

Adjusting Scotland's Block Grant for new Tax and Welfare Powers: Assessing the Options

David Bell

University of Stirling

David Eiser

University of Stirling

David Phillips

Institute for Fiscal Studies

Institute for Fiscal Studies
7 Ridgmount Street
London WC1E 7AE

Published by

The Institute for Fiscal Studies
7 Ridgmount Street
London WC1E 7AE

Tel: +44 (0) 20-7291 4800

Fax: +44 (0) 20-7323 4780

Email: mailbox@ifs.org.uk

Website: <http://www.ifs.org.uk>

© The Institute for Fiscal Studies, November 2015

Preface

This paper was supported by funding from the Nuffield Foundation. The Nuffield Foundation is an endowed charitable trust that aims to improve social well-being in the widest sense. It funds research and innovation in education and social policy and also works to build capacity in education, science and social science research. The Nuffield Foundation has funded this project, but the views expressed are those of the authors and not necessarily those of the Foundation. More information is available at www.nuffieldfoundation.org

The paper was also supported by funding by the Economic and Social Research Council (ESRC) through the Centre for the Microeconomic Analysis of Public Policy at IFS (grant reference ES/H021221/1).

The authors thank Paul Johnson for comments on a draft version of the report and numerous colleagues for helpful discussions on the issues in this paper. Any errors or omissions are the responsibility of the authors.

Contents

Executive Summary	1
1. Introduction	8
2. Existing funding arrangements for the UK's devolved governments	11
3. Proposals for further tax and welfare devolution to Scotland	15
3.1 Planned devolved tax and welfare powers	15
3.2 Smith's principles for a new fiscal framework	17
4. Assessing options for adjusting Scotland's block grant	20
4.1 When population grows at the same rate as in the rest of the UK	21
4.2 When population growth differs in Scotland	25
4.3 Illustrating using data from 1999-00 to 2013-14	29
5. Accounting for changes to tax rates in the rest of the UK	32
6. Additional considerations for welfare devolution	36
7. Difficulties with Smith's 'compensation principle'	40
8. Broader discussion and conclusions	42
References	46

Executive Summary

The Scotland Bill, currently making its way through the Houses of Parliament, will transfer a range of tax and spending powers from Westminster to the Scottish Parliament. At the same time, an adjustment will have to be made to Scotland's block grant funding from Westminster. Alongside things like changes to borrowing powers and fiscal institutions, these block grant adjustments (BGAs) form a key part of the new "fiscal framework" Scotland will require when these powers are transferred.

Unlike the tax and welfare powers, the fiscal framework is not part of the Scotland Bill. The report of the Smith Commission, on which the Bill is based, did not have the time (or perhaps the expertise) to design the fiscal framework. Instead, it laid down a number of principles for its design. This paper focuses on the effects of different options for determining the BGAs, and how these fit in with the Smith Commission's design principles. It finds that the precise way in which the BGAs are indexed over time could mean differences of over a billion pounds a year in the Scottish Government's budget after a relatively short period of time. Its analysis suggests that the Smith Commission's principles may not be workable and are not mutually compatible. It also argues that there would be real merit in a more fundamental reassessment of devolved government finance – including the Barnett Formula. A more radical reform might help address some of the difficulties this paper identifies.

The existing system of funding Scotland

- Because they currently raise relatively little revenue directly themselves, the devolved governments of Scotland, Wales and Northern Ireland rely on 'block grant' transfers from the UK government to fund spending on devolved public services. This means that the devolved governments are largely insulated from changes in both the relative and absolute levels of tax revenues in their countries (except for the relatively small taxes currently devolved). This provides a greater degree of certainty and stability in funding, but means devolved governments have less financial incentive to implement policies that boost revenues, and are subject to weaker fiscal accountability for the decisions they do make.
- The changes in these block grants is calculated by the Barnett Formula which allocates each devolved government a population share of changes in spending on 'comparable' (i.e. devolved) services in England. Because the Barnett Formula takes no account of spending needs, devolved governments' budgets are exposed to risks associated with factors that affect relative spending needs, such as differential rates of population ageing, increases in the prevalence of illness, or changes in other socio-economic factors.
- Moreover, although the Barnett Formula allocates a population share of the change in public spending in England to Scotland and the other devolved

nations, the budgets of the devolved governments still face risks and incentives associated with differential population growth. This is because the Barnett Formula only determines how the block grant changes, and while these changes reflect the latest population shares, the existing block grant (the baseline) is not adjusted to account for changes in population. Devolved governments therefore benefit from more to spend per person when their population is growing less quickly than England's, and vice versa.

Proposals for further tax and welfare devolution to Scotland

- Following the Scottish independence referendum, the Smith Commission recommended a package of substantial further tax and welfare devolution. The aim of this was to strengthen the Scottish Parliament's ability to 'pursue its own vision, goals and objectives', and to increase the parliament's 'accountability and responsibility for the effects of its decisions and their resulting benefits and costs'.
- The Smith Commission committed to retain the Barnett Formula as the mechanism for determining Scotland's block grant. But Scotland's Barnett-determined block grant will clearly need to be adjusted to reflect both the new tax-raising powers and new expenditure responsibilities being devolved. In the first year that new powers are devolved, this block grant adjustment (BGA) should be relatively straightforward, at least in principle. The initial reduction in the block grant for devolved taxes should be equivalent to the revenue forgone by the UK Government. Similarly, for welfare.
- In future years however the BGAs have to be indexed to account for inflation and likely economic growth. Otherwise they will almost certainly fall in value relative to the taxes or spending being devolved. The Smith Commission recognised this and stated that the BGAs should be 'indexed appropriately'. But it stopped short of saying how this indexation should be made. It did however set out a set of principles that the revised fiscal framework, including the BGA, should satisfy.
- These include the so-called 'No Detriment' principles. The first states that neither government should lose out solely as a result of the 'decision to devolve'. The second says that neither government should lose out as a result of the policy decision of the other in a devolved area, post-devolution. This has two specific elements. First, that if such knock-on effects occur, compensatory payments should be made between governments. And second, that changes in taxes in the rest of the UK (rUK) which are devolved to Scotland should not affect overall public spending in Scotland (the 'taxpayer fairness' principle).
- These and the other principles have implications for the type of BGA that can be used in Scotland, while meeting the spirit of the Smith Commission.

The block grant adjustment options

- We consider three options where the BGAs are linked in some way to what happens to equivalent tax revenues (or welfare spending) in rUK.
- The first approach is **Indexed Deduction (ID)**. This indexes the change in the BGA to the percentage change in total comparable tax revenues in the rest of the UK (rUK). For example, if comparable revenues in rUK grow by 5%, the BGA also grows by 5%. Under the ID adjustment method, growth in either the number of people or revenue per person in rUK would lead to an increase in Scotland's BGA. This approach therefore exposes Scotland to the risk of relatively slower population growth than in rUK. On the other hand, this mechanism would allow Scotland to capture the reward of relatively faster population growth. Scotland would therefore gain from attracting and retaining more income tax payers, for instance.
- The second approach is **Per Capita Indexed Deduction (PCID)**. This indexes the BGA per capita to the percentage change in comparable rUK revenues *per person*. This option clearly protects the Scottish budget from the risk that its population grows relatively slower than the rUK's. But equally the Scottish budget would not benefit from revenue increases that resulted from population growth. The Scottish Government would therefore lack incentives to boost growth through attracting more people to Scotland.
- The third approach is the **Levels Deduction (LD)**. This calculates the change in the BGA as a population share of the change in comparable revenues in rUK. For example, if income tax revenues increased by £10 billion in rUK, then if Scotland's population was 9% of rUK, Scotland's BGA would increase by £900m. The rationale for the LD approach is that, by being based on a population share of a cash terms change in revenue, it is symmetric with the spending side of the Barnett Formula (which calculates the change to Scotland's block grant as a population share of the cash terms change in English spending).

The effects of these options in different circumstances

- This symmetry property is useful when it comes to changes in rUK tax rates. Changes in UK tax rates for taxes that are devolved are likely to lead to a change in spending by the UK Government. To the extent that this spending is likely to benefit Scottish taxpayers in some way (either because it leads to an increase in the Scottish block grant via the Barnett formula, or because it leads to an increase in 'reserved' spending in Scotland), the block grant to Scotland would need to be adjusted to ensure that increases in taxes in rUK tax do not fund higher spending in Scotland, without a corresponding increase in Scotland's tax effort. This is the 'taxpayer fairness' element of the Smith Commission's 'no detriment' principles.
- When the additional rUK revenues are spent on services like health or education that are devolved to Scotland, Scotland gets an equivalent

population share of this spending via the Barnett Formula. An increase in taxes and spending of £10 billion in rUK would, for instance, feed through into approximately £900 million more for Scotland via the Barnett Formula. Under the **LD** approach, Scotland's BGA also increases by a population share of the change in UK revenues: again £900 million. So the increase in the BGA exactly offsets the increase in the underlying block grant, leaving Scotland unaffected. The 'taxpayer fairness' principle is satisfied.

