

# Distributional Analysis

## Challenges, Opportunities, and Next Steps

### Overview

In October 2018, the Scottish Government held a seminar on distributional analysis involving experts from 14 external organisations, together with internal analysts in finance, social security, equalities, economic modelling, and strategy. Angela O'Hagan, chair of the Equality Budget Advisory Group (EBAG), chaired the event. Presentations by EBAG, Audit Scotland, Landman Economics, and ONS on the importance and practicalities of distributional analysis were followed by breakout sessions focusing on Scotland and the UK, protected characteristics, and the social contract.

This paper summarises the outcomes of the seminar. On the one hand, it is clear that there are significant challenges to performing distributional analysis which captures the entire social contract for each of the protected characteristics, particularly in a small country with a partially devolved administration. Broadly speaking, these challenges can be divided into issues of data and issues of modelling. It is also clear, however, that there are steps which can be taken to improve and extend the analysis currently performed by the Scottish Government in both the short and long terms.

### Challenges and Opportunities

#### *Data Issues*

One of the main challenges of performing distributional analysis in Scotland is the relatively small size of the Scottish sample in UK-wide surveys such as the Family Resources Survey (FRS) and the Living Costs and Food Survey (LCF), the two main data sources for most distributional analysis in the UK. Although the Scottish FRS sample has been boosted, and the Scottish LCF sample is due to be boosted, it is still necessary to pool multiple years of data. This will continue to be necessary even after the LCF boost, as, at 1,600 households, the Scottish sample will still be too small to use single-year estimates. Pooling, however, runs the risk of concealing distributional changes within rolling averages, and also means that new or improved survey questions cannot be incorporated until they have been running for a number of years. Ethnicity is particularly affected by low numbers – and even more so when further breakdowns are required, for example by country of origin – although this issue could possibly be mitigated by weighting the data according to demographic patterns.

There are also issues with the data sources themselves. The FRS, for example, is known to underreport benefit take-up. The Resolution Foundation has developed a methodology which attempts to address this problem by allocating missing benefit income to eligible households that do not report receiving benefits, but this is inevitably a simplification. Very high and very low incomes are also mis-reported. At the lower end of the income scale,

inconsistencies may stem from fluctuations in earnings and working hours. One method to address this problem is to assume that nobody earns less than the national living wage, but more realistic methods may be possible in the future as the FRS will include a question on why the individual is earning less than the living wage. At the upper end of the income scale, the FRS already includes an adjustment for top earners based on the Survey of Personal Incomes (SPI). However, the top decile includes a much wider range of incomes than other deciles, and data quality becomes an issue when attempting to split the decile further. Usage of public services by high earners is also underreported, although this problem can be imperfectly addressed by allocating residual public spending to high-income households.

Gender poses unique challenges for distributional analysis due to the lack of data on how income is shared within households, which obscures both distributive impacts and behavioural responses. To sidestep this issue, all benefit income can be assigned to its immediate recipient, be it an individual or a household, with proportional splits applied when a benefit is received by more than one person. Ideally, however, empirical data on income sharing, such as the data collected by the EU Statistics on Income and Living Conditions (SILC), would be incorporated into the analysis, bearing in mind the limitations of self-reported surveys. It may be advisable in any case to present a range of possible impacts based on varying assumptions about how income is shared within the household. In the case of disability, meanwhile, the main data sources are not perfectly comparable, since the FRS contains a number of variables on disability while the LCF reports disability only indirectly through receipt of disability-related benefits.

Administrative data, such as the data recorded by the HMRC and the DWP as they collect tax and disburse benefits, is considered the gold standard in the field of economic modelling – not least because it avoids the pitfalls of survey methodologies, including sample deficiencies and self-reporting errors. In the long term, administrative data will likely be used to enhance and possibly replace survey data to derive more accurate information on incomes, including tax and benefits. However, the implications of the General Data Protection Regulation (GDPR) on distributional analysis, particularly with regard to administrative data, require further consideration.

In the shorter term, surveys other than the FRS and LCF may be used to enhance distributional analyses. The Scottish Household Survey (SHS) in particular contains a variety of potentially useful information. At present, the SHS only includes incomes for up to two adults in each households, but from 2018 the data collected will include income from additional adults; the benefit questions are also due to be improved. Panel surveys such as Understanding Society (US) may also have value. The Wealth and Assets Survey (WAS) is somewhat less promising, since the Scottish sample is relatively small and the survey runs biennially, meaning that pooled data sets would have to cover longer periods of time.

