 2019.

Change over time

This section attempts to replicate the data in Table 1 for earlier time periods, to test whether the Commons has got more or less representative over time – and if so, by how much.

In total, we take the data back just under thirty years. Because of problems with data availability we go back a decade each time, to the year of the census (that is, 1991, 2001,2011), in each case comparing to MPs elected in the nearest election (that is, 1992, 2001, 2010), which are at most a year distant from the census. In each case we take the position of the House of Commons as it was immediately after election.

In some cases, this is fairly straightforward. For example, the number of women in parliament has gone up decade-on-decade over this period.[11] From 9.2% of the Commons in 1992, it reached 33.8% in 2019. At the same time, the proportion of women in the wider population has remained basically constant, with just a small decline from 52.4% of UK adults to 51.1%. The result is a clear increase in levels of representation, with the point difference dropping from -43.2 to -17.3 and the proportionate figure rising from 18% to 66%.

Things are more complicated with other groups for which we have reliable data, because in each case the comparator group changes in size more significantly. Over these 30 years, for example, the proportion of young MPs has increased (up from 12.7% in 1992 to 20.6% in 2019). But at the same time the British population has aged, with young people now making up a smaller proportion of the adult population (down by about 5 percentage points). The effect of this combination of rising

---

politics. The scale of over-representation of these two groups massively dwarfs anything else, yet is never discussed.

[10] On this, however, see M. Bovens and A. Willie, *Diploma Democracy*, OUP, 2017, which sets out the many reasons why this can be a problem.

[11] But not quite election-on-election, because of a dip in 2001.

numbers of MPs representing a smaller group of people is to magnify the apparent change in representation over time. Young people in 1992 were represented at about 30% of their group size; this reached over 55% by 2019.

The changes in the size of the wider population are even more stark with non-white and non-graduate Britons. In 1992, BAME MPs made up just under 1% of the House of Commons. Between then and 2019, they increased in size more than tenfold, to 10%. Yet at the same time, the British adult BAME population has roughly trebled in size, from 4.5% of the adult population to 12.3%. The result here is to dampen the apparent growth in representation. The percentage point difference changes from -3.6 to -2.3 over these 30 years, and while there is a growth in the proportional figure from 20% to just over 80%, this is a fourfold increase rather than the tenfold increase in the raw figures.

The proportion of graduates in the UK adult population has been even more dramatic. In 1991, just 7% of the adult population had been to university. By 2019 this had almost quadrupled. There has simultaneously been a growth in the number of graduate MPs, although not as stark, up by about 15 percentage points, from 72% to 87%. In 1992, then, graduates were represented at about ten times their presence in the wider population; that is now down to just over three times. Non-graduates have seen their percentage point under-representation change from -65 to -60.3, but they have also seen their proportionate rate drop, from around 30% to below 20%.[12]

We do have data on the number of out LGB MPs. This rises from 0.2% of the Commons in 1992 (that is, one MP) to just below 7% in 2019. However, we lack reliable data for the wider population. As a thought experiment, we have kept this latter figure at the same level as in 2019 for the preceding three decades. Holding this figure constant like this means a change in the proportion represented from 7.4% in 1992 to 66.7% in 2001, to 129% in 2010 reaching 255% in 2019. None of the other groups studied achieves this rate of change – from being excluded to over-represented in just two decades.

We so far lack any reliable disability data for earlier time periods. In what follows, we have set the figure for MPs at just 3 MPs (which was the figure in 2015). There have long been disabled MPs, but they have often been remarkable by being so well known for their disability – Jack Ashley, David Blunkett, Anne Begg and so on – that it seems unlikely the actual number will be much higher (at least if we are comparing on a similar basis to the figure of five used in 2019). We have set the figure for the wider population at 19.0%, which is the figure given in the 2010/11 FRS data.[13]

Table 2 shows the percentage point variations across time. It is important to note that in one sense none of these calculations change some fundamental observations about representation in the British House of Commons. With the obvious and striking exception of LGB MPs, groups that were under-represented in 1992 remain under-represented in 2019. The extent of their under-representation changes, but this essential fact does not. Yet at the same time, it is striking that most of these groups have seen improvements in their representation. The only exception is the disabled, where our data are the most problematic (although where equally, we have no reason to think there has been anything other than the tiniest change in levels of representation).

---

[12] I have to confess I don't understand this at all. If graduates are becoming less well represented; surely non graduates should be becoming better represented? I need to put my thinking cap on.
[13] This figure is for the entire population. The figure for all adults will be slightly larger. Earlier data is collected by the FRS, but on a different basis.