- In contrast, under the **ID** or **PCID** approaches, Scotland would gain from such a tax increase. This is because tax revenues per person are lower in Scotland than in rUK. A percentage increase in the BGA is therefore smaller than a population-share based increase. Our report shows that for a £10 billion income tax increase in rUK, equal to about 2p on each income tax rate, such gains to Scotland would amount to over £100 million a year, even though Scots were paying no more tax themselves. Or vice versa for an income tax cut. These methods therefore do not fully satisfy the 'taxpayer fairness' principle.
- Tax revenues change not only because of policy changes though. Underlying growth in the economy and the tax base also affect revenues. Depending on the initial starting levels of revenues per person, revenue growth per person, and population growth, the different options we consider can have markedly different effects of the Scottish Government's budget.
- Consider the following situation: revenues start off lower per person in Scotland, grow at the same percentage rate per person as in rUK, but the population grows less quickly than in rUK. Under the **PCID** approach, the BGA increases in line with the rate of growth in revenues per person, which is the same in Scotland and rUK. Hence Scottish revenues grow at the same rate as the BGA meaning it does no better or no worse than if taxes were not devolved. This method might therefore be seen to satisfy the spirit of the principle that there should be 'no detriment from the decision to devolve'.
- On the other hand, Scotland does lose out somewhat under **ID** approach because of its lower population growth. And, under the **LD** approach it loses out even more, at least initially, because a given rate of growth in its revenues translates into less than a population-based share of the equivalent growth in rUK revenues (because rUK revenues per person started off higher). The **LD** approach therefore does not seem to satisfy the spirit of the principle that there should be 'no detriment from the decision to devolve': there will be detriment to Scotland under this approach, unless revenues in Scotland grow at a faster rate both per person, and in aggregate than in rUK. This might be seen as an unfair challenge for Scotland to meet.
- It therefore turns out that it is **impossible to design a block grant adjustment system that satisfies the spirit of the 'no detriment from the decision to devolve' principle at the same time as fully achieving the**

‘taxpayer fairness’ principle: at least while the Barnett Formula remains in place.

- What method is used to index the BGA can have a significant effect on the Scottish Government’s budget. To see this one can examine what would have happened if these methods had been in place between 1999–00 and 2013–14. While these figures are approximate and refer to the past rather than the future, they show these differences can be substantial. Relative to the **LD** method, the **ID** method could have resulted in the Scottish Government’s budget being around £1 billion higher a year after 14 years, with the **PCID** approach delivering an even bigger budget. These are quite sizeable numbers in the context of a block grant to Scotland equal to around £30 billion a year in 2013–14.
- Our analysis also shows that eventually though, if relative population decline continues, Scotland would start to do less well under the **ID** method than the **LD** method. Indeed, because the **ID** method *never* gets updated to reflect the fall in Scotland’s relative population, it can eventually imply a *negative* budget for Scotland if one far enough in the future. Clearly such an outcome would never be allowed to come to pass. But it illustrates that the **ID** method would not represent a sustainable long-term compromise between the **PCID** and **LD** methods that, in the short term, would be most beneficial to the Scottish Government and UK Treasury, respectively.

The difficulties of compensating for policy knock-on effects

- The ‘compensation principle’ set out by the Smith Commission will be impractical to fully implement. This is because the way that individuals change their behaviour in response to tax rate changes means that the counterfactual “no change” scenario cannot be observed or easily modelled. Calculating the knock on effects would therefore require a series of assumptions about behaviour, each subject to significant uncertainty, opening up the potential for frequent disagreement between the governments. Indeed, if too much attention is paid to compensating for every example of knock on effect, then the arguments and tricky negotiations that result could cause the whole system to become unworkable and unsustainable.
- These issues are likely to be particularly relevant for welfare devolution where there are complicated interactions between different parts of the benefit system, and significant scope for behavioural response to changes to benefit structure and rules.

The potential benefits of a more fundamental debate on devolution

- There is a lack of consensus, or even debate, about the type of fiscal risks and incentives Scotland (and other devolved governments) should face. Should Scotland face risks associated with relative population change? Should it face risks associated with differential demographic or economic change, over

which the Scottish Government might have only limited control. Without such a debate it is difficult to recommend a particular form of BGA as these are intimately linked to the type of fiscal risks and incentives a Scottish Government will face.

- The Smith Commission says that the UK government should bear 'economic responsibility' for its own policies, and the risk of any shocks that affect the whole of the UK, and the Scottish Government should bear the 'economic responsibility' for its policies, it says nothing about who should bear the underlying revenue or spending risks in Scotland. The methods we outline – which are perhaps the most widely discussed methods –, would mean that the Scottish Government bore all such risks associated with devolved revenues and spending. But one might want to pool at least some of these risks across the UK's fiscal union.
- The key difficulty in doing this though is isolating the effect of government policies – the consequences of which, as far as possible, you would want the government who made them to face – and other underlying risks that you might want to pool. Other countries – such as Canada, Germany and Australia – do this by trying to assess the spending needs or the revenue-raising capacity of different parts of their country, and then calculating transfers to offset some or all of the differences. A similar approach was traditionally used for local government grants in the UK. Implementing such an approach for the financing of devolved governments would require more significant changes to the fiscal framework than the Smith Commission advocated though.
- More fundamental reform could help address some of the issues and conflicts the Smith Commission principles might currently involve. For instance, the Barnett Formula could be reformed so that it allocated Scotland and the other devolved nations the same percentage change in spending per person as on comparable spending in England. This would make the formula on the spending side consistent with the **PCID** approach to indexing the BGA. It would therefore be possible to satisfy both the 'no detriment' and 'taxpayer fairness' principles.
- Replacing the Barnett Formula with a formula akin to the **PCID** approach would stop further convergence if Scotland's revenues per person kept pace with those in rUK. This **Percentage Per Capita (PPC) Formula** could be implemented with or without an initial needs or revenue capacity assessment to determine the appropriate level of block grant to give to Scotland and the other devolved nations.
- It is important to note that if, after an initial assessment, ongoing needs or revenue capacity based assessments were used to update the underlying block grant over time, that this would necessarily blunt the fiscal incentives that tax and welfare devolution otherwise entails. Recent devolution debates, at least in the case of Scotland, have been almost all about providing stronger

fiscal incentives. Moves towards *ongoing* needs assessments would therefore be a move in the opposite direction.

- That is not say such a move would be wrong. That depends on ones view about what the UK's fiscal union is for. What risks should be pooled and what ones should not. And where the balance between incentives and insurance in devolved government finance should be. There would be real merit in a proper debate about these issues.
- At the very least, because they are key to Scotland's future budget, and the incentives and risks it will face, the options available for calculating the block grant adjustments, and other elements of the fiscal framework, should be part of the public and parliamentary debate, as much as the tax and welfare powers set out in the Scotland Bill itself have been.

1. Introduction

The UK has traditionally been one of the more fiscally centralised countries in the OECD. Although the three devolved legislatures (in Scotland, Wales and Northern Ireland) have had substantial spending autonomy, tax devolution has been limited to control over relatively small recurrent property taxes. Instead, devolved spending has been funded almost entirely by block grants from the UK government, paid for by revenues pooled across the whole of the UK.

But further devolution is in the vanguard of UK fiscal policy. In particular, following the “No” vote in the Scottish independence referendum, the Smith Commission was set up to recommend a package of additional tax and welfare responsibilities to be devolved to the Scottish Government, on top of those already legislated for in the Scotland Act 2012.¹ The resulting package, including full rate- and threshold-setting powers for income tax on non-savings non-dividend income, and the assignment of half of VAT revenues, will increase the proportion of the Scottish government’s budget funded by locally raised revenue to close to 50%.² It also involves the full devolution of a number of areas of welfare spending, especially in relation to disability benefits, and the power to create new benefits and benefit top-ups in other areas of the welfare system. Legislation to introduce these new powers is currently going through the UK parliament.

The UK government also plans to partially devolve income tax to the Welsh Assembly if the Welsh people approve such powers in a referendum, and has agreed to devolve corporation tax to the Northern Ireland Assembly. And within England, the UK government has announced the devolution of business property taxes (non-domestic rates, NDR) to local authorities, including unlimited power to reduce rates and constrained powers to increase rates.³

Although the economic and political circumstances and specifics are different, in each case a key motivation for further devolution is to enhance the financial autonomy and accountability of sub-national government (SNG). Tax devolution enables SNGs to retain at least part of the proceeds of growth, providing financial incentives to pursue policies which grow the tax base and economy. It also enables SNGs to vary their budgets at the margin, by setting higher or lower tax rates. The power to change the structure and level of taxes and welfare also allows SNGs to more closely match the preference of the populations of their respective areas.

¹ Stamp Duty Land Tax, Landfill Tax, and 10 percentage points of each income tax band on non-savings non-dividends income.

² Smith Commission (2014).

³ Those authorities or groups of authorities with an elected mayor will have the ability to increase NDR rates by up to 2p in the pound above the notional national “uniform business rate” if a major of business members of Local Enterprise Partnerships agree. See <https://www.gov.uk/government/news/chancellor-unveils-devolution-revolution>.

Inevitably, however, there is a trade-off between the provision of financial incentives for growth, and insurance against short- and long-term budgetary risk. Whilst fiscal decentralisation may enhance autonomy and growth incentives, it may expose SNGs to greater fluctuations in annual budgets. Perhaps more significantly, it means that differential economic performance may result in growing disparities between SNGs, that are due at least in part to circumstances beyond the direct influence of the SNG itself.

Such a substantial package of revenue and spending devolution as proposed for Scotland therefore clearly requires changes to the broader fiscal framework under which both devolved and central government operates. This was recognised by the Smith Commission, which set out a number of principles that it felt the new fiscal framework should satisfy. This includes powers and rules over borrowing, and changes to fiscal institutions for forecasting revenues and spending, and potentially mediating in fiscal disputes between the Scottish and UK governments.

Perhaps most obviously, Scotland's annual allocation from the UK Government – the “block grant” - will have to be adjusted to reflect both the new tax revenues and new expenditure responsibilities being devolved. In the first year that new powers are devolved, these block grant adjustments (BGAs) should be relatively straightforward, at least in principle. The initial reduction in the block grant to account for devolved taxes should be equivalent to a forecast of the amount of revenue devolved (and therefore forgone by the UK government). Likewise the devolution of welfare spending will need to be accompanied by an addition to the block grant equivalent to the forecast level of spending in Scotland by the UK government in that year had the powers not been devolved.