### *Modelling Issues*

The microsimulation models used in distributional analysis can be either static or dynamic. Whereas static models assume that household behaviour remains fixed when a policy is implemented, dynamic models incorporate predicted changes in behaviour. Dynamic models are not necessarily superior, however, since they introduce additional assumptions which may be difficult to articulate and which may not be realistic. For example, a dynamic

model is liable to overestimating changes in behaviour which are actually inelastic in the short to medium term, such as place of residence or working hours, and may neglect differences in behaviour between households. In practice, the two types of models are unlikely to generate significantly different results, at least in terms of the income distribution. It is also possible to 'bolt on' behavioural responses to an otherwise static model.

An important question in distributional analysis is whether to consider the impacts of individual policies or the cumulative impact of multiple policies. This question is particularly relevant for Scotland, since policy changes by the UK and Scottish Governments both have effects. Income tax is an illustrative example, since the two administrations are responsible for different parts of the system, which interact to determine outcomes for individuals. The most appropriate approach will depend on the specific situation, including why the analysis is being conducted and what it purports to show.

A related question is how to present the results of distributional analysis. Some Scottish policies, such as the Carer's Allowance top-up and the Best Start Grant, will have a very small impact on the overall income distribution, relative to changes in, for example, income tax. However, individual households may still experience significant impacts, which will be obscured if the analysis is only based on average incomes. Setting out the 'winners' and 'losers' of a policy or set of policies, possibly complemented with qualitative analysis, is an alternative method which may be more enlightening.

There are also questions regarding which components of government activity to include in the model. On the one hand, distributional analysis which incorporates indirect tax and public services is desirable because it provides a more complete picture of the relationship between individuals and the state (the 'social contract') than an analysis which focuses solely on direct taxes and benefits. Public spending is of especial interest to the Scottish Government due to its particular budget powers and the divergences between UK and Scottish policy in this area. On the other hand, while public services tend to be progressive, they are based on the principle of universality and as such are not intended to be redistributive in the manner of tax and benefits. For this reason, performing distributional analysis on public spending is conceptually problematic.

Public spending is also the most difficult component to model from a technical perspective. Generally speaking, the first step involves determining patterns of service use, either by gathering usage data or by estimating usage patterns based on household characteristics. However, the benefits of a given service cannot always be ascertained from usage of that service. Most obviously, some services benefit everyone equally and are not exactly 'used' by anyone. This is the case with defence and environmental protection, for example, which are generally excluded from distributional analyses. Even when usage differs between groups, moreover, it may be a misleading datum on which to focus. Individuals who do not use a given service may still benefit from it, and one person's use may have knock-on benefits to others. The NHS, for example, provides insurance to all members of society, not just those who have existing health conditions, and also benefits employers by supporting a healthy workforce.

Conversely, it is often simplistic to assume that, because certain people use a public service more, they would benefit more from public spending on that service. For example, women use buses more than men, but it would be misleading to conclude that spending on buses is positive for gender equality, since asymmetric bus use is likely to reflect underlying inequalities. To properly interpret this usage pattern, and thus to draw implications for policy, would require a more sophisticated narrative. Similarly, although areas with higher crime rates in some sense 'use' police services more than others, this usage pattern is likely associated with other inequalities which would not be addressed solely through police spending.

The second step to incorporate public services is to combine service use with some kind of 'price' in order to translate in-kind benefits into monetary values. The standard approach in economics is to ask or infer what the individual or household would pay for the service if it were not provided by the government. However, this approach raises questions about the counterfactual, since private provision is inherently different from public provision. In the case of education, for example, private education may be more expensive per person than public education, but there may also be differences in quality which prevent a like-for-like comparison. Furthermore, if private education were to replace public education, its price would likely be different than its current price. To avoid these issues, market values can be eschewed in favour of allocating public expenditure to different households based on their usage of a given service use, although this approach does not take into account differences in how households value that service.

Some aspects of the social contract are easier to model. Resource expenditure, for example, is generally more straightforward than capital, the benefits of which may not accrue immediately and may be more dispersed across the population. Even resource expenditure raises problems, however, since the benefits may spill over from the immediate beneficiaries. For example, if NHS pensions were increased, it could be argued that the wider public would benefit in addition to NHS workers and their families through an improvement in the service. Finally, incorporating indirect taxes like VAT involves marrying tax rates with spending patterns, and in this sense is relatively simple – although it does require additional data, which, as noted earlier, is limited in Scotland.

## **Next Steps**

The Scottish Government is increasingly expected, and indeed obliged, to consider the impacts of its policies on equality groups through robust analysis. The budget in particular has come under scrutiny as an instrument which could be more deliberately employed to improve equality outcomes, particularly as Scotland's budgetary powers expand. Distributional analysis is one way of meeting these aims, although, as this paper clearly shows, it is not without limitations.

In the short term, the Scottish Government will continue to perform distributional analysis for individual policies, such as income tax. In the long term, it will continue to progress work on the conceptual and technical feasibility of distributional analysis, which will include seeking to improve data on intra-household distribution.