

Factual – Measured - Critical

Creating Public Attitudes to Immigration by Mis-Counting John MacInnes and Sin Yi Cheung

Immigration and ‘statistics’ have a long and intimate relationship. It usually takes the form of abusing the latter to foster hostility to the former. The most recent example is the government’s ‘go home’ poster van campaign, which the Advertising Standards Authority retrospectively banned for using misleading arrest statistics (1). That is misleading as in false, wrong, incorrect.

However the Home Office has no monopoly on misleading statistics about immigration. In September 2013 a respectable liberal broadsheet covered a story with the title, “Majority of voters think immigration is harming Britain” (The Independent, 1/09/13), based on an ‘authoritative survey’ (2). The survey was commissioned by Lord Ashcroft, former Deputy Chairman of the Conservative party. The report entitled “Small island: public opinion and the politics of immigration” indeed looks very impressive and official with @Lord Ashcroft KCMG PC 2013 on the front cover, along with a picture of the (now banned) government poster advertising its campaign against illegal immigrants (3).

The survey of public attitudes to immigration was based on ‘a poll of more than 20,000 people’ which revealed there were ‘seven segments of opinion on the subject, each with definable concerns and priorities’. The poll was supplemented by a ‘day long deliberative event’ called ‘Immigration on Trial’. Amongst other things, the poll discovered that the government’s controversial poster advertising campaign against illegal immigrants was supported by four out of five of those polled and expressing an opinion.

Lord Ashcroft, an ‘author and philanthropist’ is also Chancellor of Anglia Ruskin University. He knows that reports which contain a ‘methodology’ section will appear to be ‘authoritative’. This clearly impressed the Independent journalist who covered the report, and, perhaps, more predictably, the Daily Mail. Other papers that covered the report and passed it off as representing a genuine survey of British public opinion included the Telegraph, Sunday Times, Guardian, Huffington Post, Daily Express, Daily Mail, Daily Star and the Sun. Most newspapers referenced the feature that probably attracted most journalists’ attention: that it was a poll of over 20,000 people, variously described in the papers as ‘adults’, ‘voters’ (not the same thing), ‘British’ and so on.

It is a measure of how weak even the liberal broadsheet press is when it comes to scrutinising statistical evidence or data. Their maxim at times appears to be ‘if it is a number it is probably correct and if it is a big number it must be authoritative’.

The ‘poll’ of 20,062 people was an online one. Online research can be high quality, as the experience of organisations such as Yougov shows. However, online polls suffer

Factual – Measured - Critical

from three major drawbacks which make them unsuitable for survey research. Anyone can enter an online poll (or at least can do so if they receive an invitation, happen to come across one, or happen to surf onto the poll website). Of course, those most likely to spend their time doing so are those with strong views on the subject(s) concerned.

The report gives no information at all about how these 20,062 people were selected. The implication is that it was a 'convenience sample': that is, where those who knew about the survey and wanted to take part did so. The trouble with convenience samples is that, except in rare situations where no other options are available, they are virtually worthless. Self-selecting convenience samples produce bizarre results.

A well known example is Shere Hite's 'statistics' on infidelity which grossly overestimated the prevalence of marital infidelity in the United States. In her book *Women in Love*, based on a 'survey' of almost 5,000 women who chose to take part, she claimed that 70% of married women had had an affair within five years of marriage. Later researchers used the US General Social Survey to repeat her question using a proper random sample. They estimated the true figure at under 5%. That is why reputable pollsters and survey organisations go to great, and expensive lengths to make survey samples 'random': that is samples in which every member of the target population has a known probability of selection. Without such a sampling process the whole statistical edifice comes crashing down. But random samples are simply not possible 'online'.

Anyone can enter an online poll and claim to be somebody completely different. There is no way of verifying basic details such as age, sex, area of residence, and so on. Household pets and film stars are frequent 'respondents' to online polls. An online poll of British adults need be nothing of the sort. Anyone, anywhere on the planet, or off it for that matter, can appear to be in Britain if they have a VPN connexion to a UK site on their computer. There neither is, nor can be, an online poll of 'voters' since whether someone is on the electoral register, or eligible to be on it, is not something that can be verified 'online'.

Worse still, pets, film stars, and aliens can participate again and again. Many with 'strong views' probably do. That is why reputable online polling companies do not really run 'polls'. Rather they build up a bank of respondents, whose characteristics they can check or verify in some way, and whose participation is driven by the survey organisation rather than the respondent. They do this because otherwise they obtain results that are statistically worthless.

But, says the 'methodology' section, the results have been 'weighted' to make them 'representative of all adults in Great Britain'. If Lord Ashcroft was indeed able to do this he should patent his procedure forthwith and retire on the proceeds, as it would be of unimaginable value to survey research and marketing organisations.

Factual – Measured - Critical

Like alchemy, no such weighting is possible. What can be done is to take the assortment of ‘respondents’ as measured by their self-reported characteristics and weight them to correspond to known characteristics of the British population such as age, sex, ethnicity, occupational class, place of residence, place of birth and so on. The more characteristics controlled for, the more complex this process becomes. Of course since we have no way of knowing if any of the characteristics we are weighting for have been accurately reported anyway, this need not worry us too much.