In future years however the BGAs have to be indexed. The Smith Commission stated that the BGAs should be ‘indexed appropriately’, but stopped short of saying how this indexation should be made, even though this mechanism is central to the reallocation of risk and reward that provides a central justification for this fiscal decentralisation. This paper investigates the properties and implications of a number of indexation methods in detail. The aim is to inform both public debate and the policymaking process: the Scottish and UK governments are currently negotiating the new fiscal framework, including the method for indexing the BGA.

The methods we assess have one thing in common. In each, the BGA is linked in some way to what happens to equivalent revenues and spending in the rest of the UK. This is because, as we discuss in this paper, such an approach is a relatively straightforward way of ensuring that the UK government continues to insure the Scottish Government against revenue or welfare spending shocks that hit the whole of the UK. This is one of the principles the Smith Commission says the new fiscal framework should satisfy. However, we show that the precise way in which this link works could lead to differences in the Scottish Government's budget of more than a billion pounds a year, after little more than a decade in operation.

We also find that it is not possible to index the BGA in such a way that it fully satisfies the various “no detriment” principles set out by the Smith Commission. In fact, the principles appear to be somewhat in conflict with each other in practise, even though they seem sensible and consistent in the abstract. Furthermore, if applied in all circumstances, the idea that compensatory transfers should be made whenever “detriment” occurs to one government’s finances as a result of the other’s actions could quickly lead the system of devolved finance to become unworkable and mired in intergovernmental conflict. Some principles about what constitutes a significant and material detriment that requires compensation should be agreed in advance.

We also argue that without a more refined set of principles about the specific fiscal risks and incentives that SNGs and the UK government should face, or on the appropriate levels of relative funding for different parts of the UK, that there is no objective way to choose between the different ways of indexing the BGA. Without such principles it is even more likely that the negotiation positions of the Scottish and UK governments will simply be to maximise or minimise the likely grant, respectively. This does not seem conducive to long-term stability of the system.

The rest of this paper proceeds as follows. Section 2 describes the existing system of devolved government finance of block grants, and the risks and incentives it entails. Section 3 describes the Smith Commission proposals for the devolution of further tax and welfare powers to Scotland, and the set of principles the Commission said the associated new fiscal framework should satisfy. Section 4 assesses how different options for adjusting the remaining Scottish block grant perform under a number of revenue-growth scenarios, including historic rates of revenue growth between 1999-00 and 2013-14. Section 5 focuses on how the options interact with changes in tax policy in the rest of the UK. Section 6 highlights particular issues for the planned devolution of powers for new welfare benefits, discretionary benefit top-ups, and powers over the housing element of Universal Credit. Section 7 discusses the idea of the Scottish and UK government compensating each other if their policies have knock-on effects on the other’s revenues or spending. Finally, Section 8 draws together the analysis and offers concluding thoughts.

2. Existing funding arrangements for the UK's devolved governments

Currently, the vast majority of taxes – including income tax, National Insurance, VAT and corporation tax, among others – are paid by people all over the United Kingdom to the UK government. On the other hand, responsibility for government spending in and on behalf of Scotland, Wales and Northern Ireland is split between the UK government and the devolved governments and local authorities in each of these countries. For instance, in 2013–14, around 40% of spending on behalf of Scotland was undertaken by the UK government (with the largest item being social security benefits, followed by defence), but the remaining 60% (including spending on health, education, transport, social services, and law and order) was undertaken by the Scottish government and Scottish local authorities.⁴ Because they raise relatively little revenue directly themselves, the devolved governments of Scotland, Wales and Northern Ireland rely on transfers from the UK government to fund this spending. Thus the devolved governments face a high vertical fiscal imbalance – their spending substantially exceeds their ability to raise tax (Commission on Scottish Devolution, 2009).

These transfers come largely in the form of block grants from the UK Treasury – one for capital (investment) spending and one for current (day-to-day) spending. On top of this, devolved governments have some relatively small revenues of their own: in Scotland's case non-domestic rates (NDR), and from April 2015, revenue from the devolved versions of Stamp Duty Land Tax and Landfill Tax (these were devolved under the Scotland Act, 2012).⁵ The devolved governments then decide how to allocate this funding between different devolved services and departments.

How these block grants *change* from year to year is determined largely by the Barnett formula. The Barnett Formula determines the change to a devolved governments grant in a given year as a population share of changes in spending on 'comparable' (i.e. devolved) services in England:⁶

$$\begin{array}{ccccc} \text{Cash change in} & & & & \text{Appropriate} \\ \text{spending by UK} & \times & \text{Comparability} & \times & \text{population} \\ \text{Govt Department} & & \text{percentage} & & \text{proportion} \end{array}$$

For example, if the UK government announces a £100m increase in Department of Health spending, if 99% of that department's budget is spending in England on responsibilities that are devolved to Scotland, and if Scotland's population is 10% of England's, then the Scottish Government's budget would increase by £9.9

⁴ Scottish Government (2015).

⁵ Land and Buildings Transactions Tax, and Scottish Landfill Tax, respectively.

⁶ A fuller explanation of how the Barnett Formula identifies changes in comparable spending can be found in Phillips (2014a)

million. Any Barnett-calculated change is added to or subtracted from the existing grant (the 'baseline').

This means that the size of the block grants today depends on the historic level of spending in Scotland, Wales and Northern Ireland before the formula was introduced in the late 1970s and on the application of the Barnett formula to changes in spending since then.⁷ Therefore there is no link between grant size and spending needs in a particular country, nor the amount of tax revenues raised in that country.

This means that the devolved governments are largely insulated from changes in both the relative and absolute levels of tax revenues in their countries (except for the relatively small taxes currently devolved). This provides a greater degree of certainty and stability in funding, but means devolved governments have less financial incentive to implement policies that boost revenues, and are subject to weaker fiscal accountability for the decisions they do make. On the other hand, devolved governments' budgets are exposed to risks associated with factors that affect relative spending needs, such as differential rates of population ageing, increases in the prevalence of illness, or changes in other socio-economic factors.

Moreover, although the Barnett Formula allocates a population share of the change in public spending in England to Scotland and the other devolved nations, the budgets of the devolved governments still face risks and incentives associated with differential population growth. This is because the Barnett Formula only determines how the block grant changes, and while these changes reflect the latest population shares, the existing block grant (the baseline) is not adjusted to account for changes in population. This means the amount of money the Scottish Government has to spend per person increases more quickly when its population is growing *less* quickly than England's, and vice versa. The Scottish Government therefore has a theoretical incentive to reduce population growth – because the existing block grant and any increments to it then has to be shared among fewer people than it otherwise would. Box 1 explains in more detail how the Barnett Formula works and its sensitivity to relative population growth.

Box 1. The Barnett formula and differential population growth

If the rate of population growth in a devolved country was the same as in England, the Barnett formula implies convergence to the same level of spending per person when spending is increasing in nominal terms (and divergence when spending is being reduced). This is because the Barnett Formula gives the same cash-terms increase in spending per person to devolved governments as in England, which will be smaller in percentage terms if spending in the devolved

⁷ HM Treasury also sometimes makes changes to grants that are not based on the Barnett formula (a process sometimes termed 'formula bypass'). In addition, grants are also changed when there are changes to what services are devolved.

country starts off higher, and vice versa. The effect of the different initial spending level (the baseline) therefore becomes proportionately less over time.

This convergence is faster the faster the rate of spending growth. Thus for Wales, Scotland and Northern Ireland, where spending per person is higher than in England, if population grew at the same rate as in England, then the Barnett Formula would lead to spending converging down to the level in England: a process known as the 'Barnett Squeeze'. The black line in Figure 1 illustrates this.

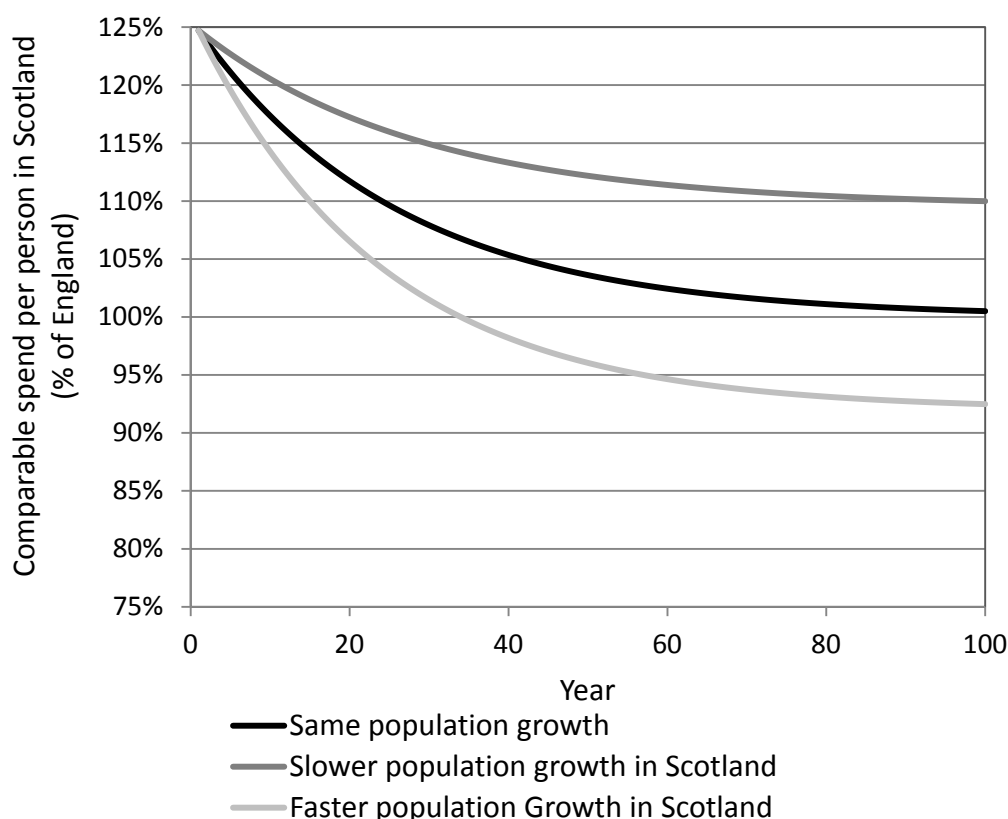
However, when population growth differs between England and a devolved country, this process of convergence is different and the Barnett Squeeze may not happen. To see this, imagine that population increases in England but not in Scotland. If the UK government keeps spending on comparable services in England fixed in per capita terms, there will necessarily be an increase in aggregate spending in Scotland. This feeds through to increased grant for Scotland, despite the fact that Scotland's population is unchanged. This causes spending to rise in Scotland relative to England.

