Improving Distributional Analysis of the Scottish Government Budget

Scoping Paper

Background

The Budget Process Review Group (BPRG) has recommended that the Scottish Government "explore the feasibility of providing a distributional analysis, by equality characteristic, of the taxation, expenditure and social security proposals in the Budget (recommendation 43)". After providing an overview of distributional analysis, this paper considers the models and data that have been used or could be used to conduct distributional analysis according to the BPRG's recommendation.

Distributional Analysis

Distributional analysis is a method for estimating the financial impact of a given set of policies on different groups. This kind of analysis is useful because different groups, including those with protected characteristics, are likely to gain or lose more than others from any given policy. Distributional analysis can therefore help determine whether policies achieve their objectives, whether they are consistent with overarching strategies, and whether they meet social criteria such as fairness and equality.

When it comes to large-scale, long-term policies, like those pertaining to the budget, the primary tool for conducting distributional analysis is a microsimulation model. A microsimulation model is essentially a set of relationships between key variables which allows the user to estimate the effects of a policy on households, individuals, or other 'micro units'. In order to isolate the effects of the policy from those changes which would have occurred in the absence of the policy, microsimulation models typically compare a 'baseline' scenario (which excludes the policy) against a 'counterfactual' scenario (which includes the policy). Microsimulation models can be either 'static', meaning that relationships between variables do not change when a policy is introduced, or 'dynamic', meaning that those relationships can change as households are predicted to alter their behaviour.

Distributional analysis is most commonly used to estimate the effects of direct taxes[1] and benefits on people with different levels of income[2]. The BPRG's recommendation for the

[1] Direct taxes are taxes levied on individuals or their property, such as income tax. Indirect taxes, on the other hand, are taxes levied on transactions, such as VAT.

[2] These levels are usually defined by income deciles, which are delineated by ranking the population by income and then splitting it into ten groups of equal size. The first decile therefore contains the poorest 10% of households, the second decile contains the next-poorest 10%, and so on.

Scottish budget is to explore the feasibility of a broader methodology which additionally includes indirect taxes and public services and which examines the effects on different protected characteristics. As discussed below, such an analysis would be highly valuable, but is significantly more demanding in terms of data requirements and modelling capacity.

Models

A number of microsimulation models exist which provide distributional analysis, some of which have been used in a Scottish context. Table 2 (Appendix) summarises the relevant details of these models as discussed in this section.

Models Used for Scotland

In November 2017, the Scottish Government published 'the Role of Income Tax in Scotland's Budget' which provided distributional analysis, by income, of various income tax proposals on individuals[3]. Impacts on individuals were estimated using the Scottish Government's **Income Tax Model**, which is based on ONS population forecasts, macroeconomic forecasts, and various data sources[4]. The Model is static, although behavioural responses are added in a second stage by applying taxable income elasticities. This model was used to perform distributional analysis of changes to income tax policy on age and gender groups[5].

Impacts on households, on the other hand, were estimated using the Scottish Government's **Tax and Welfare Model**, which is a static model based on the Family Resources Survey (FRS)[6]. Although the publication focused on income tax, and only considered impacts by income category, the Tax and Welfare Model is capable of including benefits and can be adapted to consider the protected characteristics included in the FRS, which are discussed in the final section. Indeed, the model was used to perform distributional analysis of changes to income tax policy for households by disability status and income. The model does not incorporate indirect taxes or public services, but could potentially be extended to do so with varying degrees of difficulty; VAT, for example, would be easier to incorporate than corporation tax.

---

[3] Scottish Government, 2017, 'The Role of Income Tax in Scotland's Budget', available at http://www.gov.scot/Publications/2017/11/5307.

[4] Namely the Survey of Personal Incomes (SPI), the Annual Survey of Hours and Earnings (ASHE), and the Annual Population Survey (APS).

[5] Scottish Government, 2017, 'The Scottish Government's Income Tax Policy (Stage 1): Analytical Note on Impacts on Income Levels and Equality', available at http://www.gov.scot/Topics/Government/Finance/scottishapproach/Scottishincometax2018-2019/documents.