Thus, let us run a thought experiment in which such weighting has been done with great attention to detail and by some magic process we can be sure that all respondents have conscientiously and accurately reported to us their various characteristics. Do we now have a ‘representative sample’?

The answer is no. We can only weight the sample according to known characteristics of the population from which it has been drawn. If we don’t know the distribution of the relevant characteristic in the population we have no way of estimating the weight. But why did we carry out the poll in the first place? Because we wanted to know about unknown characteristics of the population, such as their attitudes to various aspects of immigration. We cannot weight for these. Yet it is precisely these attitudes on which we might expect our online survey to be wildly non-random (in the sense of every member of the target population having an equal probability of selection) and non-representative, since those with an interest in immigration issues, regardless of their other characteristics, will have had more motivation to participate. We could apply as many weights as we wished and not remedy this defect. A poll which is not based on a randomly drawn sample, or the closest practicable analogue to it, is worthless. Full stop. There are no statistics under the sun that can turn a convenience sample into a random one.

Lord Ashcroft doesn’t bother to tell us in the ‘methodology’ section how the online ‘interviews’ (there were two lots: one of 20,062 adults, one of 2006) were organised. This hardly matters, since there is no conceivable way they could have been organised to overcome the various problems outlined above. Actually the methodology section comprised a grand total of 152 words, most of which are used to say that ‘discriminant analysis’ revealed that there are ‘seven segments of public opinion’. No indication is given of where this comes from. Presumably from factor analysis of the responses on some variables. However, without a random sample, these ‘segments’ can be nothing more than ‘segments’ of the 20,000 people who took part in the poll. There is no way to generalise the results from such a poll to all ‘adults in Great Britain’.

The journalists covering the story could at least have cast an eye over the survey instrument. It features such examples of question design as:

Factual – Measured - Critical

“Which two of the following, if any, most concern you about immigration into Britain?

Immigrants claiming benefits and using public services when they've contributed nothing in return
Increasing the pressure on public services like schools and hospitals
Immigrants being given priority over established residents when it comes to benefits or public services
Immigrants taking jobs that would otherwise go to British workers, or pushing down wages in general
Changes in the character of local areas with large numbers of people not originally from Britain
No concerns”

Leaving aside the point that having ‘no concerns’ cannot constitute ‘two’ options, this question is not so much leading as one that won the race long ago and has since retired. Conspicuous by their absence from the survey are any questions asked in other, more robustly organised, enquiries, such as the British Social Attitudes Survey. Such inclusion is standard methodological good practice, as it allows benchmarking.

The ‘methodology’ section contains no details about whether question items or scales were randomised (essential to avoid question or probe order effects) but the absence from the ‘poll results’ section of data on non-response or “don’t know’s” makes one suspect that such methodological basics are unlikely to have been followed.

In short, the ‘report’ is methodological nonsense. But, of course, that didn’t matter. Dressed up with some erudite speculation about the nature of the ‘segments’ of public opinion revealed by the ‘discriminant analysis’ it did the essential job of hoodwinking a gullible press, some of it eager to be so hoodwinked, into thinking that here was a ‘scientific’ report of public attitudes. What it could only be, of course, was a cynical exercise in creating and shaping public opinion by asserting, on the basis of no scientific evidence whatsoever, that the opinions found in the online ‘survey’ represented a cross section of British opinion. It could do this, unfortunately, because most of the British press (unlike its US counterpart) has little knowledge of, or interest in, statistics. Such ignorance allows the likes of Lord Ashcroft to lead it by the nose in whatever direction they please.

This is no isolated example. The widespread respect for any statistics and for those who produce them by the British media is also evident in the BBC coverage of “Romania and Bulgaria immigration” based on the estimate by Migration Watch (4). Their Home affairs correspondent even praised the track record of this pressure group on ‘analysing future population trends’ and declared that ‘its latest forecast should be taken seriously’. Headlines like “50,000 Romanians and Bulgarians will arrive in the UK...” was all over the front pages of the Telegraph, the Daily Mail and

Factual – Measured - Critical

the Financial Times. However, at a recent migration statistics conference, Migration Watch admitted that their estimate is based on ‘a judgement’ and ‘an educated guess’ rather than any statistical methods at all.

For the sake of democracy, the British press should get to grips with statistics.

References:

1. http://www.asa.org.uk/Rulings/Adjudications/2013/10/Home-Office/SHP_ADJ_237331.aspx
2. <http://www.independent.co.uk/news/uk/politics/majority-of-voters-think-immigration-is-harming-britain-8793233.html>
3. <http://lordashcroftpolls.com/wp-content/uploads/2013/08/LORD-ASHCROFT-Public-opinion-and-the-politics-of-immigration2.pdf>
4. <http://www.bbc.co.uk/news/uk-21039087>

John MacInnes is Professor of Sociology at the University of Edinburgh and ESRC Strategic Advisor on Quantitative Methods Training.

Sin Yi Cheung is Senior Lecturer in the School of Social Sciences at Cardiff University.