Scotland's population growth has indeed been slower than that in England since devolution (and indeed, since long before): growth has averaged 0.35% per year compared to 0.7% a year in England.^a The dark grey line in Figure 1 illustrates that this slows the Barnett Squeeze, and means that spending in Scotland converges to a level higher than 100% of that in England (the precise level depends on how different population growth is, and how quickly spending is growing).^b In contrast, the pale grey line shows what happens when this pattern of relative population growth is reversed. In that case the Barnett squeeze happens more quickly, and in fact, relative spending in Scotland would converge to a level below 100% of that in England.

The devolved government in Scotland (and Wales and Northern Ireland) therefore faces some risk to its budget from differential population change: it benefits from relatively slow population growth, and loses from relatively fast population growth.

- a. This is one of the reasons why the convergence between spending per person in Scotland and England that the Barnett Formula is often said to result in has not actually happened to any significant extent. Phillips (2014a) discusses an additional reason related to flaws in the way the Barnett Formula has treated non domestic rates revenues.
- b. This feature of the Barnett Formula is discussed in more detail in Cuthbert (2001).

Figure 1. The impact of differential population growth on the 'Barnett Squeeze'



Note: Assumes comparable spending growth of 4% per year.

Source: Authors' calculations.

Recent years have seen increasing demands to reform the system. In Scotland's case, in particular, this includes calls for the devolution of additional tax powers and revenues to address the Scottish Government's high vertical fiscal imbalance and increase the fiscal responsibility of Scottish politicians to their electorate.⁸ As a result, the Calman Commission recommended devolution of a few small taxes and the partial devolution of income tax.⁹ As already mentioned, this was legislated for in the Scotland Act 2012. But it was after Scotland voted "No" in the independence referendum that the most substantial changes to its fiscal powers and framework were proposed.

⁸ It might be seen as somewhat unfair that devolved government budgets should be exposed to such a wide range of spending risks that lay at least in part outside the control of the devolved government. A major reason why this apparent unfairness has not been seen as an issue under the existing system, at least in Scotland's case, is that the existing funding settlement is generous relative to Scotland's needs in other ways. As discussed in further detail in Phillips (2014a), the baseline grant when Barnett first began operating in 1979 was already generous to Scotland, and has remained so due to a number of features and flaws in the Barnett Formula.

⁹ The revenue from 10 percentage points of each tax band on non-savings non-dividends income, and the power to vary each rate up or down in lock-step (so that if the Scottish Government would have to increase or cut all rates of tax by the same amount). See Commission on Scottish Devolution (2009).

3. Proposals for further tax and welfare devolution to Scotland

In the run-up to the Scottish independence referendum, the “No” campaign pledged that there would be further fiscal devolution if Scotland voted to remain in the UK. Following the “No” vote, the Smith Commission was set up to reach agreement between the five parties represented in the Scottish Parliament on a new devolution settlement for Scotland, including a package of additional tax and welfare powers.¹⁰

The Commission, which reported in November 2014 (just two months after it commenced), set out recommendations for quite substantial new fiscal powers. These were designed to strengthen the Scottish Parliament’s ability to ‘pursue its own vision, goals and objectives’, and to increase the parliament’s ‘accountability and responsibility for the effects of its decisions and their resulting benefits and costs’.¹¹ Legislation to implement the agreed devolution package (the Scotland Bill 2015) has just cleared the UK House of Commons and is now being debated in the House of Lords.¹² In parallel, the UK and Scottish governments are also negotiating the new fiscal framework that these new tax and welfare powers will operate under.

This section first explains the tax revenues and powers, and welfare spending and responsibilities that will be devolved to Scotland (sub-section 3.1).¹³ It then sets out the principles the Smith Commission says the new fiscal framework should satisfy (sub-section 3.2).

3.1 Planned devolved tax and welfare powers

On the tax side, the Smith Commission recommended the devolution of income tax rates and bands on non-savings and non-dividend income, and all associated revenues. It also recommended the devolution of Air Passenger Duty (APD) and the Aggregates Levy to the Scottish Parliament, and the assignment of half of VAT revenues raised in Scotland to the Scottish Parliament (assignment means the Scottish Government gets the revenue from a tax but does not have the power to vary the tax rate). The UK government has agreed to this package, which is being legislated for in the Scotland Bill 2015. As a result, the composition of the Scottish Parliament’s income will change significantly (Table 1).

¹⁰ The parties represented in the Smith Commission were the Scottish Conservatives, Scottish Greens, Scottish Labour Party, Scottish Liberal Democrats, and the Scottish National Party.

¹¹ Smith Commission (2014).

¹² The UK Government published an initial Command Paper setting out its response to the Smith Commission and draft legislative clauses to implement the devolution package. See HM Government (2015). The initial legislative arrangements have been amended to further increase powers in some areas. See, for instance, <https://www.gov.uk/government/news/scottish-secretary-scotland-wins-if-holyrood-uses-new-powers-well>.

¹³ A fuller discussion of the tax and welfare powers to be devolved can be found in Phillips (2014b).

Table 1: Revenues devolved to the Scottish Government or Scottish local authorities (£s millions, 2013-14 values_

	Historic Situation	Scotland Act 2012	Smith Commission Package
Non-domestic Rates	1,927	1,927	1,927
Council Tax	1,941	1,941	1,941
Income Tax		4,258	10,911
Stamp Duties (Land and Buildings)		385	385
Air Passenger Duty			251
Landfill Tax		105	105
Aggregates Levy			50
Assigned VAT		0	5,030
Total devolved revenues	3,868	8,617	15,571
Total devolved and assigned revenues	3,868	8,617	20,600
Devolved expenditure	40,813	40,813	43,334
Devolved revenue as % of estimated devolved expenditure	9%	21%	36%
Devolved and Assigned revenue as % of estimated devolved expenditure	9%	21%	48%

Note: This table includes Scottish local authority spending and revenues (council tax) as ultimately these powers and revenues are subject to the control of the Scottish Government.

Source: Adapted from Government Expenditure and Revenue Scotland (Scottish Government, 2015)

In particular, the share of the Scottish Government's spending that will be funded by tax revenues either devolved to it, or assigned to it, will increase from less than one-tenth historically, and two-tenths under provisions of the Scotland Act 2012 to around one-half.

Table 2: Benefits to be devolved to Scotland, (£s millions, 2013–14 values)

Attendance Allowance	481
Carer's Allowance	182
Disability Living Allowance	1,473
Industrial Injuries Disablement Benefit ¹	83
Personal Independence Payment ²	17
Severe Disablement Allowance	91
Cold Weather Payment	1
Funeral Payment ³	4
Sure Start Maternity Grant	3
Winter Fuel Payment	185
Total expenditure on benefits to be devolved	2,521

Source: Adapted from Government Expenditure and Revenue Scotland (Scottish Government, 2015)

On the welfare spending side, the Smith Commission proposed fully devolving a number of social security benefits that are currently reserved to the UK Parliament. This includes the power to reform the structure of these benefits (and indeed create new benefits in these areas). The combined value of these

devolved benefits is around £2.5bn per year (Table 2). Most welfare spending will remain reserved to Westminster though.

The Scottish Government will also have power to vary the housing element of Universal Credit (UC), and make changes to UC payment arrangements (such as making payments weekly rather than monthly). In addition, the Scottish Government will have the power to top-up other non-devolved benefits, and provide discretionary financial support to households suffering hardship.

3.2 Smith's principles for a new fiscal framework

The Smith Commission recognised that alongside these new powers, revenues, spending responsibilities, Scotland will require a substantially revised fiscal framework. This means decisions on: the scope of increased borrowing powers which will be required to reflect the additional economic risks and revenue volatility that the Scottish Government will face; the extent and scope of fiscal rules governing Scottish Government deficits and debt; arrangements for independent fiscal scrutiny, including fiscal forecasting; and arrangements for governing the increasingly complex interactions between Scottish and UK fiscal policy, including dispute resolution. Most obviously, of course, the block grant currently given to the Scottish Government by the UK Treasury will have to be adjusted to account for the revenues that the Scottish Government will have, and the additional spending responsibilities it will take on.

In the first year that new powers are devolved, these block grant adjustments (BGAs) should be relatively straightforward, at least in principle. The initial reduction in the block grant to account for devolved taxes should be equivalent to a forecast of the amount of revenue devolved (and therefore forgone by the UK government). Likewise the devolution of welfare spending will need to be accompanied by an addition to the block grant equivalent to the forecast level of spending in Scotland by the UK government in that year had the powers not been devolved.

The Smith Commission Agreement (SCA) recognised this in its 1st “No Detriment” principle: that neither the Scottish nor UK Government’s budgets should be larger or smaller simply as a result of the initial transfer of tax and/or spending powers, before considering how these are used.¹⁴

In future years however the BGAs have to be indexed. If the BGA was not indexed, but fixed at the year 1 amount, then in the face of inflation and economic growth its relative value would be eroded over time. The implication would be that the Scottish Government would benefit and the UK Government suffer “detriment” from devolution over time, as tax revenues grow, but the amount taken off the block grant to account for this remains constant. Likewise, the addition to

¹⁴ Paragraph 95.3 of Smith Commission (2014).