[6] The Tax and Welfare Model has also been used alongside **eTABOSH** (Tax and Benefits of Scottish Households), which provides a 'case study' perspective on the impacts of tax and benefit changes by different household types but does not have the capacity to provide full-scale distributional analysis.

A number of publications – including work commissioned by the Scottish Government to forecast child poverty[7] and work commissioned by the Equality and Human Rights Commission (EHRC) to analyse tax and welfare reforms[8] – have used the **Tax Transfer Model** developed by Landman Economics for the Institute for Public Policy Research (IPPR). Based on FRS and LCF data, the Tax Transfer Model is static, but considers multiple protected characteristics, provides Scotland-specific estimates, and includes tax, benefits, and public services.

Researchers from **Sheffield Hallam** University have estimated the distributive impact of UK Government welfare reforms by local authority area and household type across Great Britain[9]. This analysis does not involve a micro-simulation model per se; rather it allocates the financial savings of welfare reforms reported by the Treasury on the basis of the claimant data reported by DWP and HMRC, in addition to information contained in Equality Impact Assessments and pilot schemes. While this approach could possibly be extended to include direct taxes, it does not appear suitable for analysing the impacts of indirect taxes or public services.

Other UK Models

HMRC publishes a distributional analysis by income to accompany the UK budget, based on HMT's **Intra-Government Tax and Benefit Model**[10] (IGOTM). IGOTM is based on the Living Costs and Food (LCF) Survey, supplemented with FRS and ASHE data. Like the Tax Transfer Model, IGOTM incorporates direct and indirect taxes, benefits, and public services; and like the Income Tax Simulation Model, behavioural responses can be added to what is otherwise a static model. The Scottish Government currently possesses a version of IGOTM, but due to the limited size of the Scottish LCF sample it is generally considered to be too unreliable for producing official advice. In addition, this particular version does not include public services.

DWP's **Policy Simulation Model (PSM)** shares many characteristics with the Scottish Government's Tax and Welfare Model – it is a static model based on FRS data which models direct tax and benefits. However, PSM contains certain features which are absent from the Tax and Welfare Model: it can model the transition from legacy benefits to Universal Credit,

---

[7] Reed, Howard, 2018, 'Forecasting Child Poverty in Scotland', Scottish Government, available at http://www.gov.scot/Publications/2018/03/2911/downloads.

[8] Portes, Jonathan and Reed, Howard, 2018, 'The Cumulative Impact of Tax and Welfare Reforms', Equality and Human Rights Commission, available at https://www.equalityhumanrights.com/sites/default/files/cumulative-impact-assessment-report.pdf

[9] Beatty, Christina and Fothergill, Steve, 2016, 'The Impact on Scotland of the New Welfare Reforms', Scottish Parliament Social Security Committee, available at http://shura.shu.ac.uk/15885/1/impact-scotland-new-welfare-reform.pdf.

[10] HM Treasury, 2017, 'Impact on Households: Distributional Analysis to Accompany Autumn Budget 2017', available at https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/661465/distributional_analysis_autumn_budget_2017.pdf

rather than just a legacy-only or UC-only world; it adjusts for incomplete take-up of benefits, rather than assuming complete take-up; and can uprate incomes for future years, rather than relying on incomes as they are recorded in the FRS. The Scottish Government has access to a Scotland-only version of the PSM, which it uses internally for forecasting benefit caseloads. However, the model is constrained in terms of usage and publication of findings, so the Tax and Welfare Model is generally preferred when possible.

Euromod, owned by the Institute for Social and Economic Research (ISER), also shares the key features of the Tax and Welfare Model – static, based on FRS data, models direct tax and benefits – but for a number of reasons is generally considered less useful than either the Policy Simulation Model or the Tax and Welfare Model for the purposes of the Scottish Government, which holds an older version of the model. For example, the analysis lags behind policy changes and uses a coding language that is unfamiliar to most analysts in the Scottish Government.