2. Percentage point difference from proportionality, 1992-2019

	Election			
Group	1992	2001	2010	2019
Non-white	-3.6	-4.9	-7.0	-2.3
Women	-43.2	-34.3	-29.5	-17.3
Young	-28.9	-26.9	-17.9	-15.6
LGB	-2.5	-0.9	0.8	4.2
Disabled	-18.5	-18.5	-18.5	-22.0
Non-graduates	-65.2	-51.1	-50.0	-60.3
Average	**27.0**	**22.8**	**20.6**	**20.3**
Weighted average	**16.5**	**12.2**	**10.5**	**10.7**

The second from bottom row presents an average of the absolute values of point deviations, to create one measure of overall descriptive representation. This sees a fall in the overall unrepresentativeness of parliament, down from a score of 27.0 in 1992 to 20.3 in 2019. The fall is decade-on-decade, although the change in the last decade has been rather slight.

Presenting the data in this way might indicate that all these characteristics are broadly equal in their importance. They are certainly not all equal in size, though. Imagine, for example, two groups, one of which comprises 50% of the population, another just 5%. And imagine two parliaments. In Parliament A, the larger group was represented relative to its size in the population, whereas the smaller one was not represented at all. In Parliament B, the situation was reversed, and the smaller characteristic was represented in full, the larger one not at all. We would almost certainly think that Parliament A was doing a better job at representing the wider population than Parliament B.

The final row in the table is therefore a weighted average, averaging the absolute difference in each case multiplied by the percentage the group makes up in the wider population. This too shows a clear improvement in levels of representation to begin with, but with representation becoming very slightly worse between 2010 and 2019.

Table 3 repeats the exercise, but showing representation as a proportion. As with Table 2, most groups see an increase in the levels of representation over this time period. In 1992, all but one of these groups could be classed as 'excluded' (and the one exception, the young, sneaked in just 0.5 percentage points over the bar); by 2019, only two of them were still so classified. The penultimate row calculates an average measure (taking the mean average of the absolute value of 100 minus the relative size of the group's presence). Because this is calculated as the difference from perfect representation, lower values indicate better representation. The table shows increasing levels of representation until 2010, worsening thereafter. The final row calculates a weighted average, multiplying the scores for each group by their size in the wider population. This shows levels of representation getting better over time, but only very slightly after 2010.

3. Proportional representation of social groups, 1992-2019

	Election			
Group	1992	2001	2010	2019
Non-white	19.9%	26.7%	37.6%	81.3%
Women	17.6%	34.3%	42.7%	66.1%

Young	30.5%	31.9%	51.9%	56.9%
LGB	7.4%	66.7%	129.6%	255.6%
Disabled	2.6%	2.6%	2.6%	3.5%
Non-graduates	29.9%	35.9%	31.5%	17.7%
Average (deviation from 100)	**82.0**	**67.0**	**60.5**	**71.7**
Weighted average	**27.0**	**22.8**	**20.6**	**20.3**

We are aware of arguments that say that it is not a problem when previously marginalised groups become over-represented. In all of the above deviation above 100% representation is counted in the same way as deviation below 100%. It would be perfectly possible to decide to label deviations from proportionality above 100% at zero. At the moment, this only applies to LGB MPs since 2010, and it would make almost no difference to the findings.[14]

Again, as with the initial analysis, there are multiple problems of data collection, even going back just this far. Occasionally, data are not perfect like-for-like measures. As a result, it would be unwise to place too much stress on any individual data point. However, you measure it, the overall trends seem clear. There was an increase in the representativeness of the British parliament after 1992, until 2010. After that, depending on the measure utilised things either slightly improved or regressed.

Of these four indices, our preferred measure is the final one, the weighted proportional one. This takes into account different group sizes but also captures better the relative presence or absence of a group. By this measure, then, the Commons has got progressively more representative over time since 1992 – with its descriptive unrepresentativeness reduced in total by around a quarter over those 30 years.

Ideally, we would go back before 1992, but data issues prevent this. We can for some variables, but it becomes increasingly difficult to sustain for these same six variables.

Party

Measures of partisan representation are well established in the academic literature. Table 4 shows the two most commonly utilised measures – the Loosemore-Hanby and the Gallagher indices, for each of these four elections. Each, in different ways, measures the deviation from proportionality of the election results. The final two rows of data then average each of these measures with the weighted proportional representation score from above. Regardless of which is used, the overall result is the same: the Commons becomes more representative – measuring both partisan and descriptive characteristics – over time.

4. Political and overall representativeness

	Election			
Indices	1992	2001	2010	2019

---

[14] It would matter only with the unweighted proportional index, which is the one most prone to being skewed by an individual outlier, but for all others the difference is marginal.

Gallagher	13.6	17.8	15.1	11.9
Loosemore-Hanby	18.0	22.0	22.8	17.2
Overall representativeness (using Gallagher)	20.3	20.3	17.9	16.1
Overall representativeness (using Loosemore-Hanby)	22.5	22.4	21.7	18.8

Discussion and conclusion

To be written.

Private schooling will show an increase in levels representation.

Class will show a decrease in representation, although my suspicion will be that, like graduates, it will be from such a high base that it won't actually make much difference.

When constructing the index, we want variables that are independent of one another (at least theoretically). Else, there is a danger of double counting some things.

Want to try to do it comparatively? I think any paper might at least need an attempt at that – but it is obviously going to be hellishly difficult. All the data problems will just increase, massively.