Scotland's block grant to account for devolved welfare powers will need to be indexed to account for likely spending growth in the future.

The SCA recognised this and stated that the BGAs should be 'indexed appropriately'. But it stopped short of saying how this indexation should be made, even though this mechanism is central to the reallocation of risk and reward that provides one of the main justifications for tax and welfare devolution in the first place, and can play a significant role in determining just how much money the Scottish and UK governments have to spend.

However, the SCA does set out a number of further principles that the revised fiscal framework, including the BGA, should satisfy.

First, that the Scottish Government should have "economic responsibility" for its policy decisions.¹⁵ This means that the Scottish budget should benefit in full from the policy decisions of the Scottish Government that increase revenues or reduce expenditure, and bear the full costs of policy decisions which reduce revenues or increase expenditure.

This immediately rules out indexing the BGAs to what happens to devolved tax revenues or welfare spending in Scotland. To see this, suppose that Scottish Government policy leads to faster revenue growth – whether by higher tax rates, or through faster growth in the tax base. If the BGA were indexed to this, these higher revenues would be clawed back by an equivalent increase in the BGA. This would remove the incentive for the Scottish Government to grow its revenues (and constrain its spending), and would mean it did not bear the costs or benefits of its own decision making. However, as we will discuss later, the distinction between revenue or expenditure changes which are due to Scottish Government policy decisions and those that are due to other factors (such as underlying economic or demographic factors) is not easy to make, and the indexing methods we discuss devolve responsibility for both types of fiscal risk.

Second, the SCA states that the UK Government should continue to manage risks and shocks that affect the whole UK.¹⁶ This means, for instance, that if there was a recession that led to falls in income tax revenues across the whole of the UK, including Scotland, then the UK Government should use its borrowing powers to smooth this, rather than the Scottish Government having to borrow itself to make up for a (temporary) shortfall in revenue. This principle of risk sharing for shocks that affect the whole of the UK seems reasonable within the context of an ongoing fiscal union. It also seems sensible given the potentially greater costs of Scottish Government borrowing relative to UK government borrowing.

Third, in addition to the aforementioned principle that there should be no detriment from the decision to devolve a revenue stream or spending responsibility, the SCA identifies an additional "No Detriment" principle. This is

¹⁵ Paragraph 95.1 of Smith Commission (2014).

¹⁶ Paragraph 95.8 of Smith Commission (2014).

the principle that there should be no detriment to one government as a result of policy decisions of the other government post-devolution.¹⁷ This additional ‘no detriment’ principle actually includes two elements:

- Where either the UK or Scottish Governments make policy decisions that affect the tax receipts or spending of the other, that the decision-making government will either reimburse the other if there is an additional cost, or receive a transfer from the other if there is a saving. This is the principle of compensation.
- Changes to devolved tax rates in Scotland should only affect public spending in Scotland. Changes to tax rates in the rest of the UK, for which responsibility in Scotland has been devolved, should only affect public spending in the rest of the UK. This is the principle of taxpayer fairness.

As we discuss in detail later in the paper, while both elements of this additional ‘no detriment’ principle may seem sensible, implementing them in practise is likely to be problematic.

Fourth, the SCA also states that the underlying Scottish block grant should continue to be adjusted each year according to the workings of the Barnett Formula.¹⁸ This means any process of indexing the BGA will work alongside the Barnett Formula and changes in devolved tax revenues in determining the Scottish Government’s budget each year.

The final principle is that the resulting framework should be implementable and sustainable.¹⁹ More particularly, that once the framework is agreed, it should, as far as possible, operate in a systematic and mechanical way, without the need for ‘frequent ongoing negotiation’. This means avoiding systems for indexing the BGA that require frequent ad-hoc adjustments to meet the other principles (such as the ‘compensation’ and ‘taxpayer fairness’ elements of the ‘no detriment’ principles).

Taken together, these principles help in determining which way to index the BGA. But they also set a high bar for the new fiscal framework and BGA indexation mechanism. As we show in the following sections, this is a bar which cannot be fully met.

¹⁷ Paragraph 95.4 of Smith Commission (2014).

¹⁸ Paragraph 95.2 of Smith Commission (2014).

¹⁹ Paragraph 95.6 of Smith Commission (2014). Additional recommendations for the fiscal framework that do not directly relate to the BGA include the provision of necessary borrowing powers; strengthened independent fiscal scrutiny of the Scottish Government’s fiscal decisions and Scotland’s public finances; and, that the two governments should update both parliaments on any changes to Scotland’s fiscal framework.

4. Assessing the options for adjusting Scotland's block grant

In this section we assess a number of specific BGA indexation approaches that share one thing in common: in each, the BGA is linked in some way to what happens to equivalent tax revenues in the rest of the UK (rUK). We focus on such approaches because they offer a relatively automatic way of ensuring that the UK Government continues to manage fiscal risks that affect the whole of the UK. To see this consider what happens if a recession hits the whole of the UK and revenues fall. In these circumstances, Scotland's revenues would fall, but lower revenues in rUK would mean the BGA is also reduced: an increase in the amount the Scottish Government would therefore offset the fall in tax revenues. This way of indexing the block grant reduction therefore acts to smooth the Scottish Government's budget, which would mean less need to rely on borrowing to smooth the ups and downs of tax revenues. This might be attractive to both the Scottish Government – which might find it relatively difficult and expensive to borrow –, and the UK government – which might be wary of the risks of giving Scotland large and wide-ranging borrowing powers.

On the other hand, if revenues from the devolved tax grow faster than comparable revenues in the rest of the UK, then they also grow more than the BGA under such an approach: this would act to increase the Scottish Government's budget. Alternatively, if revenues grow less quickly than those in the rest of the UK, they grow less quickly than the BGA: this would act to reduce the Scottish Government's budget. The Scottish Government therefore still has financial incentives to grow its revenues. This approach also goes a significant way to ensuring that the Scottish government automatically bears 'economic responsibility' for its decisions – at least the responsibility for effects on the relative growth of devolved revenues in Scotland.

Other methods not linked to tax revenues (or welfare spending) in rUK do exist, of course. For instance, deducting a fixed proportion off the block grant equal to the proportion deducted in the first year.²⁰ Hence, if the initial block grant were £30 billion and £10 billion in tax revenues were devolved, the BGA is set equal to one-third of the underlying block grant every year going forwards. This has some similarities with the way in which business rates have been devolved.²¹ Alternatively, the BGA could be indexed at some fixed rate.²² But this begs the

²⁰ Independent Commission for Funding & Finance for Wales (2010).

²¹ This was the underlying principle, although as shown in Phillips (2014a), there were errors in implementing this via the Barnett Formula's comparability factors for the UK's Department for Communities and Local Government.

²² See for instance, Dr James Cuthbert's supplementary evidence to the House of Lords Economic Affairs Committee enquiry on Scotland's new fiscal framework, available at: <http://data.parliament.uk/writtenevidence/committeeevidence.svc/evidencedocument/economic-affairs-committee/the-devolution-of-public-finances-in-the-united-kingdom/written/21826.pdf>.

question, what rate? The rate chosen would have significant implications for the size of the net block grant given to Scotland and it would therefore be unsurprising if it become subject to controversy and dispute. Furthermore, under both these approaches the Scottish Government would be exposed to revenue shocks that affect the whole of the UK, violating one of the Smith Commission's principles. This is because the BGA does not automatically adjust when revenues elsewhere in the UK decline – it would instead be linked to the amount spent on devolved services in rUK (which is unlikely to fall following a revenue shock, at least straight away) or increase in some arbitrary fixed way. Changes in the BGA will therefore not act to offset changes in revenues caused by common shocks in any way.

There are a number of different ways in which the BGAs can be indexed to tax revenues in the rest of the UK. The first, suggested by Gerald Holtham, is to index the initial BGA to the subsequent percentage change in revenues from the equivalent taxes in rUK.²³ Following his terminology we label this the Indexed Deduction (ID) approach. This approach can be modified so that instead of indexing to the aggregate percentage change in revenues, one indexes to the percentage change in revenues per person. We label this approach the Per Capita Indexed Deduction (PCID). Finally, we consider an approach set out in Bell and Eiser (2014) which they term the Levels Deduction (LD) approach. This updates the BGA each year by adding a population share of the change in comparable revenues in rUK.

In what follows we first assess how these different options perform when population grows at the same rate in Scotland as in the rest of the UK (sub-section 4.1), before examining the difference that differential population growth can make (sub-section 4.2). We find that the different approaches could mean the Scottish Government's budget evolving quite differently over time, and involve Scotland bearing different sorts of risks and incentives. We then look at how Scotland would have fared under the different approaches in the period between 1999-00 and 2013-14 using tax revenue data from HMRC and benefit spending data from DWP to get some sense of how big these differences might be in practise (sub-section 4.3). Section 5 then discusses particular issues when changes in revenues in rUK are driven not by underlying revenue growth but by policy changes, and Section 6 discusses particular issues for the BGAs to account for additional welfare spending responsibilities.

4.1 When population grows at the same rate as in the rest of the UK

In order to analyse the effects of different indexation methods a little simple algebra and a few graphs are very helpful.²⁴ To keep the exposition as simple as

²³ Independent Commission for Funding & Finance for Wales (2010).

²⁴ The Appendix contains further detail on the algebra that explains how the different BGA indexation methods perform under different circumstances.

possible we illustrate this discussion with reference to devolved income tax, but the same principles apply to other taxes.