The IFS's **TAXBEN** model, based on FRS and LCF data, is dynamic but does not include public services. Conversely, the **ONS's microsimulation model**, based on LCF data, is static but includes public services. The methodology for apportioning in-kind benefits within this model has recently been improved using FRS data to take account of the effects of income, tenure, and other socio-economic variables on service use. The Scottish Government does not currently hold either of these models and it is not clear whether they could be adapted to provide Scottish estimates.

Two models were commissioned by the Joseph Rowntree Foundation (JRF) to estimate the effects of policies on poverty. **The Lifetime Income Distributional Analysis Model** (LINDA), developed by the National Institute for Economic and Social Research (NIESR), is a dynamic model based on the Wealth & Asset Survey (WAS) with input from other data sources[11]. LINDA includes direct and indirect taxes and benefits but not public services. The second model, developed at **Herriot-Watt University**, links a dynamic macro-simulation with a static micro-simulation based on Understanding Society (UKHLS) data, but is primarily geared to examining issues of housing and is more ad hoc in its treatment of tax and benefits. The Scottish Government does not hold either of these models, which are both constructed at the UK level; the Herriot-Watt model could be adapted to produce Scottish estimates, but the potential to adapt LINDA is unclear.

Data Sources

Income-based Sources

The two most relevant data sources for distributional analysis in a Scottish context are the Family Resources Survey and the Living Costs and Food Survey. These data sources are used in many of the models outlined above.

[11] Namely the FRS, the LCF, the ASHE, the Labour Force Survey (LFS), and the British Household Panel Survey (BHPS).

The **Family Resources Survey** (FRS) is an annual repeated cross-sectional survey[12] focusing on income, housing, and pensions. The FRS is widely considered to be the best source of data on individual, family, and household incomes in the UK and is accordingly used to produce the DWP's dataset on Households Below Average Income, which provide the official estimates of poverty and income inequality. Thanks to an ongoing boost funded by the Scottish Government, the Scottish sample is overrepresented in the data, with around 2,700 (15%) of the 20,000 total households located in Scotland. In terms of protected characteristics, the FRS includes data on age, disability, marriage and civil partnership, maternity, race, and sex, but not gender reassignment or pregnancy. The sample sizes for a selection of these characteristics, pooled for the years 2014-2017, are shown in Table 1 (below).

Table 1: Sample Sizes for Selected Protected Characteristics, FRS, 2014-2017

Gender – single adults	**Men**	**Women**
Single Pensioner	508	1,054
Single without dependent children	1,576	1,344
Single with dependent children	46	539
All single adults	2,130	2,937
Ethnicity	**People**	
White – British	8,926	
White – Other	400	
Asian or Asian British	154	
Mixed, Black or Black British and Other	116	
Disability	**People**	
Disabled person in the family	3,605	
No disabled person in the family	5,991	

Like the FRS, the **Living Costs and Food Survey** (LCF) is an annual repeated cross-sectional survey focusing on gross and disposable incomes of individuals, families, and households in the UK. Although the income data is less detailed than the FRS, the LCF includes additional data on expenditure of goods and services at the household level. At around 5,000 households, the LCF's sample size is approximately one-quarter that of the FRS, with only 420 households in Scotland, although this will increase with an upcoming boost. The LCF includes the same protected characteristics as the FRS minus disability and with a less detailed race classification.

Other Sources

[12] Unlike a panel survey, which follows a cohort of households over time, a repeated cross-sectional survey interviews a new set of households each year.

In addition to the FRS and the LCF, two data sources are available which are less directly focused on income and expenditure but contain larger sample sizes for Scotland: the Scottish Household Survey and Understanding Society.

The **Scottish Household Survey** (SHS) is a biennial repeated cross-sectional survey covering a range of topics in addition to income, including demographics, housing, and transport, as well as health, employment, and education. The sample consists of around 10,000 Scottish households, with classifications for age, sex, religion, disability, race, and marriage/civil partnership. There is currently an intention to improve the income-related questions in the SHS.