First, let's start with a system where the Scottish Government's spending is fully funded by the block grant (a rough approximation of the current situation). Under the Barnett Formula, Scotland's block grant, G , is equal to grant in the previous period plus a population share of the change in comparable spending (C) in England. P denotes population and the superscripts S and E denote Scotland and rUK respectively:²⁵

$$G_t = G_{t-1} + \frac{P_{t-1}^S}{P_{t-1}^E} (C_t^E - C_{t-1}^E)$$

When income tax is devolved, Scotland's block grant is calculated as before, but is then adjusted by the block grant adjustment (BGA).

$$G_t = \left(G_{t-1} + \frac{P_{t-1}^S}{P_{t-1}^E} (C_t^E - C_{t-1}^E) \right) - BGA_t$$

In the first period that the tax is devolved, the BGA is equal to the value of the devolved tax revenue in Scotland:

$$BGA_t = T_t^S$$

The first approach to indexing the BGA is **Indexed Deduction** (ID). This indexes the change in the BGA to the percentage change in total comparable revenues in rUK. In time $t+1$, the ID method calculates the BGA as:

$$BGA_{t+1} = \frac{T_{t+1}^E}{T_t^E} \cdot BGA_t$$

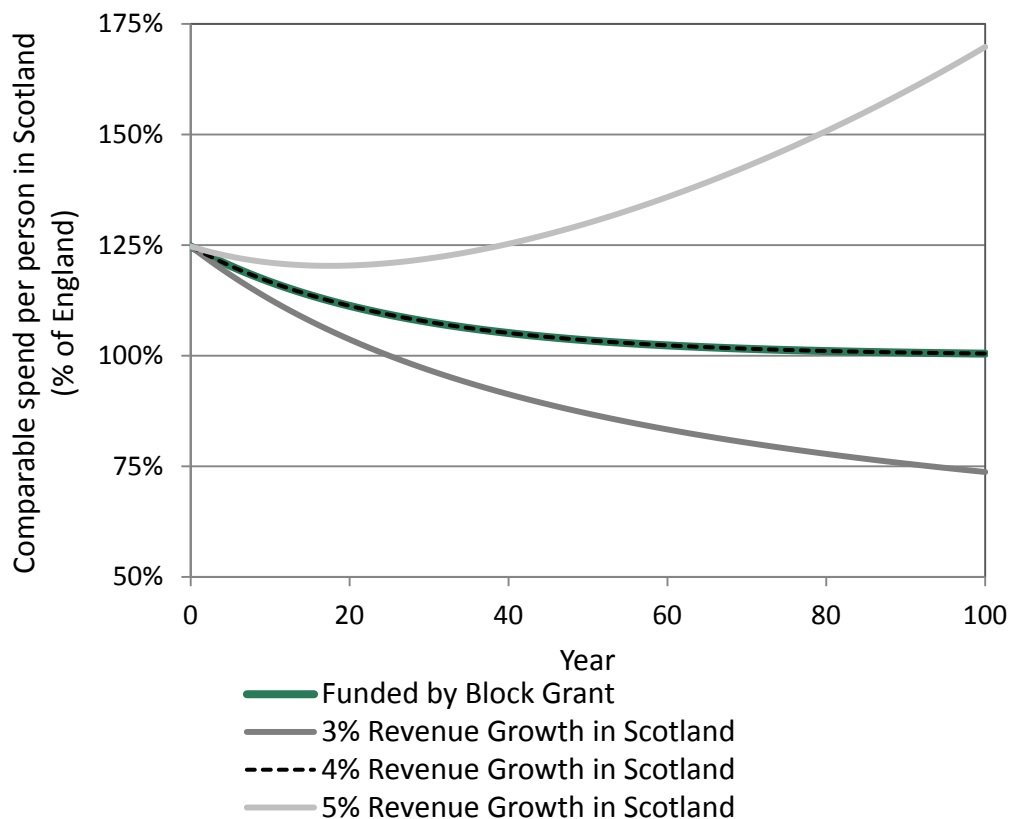
For example, if comparable revenues in rUK grow by 5%, then the BGA also grows by 5%. When population grows at the same rate in Scotland as in the rest of the UK, the ID and **Per Capita Indexed Deduction** (PCID) methods for indexing the BGA are equivalent and produce the same outcome. Under both these BGA indexation approaches, the Scottish Government's budget will be higher than it would have been under the status quo (i.e. no tax devolution) if income tax revenues in Scotland grow relatively more quickly in percentage terms than those in rUK. Conversely, the Scottish Government's budget will be lower than it would have been under the status quo if income tax revenues in Scotland grow relatively more slowly in percentage terms than those in rUK.

This is illustrated in Figure 2 which shows three scenarios. First, where Scotland's income tax revenues grow by 3% per year, second where Scotland's income tax revenues grow by 4% per year, and third where Scotland's income tax

²⁵ Strictly speaking, the Barnett Formula applies to spending in England rather than rUK. We use the term rUK here for simplicity of exposition and consistency with plans for the BGAs.

revenues grow by 5% a year . In each case income tax revenues are assumed to grow by 4% per year in rUK.²⁶

Figure 2. The impact of differential rates of growth of revenue on Scottish Government spending power under the ID approach



Note: Assumes comparable spending growth of 4% per year.
Source: Authors' calculations.

Under the **Levels Deduction** (LD) approach the change in the BGA is equal to a population share of the change in comparable revenues in rUK:

$$BGA_{t+1} = \frac{P_t^S}{P_t^E} (T_{t+1}^E - T_t^E)$$

For example, if income tax revenues increased by £10 billion in rUK, then if Scotland's population was 8% of rUK, Scotland's BGA would increase by £0.8 billion.

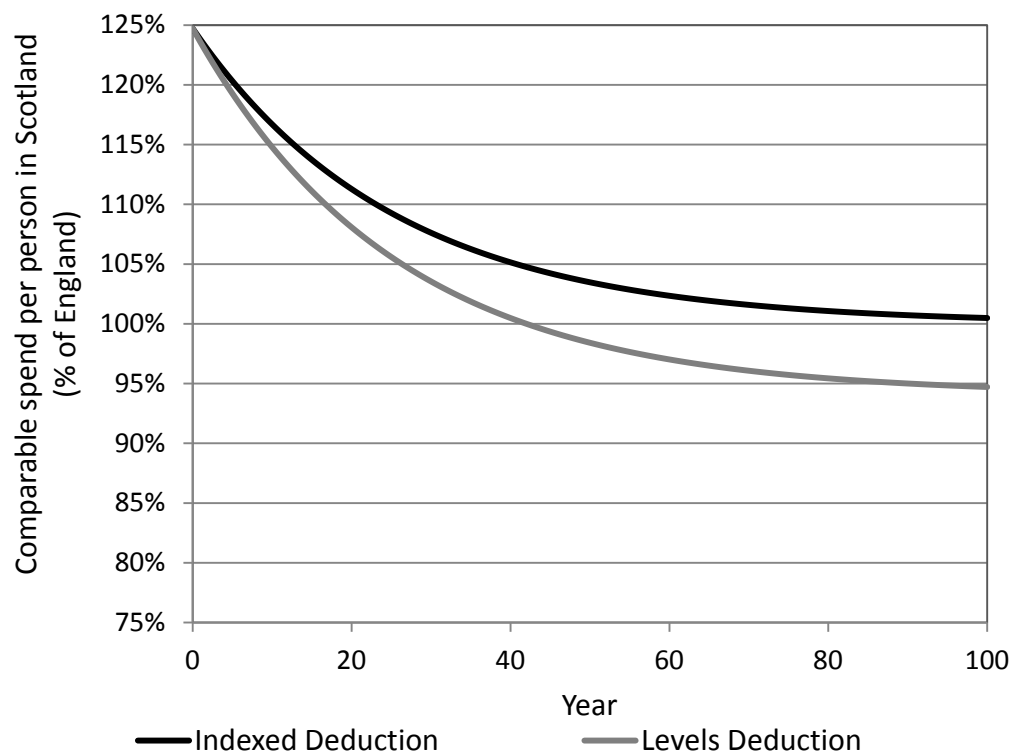
If income tax revenues per person are the same in Scotland as in rUK , then the LD and the ID and PCID lead to the same change in the BGA. But when revenues per person differ, a percentage change in the BGA is not the same as a population-share based change. For instance, income tax revenues per person have

²⁶ Additional assumptions include that devolved spending in Scotland starts at 125% of the level of spending on comparable services per person in rUK, that spending growth in rUK on comparable services grows at 4% per annum, that population growth is 0% in both Scotland and rUK, and that income tax revenues in Scotland per person begin at 88% of the average for rUK.

historically been lower in Scotland than in rUK. This means any given percentage increase in rUK revenues is bigger in cash terms per person than the equivalent percentage increase in Scottish revenues. While Scottish revenues per person are lower than those in rUK, the Scottish Government therefore benefits from bigger budget increases under the ID and PCID approaches than under the LD approach.²⁷

To see this, imagine a hypothetical scenario where income tax revenues per capita are £1000 in rUK but only £880 in Scotland, and that there are 1000 taxpayers in rUK and 90 in Scotland (and thus total tax revenues are £1,000,000 in rUK and £79,200 in Scotland).²⁸ If rUK revenues grow by 10% (to £1,100,000), then Scotland's BGA under the ID approach would grow by 10% to £87,120, a nominal increase of £7,920. However, the LD approach would adjust Scotland's BGA by a population share of the nominal increase in rUK revenues. The nominal increase in rUK revenues is £100,000, and a population share of this is £9,000, which is clearly higher than the increase in the BGA of £7,920 under the ID.

Figure 3. Comparing the LD and ID approaches when population grows at the same rate in Scotland and rUK



Note: Assumes comparable spending growth of 4% per year.
Source: Authors' calculations.

²⁷ Of course, the LD approach conversely advantages Scotland when revenues are falling.

²⁸ These figures are hypothetical, but are chosen to represent approximate values for relative income tax revenues and population.

This is further illustrated in Figure 3 which assumes revenues grow by 4% a year in both Scotland and rUK.²⁹ It shows that the divergence in overall funding for the Scottish Government under the ID and LD approaches grows over time and can become notable. In this example, for instance, after 30 years, the gap between the two methods is equivalent to around 4% of the Scottish Government's budget. In cash terms, that would be equivalent to around £1.2 billion in today's terms, a non-trivial sum.

Indeed, a feature of the LD approach in such circumstances is that Scottish revenues would have to grow at a higher rate (in percentage terms) than those in rUK to keep up with the BGA (which is determined by a population-based share of equivalent revenue growth in rUK). This might be seen as in conflict with the first 'no detriment' principle: that neither government should lose simply as a result of the decision to devolve a tax or spending item. While a strict interpretation of this might mean this only applies in the first year of devolution, adopting a system that might be expected to lead to such an outcome in subsequent years (unless revenues grow at a faster rate in Scotland) does not seem in the spirit of this principle.

Section 2 showed that the Barnett Formula leads to convergence in spending per person in Scotland and rUK when population grows at the same rate in each. Assuming Scotland starts from a position of lower revenue per person than rUK, the LD method seems likely to speed up this 'Barnett Squeeze'. This is because unless revenues grow at a higher percentage rate in Scotland than rUK, they will grow less quickly than the BGA, putting downward pressure on the Scottish Government's budget. As Figure 3 shows, eventually this would actually lead to 'overshooting' and spending per person in Scotland converging to a level below that in rUK.³⁰

4.2 When population growth differs in Scotland

When population grows at different rates in Scotland than in the rest of the UK, the ID and PCID methods for indexing the BGA are no longer equivalent. ID is calculated as before, with PCID calculated as:

$$BGA_{t+1} = \left(\frac{T_{t+1}^E}{T_t^E} \cdot \frac{P_t^E}{P_{t+1}^E} \cdot \frac{P_{t+1}^S}{P_t^S} \cdot BGA_t \right)$$

The ID method now exposes Scotland to risks associated with its population growing at a different rate to that of rUK. For instance, if Scotland's population grew less quickly than that in rUK (as it has typically done), this would likely act as a drag on growth in total revenues in Scotland. By indexing the BGA according to aggregate rather than per person revenue growth in rUK, the ID method would therefore see the Scottish Government losing out unless its revenues per person

²⁹ The other assumptions are the same as in Figure 2.

³⁰ This level is a weighted average of 100% and Scotland's devolved income tax revenues per person as a percentage of those in rUK.

grew more quickly. Conversely, the Scottish Government would see a budgetary gain if its population grew faster than in rUK. We show below that, in the long run, these gains or losses from differential population change grow exponentially, and can become very large.

The PCID option clearly protects the Scottish budget from the risk that its population grows relatively slower than rUK's. But equally, it means the Scottish Government's budget would not benefit if revenues increased as a result of a relatively faster growing population. This blunts incentives for the Scottish Government to boost the Scottish economy by attracting more people to live and work in Scotland (Scotland could even be worse off if those migrants paid less income tax, on average, than the existing workforce). Furthermore, to the extent that Scottish Government policies affect the population of Scotland – through in or out migration, say –, it would mean that Scotland would not necessarily face the full 'economic responsibility' for its policies.

Differential population growth has a rather complex effect on comparisons between the ID and LD methods. If revenues per person started off equally, then the Scottish Government's budget would subsequently do better with LD when tax revenues are increasing. This is because whereas the ID approach takes no account of differential population change and indexes the BGA upwards by the full aggregate increase in tax revenues in rUK, the LD approach uses updated population figures each year when calculating the *increments* to the BGA. It therefore takes account of population change when calculating these increments (even if it does not adjust the existing BGA, much like the Barnett Formula and the underlying block grant – See Box 1, above).

To illustrate this, return to a hypothetical case where rUK population is 1000, rUK revenues per capita are £1,000, and Scottish population is 90 with revenues per capita now also £1,000. Suppose that rUK population increases by 10% over the space of 10 years, while all other parameters remain unchanged.

- The ID approach would simply adjust Scotland's BGA up by 10% (the rise in total rUK revenues) which would be equivalent to £9,000 – so that the Scottish Budget would be £9,000 worse off than under the status quo Barnett approach.
- The PCID approach would completely insulate the Scottish Government's budget from the rise in rUK population – there has been no change in rUK revenues per person, and thus no change in the BGA.
- The LD approach would adjust the BGA by a population share of the rUK tax increase each year. Over time, the Scottish population share would have fallen from 9% to 8.18%. The result of applying these population-based increments to the BGA each year would be to increase it by £8,619 after the ten years. This is clearly less than the £9,000 increase under the ID approach.

However, if Scottish revenues per capita are lower than those in rUK (as is currently the case), then the Scottish budget may do worse from the levels

approach than the ID approach, at least ‘initially’. This is because an equivalent percentage increase in the BGA may be less, in cash terms, than the population share of the cash terms increase in rUK revenues. To illustrate, return to the previous example but assume that Scottish per capita tax revenues are £880 rather than £1,000, so that Scotland’s total revenues and BGA start off at £79,200. A 10% increase in Scotland’s BGA under the ID approach is now equivalent to £7,920, which is less than the £8,619 adjustment made by the LD approach.

But, even if Scotland’s initial revenues per capita are lower than rUK’s, if Scotland’s population continues to decline relatively over time, there will come a point where the LD approach results in a more generous grant to Scotland than ID. The fact that the levels approach adjusts for Scotland’s (declining) population share at the margin means after a while it starts to become relatively more generous to Scotland at the margin compared to the ID approach. Eventually there must be a crossing point where it becomes more generous in absolute terms to Scotland. The larger the initial gap between Scottish and rUK revenues per person, and the narrower the difference in population growth between the two countries, then the ID approach will provide a relatively more generous budget for the Scottish Government than the LD approach for longer.

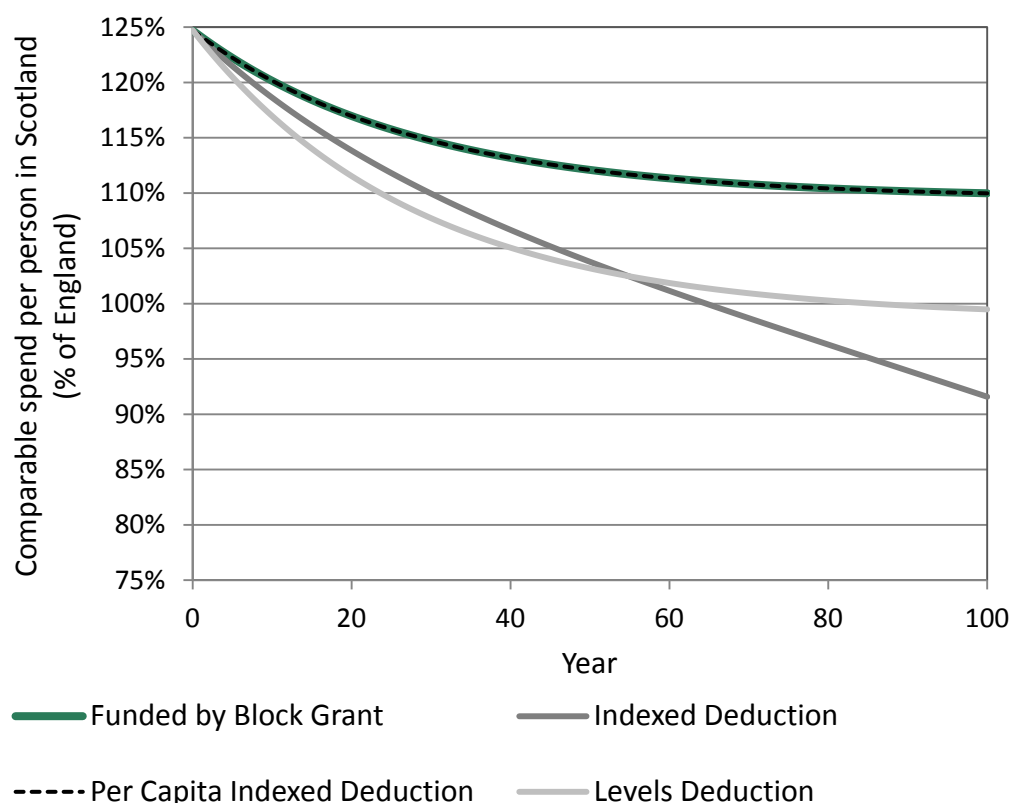
Figure 4 shows what happens when revenues per person in Scotland and rUK both grow at 3.3% per annum, but population grows by 0.35% a year in Scotland and 0.7% a year in rUK (so total revenues in rUK grow by about 4% a year compared to about 3.65% in Scotland). As in Figures 2 and 3, we assume that revenues per person in Scotland are 88% of the equivalent revenues in rUK.³¹

Clearly, the PCID method is most beneficial for Scotland as it insulates Scotland fully from both from the starting position of lower revenues per person, and subsequent differential population growth. Scotland does less well under the ID approach, but initially not as poorly as under the LD method. This is because the effect of relatively lower revenues per person initially outweighs the impact of differential population growth. This gap grows for the first 24 years (peaking at about 2 percent of the Scottish Government’s budget), but then starts to shrink as the cumulative impact of differential population starts to outpace the differences in revenues per person. Eventually, after around 55 years ‘crossover’ occurs where the Scottish Government’s budget under ID falls below that under LD. Beyond this point, the Scottish Government’s budget becomes progressively relatively less generous under ID relative to LD if relative population decline continues.

Thus while ID might seem like a compromise between PCID and LD in the short term, in the long term there are even bigger differences between budget outcomes under the ID and PCID methods when population growth differs between Scotland and rUK, than there are between the LD and PCID methods.

³¹ These rates of growth are approximately equal to the averages for Scotland and rUK since the advent of devolved government in 1999. Other factors are assumed to be the same as in Figure 2.

Figure 5. Comparing the LD, ID and PCID approaches when population grows more slowly in Scotland than rUK



Note: Assumes comparable spending growth of 4% per year.

Source: Authors' calculations.