The UK Household and Longitudinal Survey (UKHLS or '**Understanding Society**') is an annual longitudinal survey covering a range of topics including income, health, employment, education, social life, and education. The sample consists of around 40,000 households in the UK with around 3,300 in Scotland, containing data on age, sex, religion, disability, pregnancy, race, and marriage/civil partnership.

Discussion

Three main issues emerge from the BPRG's recommendation:

1. which aspects of government activity can be modelled (direct tax, indirect tax, benefits, public services);
2. which protected characteristics can be modelled (age, disability, gender reassignment, marriage and civil partnership, pregnancy and maternity, race, religion or belief, sex, sexual orientation);
3. whether Scotland can be modelled separately.

Tax, Benefits, and Public Services

A methodology which models the distributional effects of all taxes, benefits, and public services can be used to estimate the total value that each micro-unit gains from government policies (the 'social wage'). This 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. For example, a distributional analysis conducted by the Institute for Fiscal Studies (IFS) indicated that the Labour Party's 2017 manifesto was regressive, i.e. that it would place a larger burden on poorer households[13]. However, the analysis only considered personal tax and benefits; if it had included other

[13] IFS, 2017, 'General election analysis 2017', available at https://www.ifs.org.uk/uploads/Presentations/Rob%20Joyce%2C%202017%20General%20Election%2C%20manifesto%20analysis.pdf.

taxes (e.g. corporation tax) and public spending, the results would likely have been inverted[14].

Despite their advantages, distributional analyses which include indirect taxes and public services are relatively uncommon because they involve linking household characteristics to spending behaviour and service use, which requires not only additional data sources and modelling capabilities but also interpretive care. For example, if poorer households use health services more than richer households, this could be interpreted either as a correlation between low income and health problems or as a correlation between high income and usage of private healthcare. Although this distinction is significant when it comes to designing policy, it will not be captured in a distributional analysis.

Protected Characteristics

While distributional analysis is most commonly performed by income level, other distinctions, including protected characteristics, are also relevant. In the case of gender, economists have demonstrated that budgetary items often have unequal effects on men and women, and that distributional analysis can be used to design effective, gender-positive policies. In Andalusia, for example, distributional analysis by gender has prompted an expansion of after-school activities for children and a top-up in the state pension[15].

The main reason that distributional analysis by protected characteristic is relatively uncommon is that the requisite data is limited, particularly in the case of gender reassignment, sexual orientation, and pregnancy/maternity, and to a lesser extent disability, race, and religion. Even when surveys record these characteristics, the samples are often too small to be meaningful, particularly when it comes to producing a disaggregated analysis for Scotland. In the case of gender, meanwhile, distributional analysis can be hindered by the fact certain data, such as receipt of certain benefits, only exist at the household- or family level and therefore fail to reveal the distribution of resources between men and women.

Disaggregation for Scotland

The Scottish Government's fiscal powers will be most effectively exercised when its policies are tailored to Scotland's distinct demography and economy. Even the most powerful microsimulation model will therefore be inappropriate if it is unable to produce

---

[14] IPPR, 2017, 'The perils of distributional analysis', available at https://www.ippr.org/blog/the-perils-of-distributional-analysis-was-labour-s-manifesto-really-regressive. The IFS analysis also included Universal Credit in the counterfactual rather than the baseline, effectively amalgamating its effects with those of Labour's policies.

[15] Junta de Andalucia, 2012, 'Gender impact assessment report on the budget of the Autonomous Community of Andalusia for 2012', available at https://www.juntadeandalucia.es/economiayhacienda/haciendayadministracionpublica/planif_presup/genero/informe/informe2012/informe_ingles.pdf.

disaggregated analysis for Scotland as opposed to the UK as a whole. In addition, the model must be capable of isolating the impacts of Scottish Government policies from those of UK Government policies. The Scottish Government must also have the capacity to run and update the model.

As intimated above, sample size is often the primary constraint to producing a Scotland-specific analysis – particularly when it comes to protected characteristics, each of which represents a fraction of the Scottish sample. Among the data sources discussed above, the issue of sample size is most acute in the LCF, which only contains around 400 Scottish households. There are methods for addressing this issue, for example by aggregating multiple years of data, combining different data sources, or imputing missing data[16], but these methods introduce their own complications and ultimately there is only so much that can be done with limited data.

Conclusion

This paper has explored the feasibility of producing a distributional analysis for the Scottish budget, including taxes, benefits, and public-service expenditure, by equality characteristic. Such an analysis, while possible in the main, is not without its challenges.

The FRS has formed the basis for distributional analyses in Scotland so far and this will most likely continue to be the case. Not only is the FRS regarded as the best source of income data in the UK, it also records most of the protected characteristics and contains a boosted sample for Scotland. In addition, incorporating public services will most likely require the LFS, possibly augmented with other data sources such as the SHS and Understanding Society. Neither the FRS nor the LCF include gender reassignment or pregnancy, and Scottish sample sizes for some characteristics may be insufficient, but alternative data sources are unlikely to provide superior granularity.

Landman's Tax Transfer Model demonstrates that the FRS, combined with LCF data, can provide a Scotland-specific methodology, covering tax, benefits, and public services, for most protected characteristics. The Tax Transfer Model does encounter some sample-size issues, e.g. for ethnicity, and is unable to cover those characteristics excluded from the data. However, most of the other models discussed in this paper only examine the income dimension, and moreover have not been used to produce disaggregated analyses for Scotland. The Scottish Government's in-house Tax and Welfare Model is Scotland-specific and can be readily expanded to cover the same characteristics as the Tax Transfer Model, but expanding it to incorporate indirect taxes and public services would be more challenging.

In sum, distributional analysis as per the BPRG's recommendation would be possible, excepting breakdowns by certain equality characteristics. If the Scottish Government were to pursue such an analysis, it could use Landman's Tax Transfer Model (whether by

[16] Imputation is a way of preserving observations that contain missing values and thus maximising the effective sample size.

purchasing the model or by commissioning Landman Economics), seek to expand its own Tax and Welfare Model (both to cover protected characteristics and to include indirect taxes and public services), or commission the development of a new model.

Appendix

Table 2: Microsimulation Models

Model	Static or Dynamic	Coverage	Data Sources	Possessed by Scottish Government?	Capable of Scotland-specific analysis?	Comments
Income Tax Model	Static with behavioural add-on	Income tax	SPI, ASHE, APS	Yes – in-house model	Yes	Used for SG publications
Tax and Welfare Model	Static	Direct tax and benefits	FRS	Yes – in-house model	Yes	Used for SG publications
Tax Transfer Model	Static	Direct and indirect tax, benefits, and public services	FRS, LCF	No – owned by Landman Economics	Yes	Used for recent SG publications
Sheffield Hallam analysis	N/A	Benefits	Treasury, DWP, HMRC	No	Yes	Not a microsimulation model per se
Intra-Government Tax and Benefit Model (IGOTM)	Static with behavioural add-on	Direct and indirect tax, benefits, and public services	LCF, FRS, ASHE	Partially – version acquired from HMT	Unreliable	Version possessed by SG does not include public services
Policy Simulation Model (PSM)	Static	Direct tax and benefits	FRS	Partially – version acquired from DWP	Yes	Version possessed by SG has restricted use
Euromod	Static	Direct tax and benefits	FRS	Partially – version acquired from ISER	Yes	Version possessed by SG is outdated
TAXBEN	Dynamic	Direct and indirect tax and benefits	FRS, LCF	No	Uncertain	

ONS Microsimulation Model	Static	Direct and indirect tax, benefits, and public services	LCF	No	Uncertain	
Lifetime Income Distributional Analysis Model (LINDA)	Dynamic	Direct and indirect tax and benefits	WAS, FRS, LCF, ASHE, LFS, BHPS	No	Uncertain	
Heriot-Watt Housing Model	Dynamic	Direct tax and benefits	UKHLS	No	Potential	Primarily geared to housing issues