Furthermore, as highlighted in Cuthbert (2015), when population grows relatively more slowly in Scotland, ID would lead to continual and accelerating declines in government spending per person in Scotland relative to rUK (unless faster revenue growth per person offsets the population decline). Indeed, ultimately, this approach would imply a *negative* budget, although of course, in practice reforms would take place long before such an eventuality arose. But this implies ID might not be sustainable as a long-term solution if relative population decline were to continue, and it were felt the Scottish Government did not have the tools to address and reverse this trend.

The LD method could also result in government spending per person in Scotland falling significantly below that in rUK if Scotland's relative population continually declined.³² Alongside the fact that revenues per person need to grow more quickly to keep up with the BGA even if population growth is the same as in rUK, this may make the LD seem particularly undesirable. But as we see in Section 5 though the way in which the LD approach mimics the way the Barnett Formula works has an advantage with respect to changes in tax policy in rUK and the

³² Although the Scottish Government's budget could not become negative in this instance (again, see Cuthbert (2015)).

Smith Commission's principle that .(Indeed, this is why we consider this approach despite its other less desirable features).

4.3 Illustrating using data from 1999-00 to 2013-14

In this sub-section we illustrate what sort of difference the factors discussed above can make during a medium term horizon (14 years) using historical data.³³ In doing this a few things should be noted:

- First, that the specific dates chosen correspond to the period for which we have some data available, and to an extent are arbitrary. The idea here is to get a sense of the magnitude of changes that can reasonably occur over the space of a few years, rather than assess what would have happened in these specific years. If we had data for a longer or shorter time horizon and used that, the specifics would likely have been very different.
- Second, that the data on tax revenues available relates to overall revenues (whereas for income tax and VAT, only a proportion of revenues are being devolved), and includes the effects of tax policy changes as well as underlying revenue changes. This means that the figures here provide an approximate rather than exact estimate of the effects of the different BGA indexation mechanisms during this period.

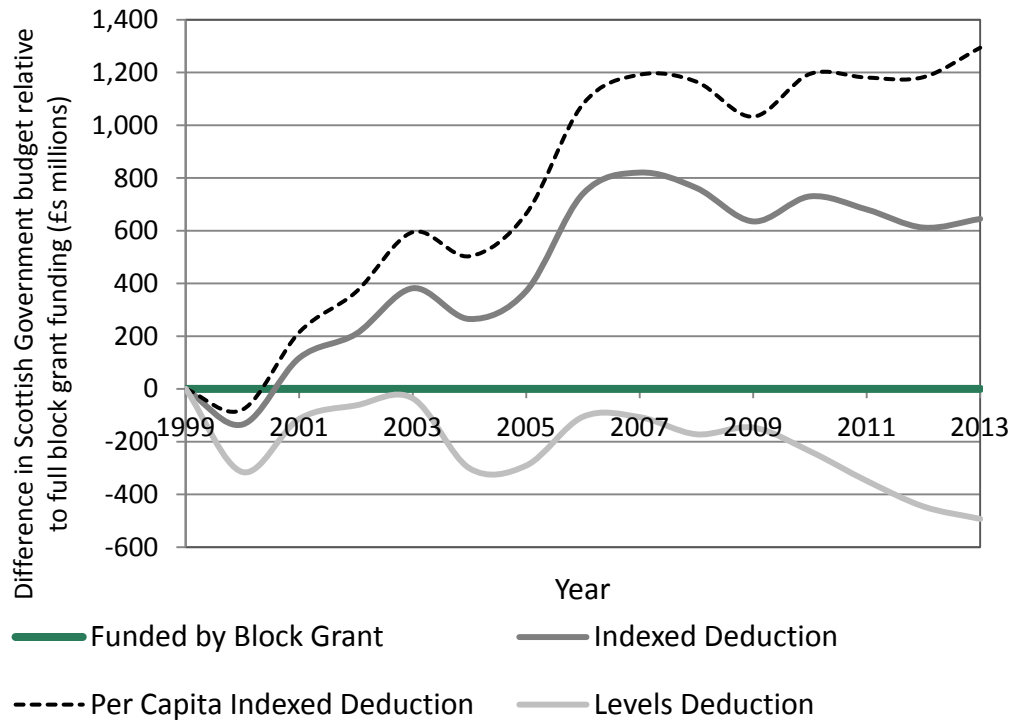
During this period as a whole, income tax revenues grew by 74% in Scotland, compared to 67% in rUK. Overall revenues from the taxes to be devolved or assigned grew by 79% and 72%, respectively. This faster rate of growth took place despite slower population growth in Scotland: 5.0% compared to 9.6% in rUK. So the difference in revenue growth per person was even greater. However, while the rate of growth of Scotland's revenues was quicker than in rUK, it was still less than Scotland's population-based share of rUK revenues as rUK revenues per person started off higher.

Figure 5 shows what this all would imply for Scotland's budget. In it, we assume that revenues from all devolved taxes are pooled and a single BGA made. Such a pooled approach would have the benefit of simplicity and consistency has been discussed before, although it may be considered inappropriate to use the same method to adjust for each tax, if the fiscal risks associated with them differ significantly. Figure 6 repeats the analysis but for income tax only. Spending power is shown in cash terms (millions of pounds) relative to a no-devolution scenario.

The Scottish Government's budget would unsurprisingly have done best under the PCID approach as it would have gained in full from its relatively faster rate of revenue growth per person. Under the ID approach, Scotland's relatively slower

³³ Disaggregated Tax Revenues (HMRC) and Sub-National Benefit Expenditure Estimates (DWP). Available at and at respectively. Accessed in August 2015.

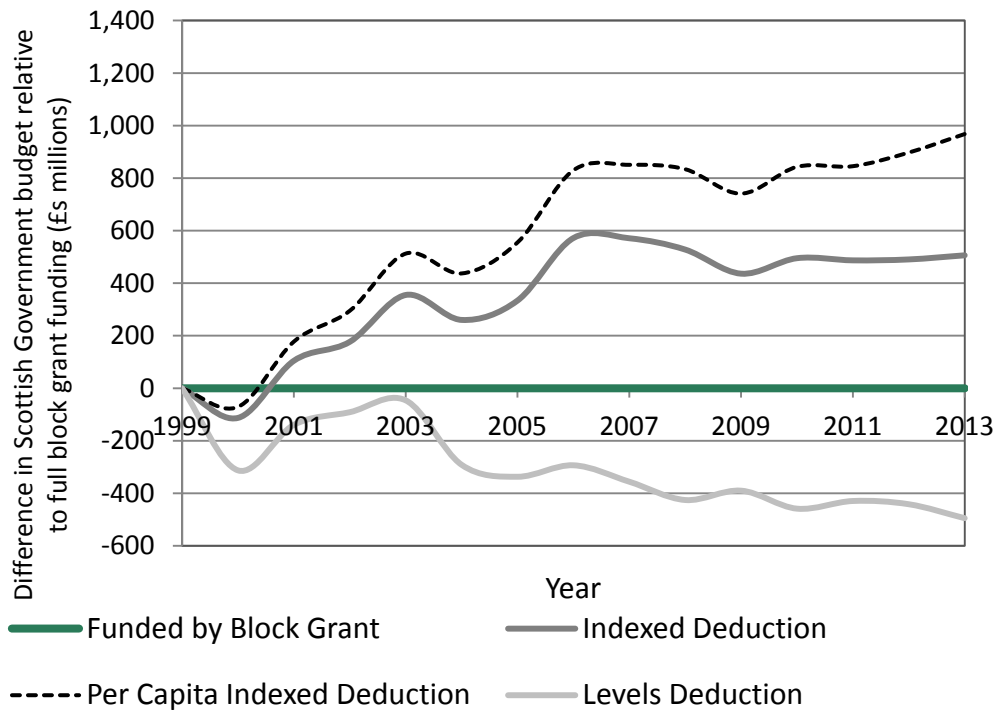
Figure 5. Comparing the LD, ID and PCID approaches using data from 1999–00 to 2013–14 for all taxes



Note: Assumes comparable spending growth of 4% per year.

Source: Authors' calculations

Figure 6. Comparing the LD, ID and PCID approaches using data from 1999–00 to 2013–14 for income tax only



Note: Assumes comparable spending growth of 4% per year.

Source: Authors' calculations

population growth would have acted as a drag on this. After the 14 years in question, this would have equated to a difference of around £650 million a year relative to PCID deduction for all taxes, and £460 million for income tax alone. The relatively low starting levels of revenue in Scotland mean that the Scottish Government's budget would have done least well under the LD approach. If these methods were applied to all taxes to be devolved under the Scotland Bill proposals, then the gap between LD and ID would have been £1,150 million after this 14 year period. The gap between LD and PCID would have been £1,790 million a year by the end of this period. These differences are equivalent to over 3% and almost 6% of the Scottish Government's block grant, respectively, in 2013–14. Figure 6 shows that most of this difference results from Income Tax: both because this is the largest component of taxation to be devolved, and the tax where Scottish revenues per person are particularly low relative to rUK.

If these patterns were projected forward, the gaps between the LD and indexed deduction approaches would not be so big. This is because the stronger growth in revenues per person in Scotland during the period in question has narrowed the difference in revenues per person between Scotland and rUK. As discussed above, this smaller gap means that the LD and ID approaches will now have more similar (although still markedly different) impacts than they would have previously.

To reiterate, it would be unwise to pay too much attention to the precise figures here (they are subject to margins of error). But this analysis does show that the differences between the different methods can quickly become fairly large.

5. Accounting for changes to tax rates in the rest of the UK

Changes to Scottish income tax rates clearly have no bearing on the BGA – the Scottish budget must bear the consequences. Thus a Scottish tax rate change which increases the Scottish Government's revenues relative to revenues in rUK will feed into the Scottish budget. However, although there are no direct consequences for the BGA, a tax rate change by the Scottish Government may have indirect implications for the UK Government budget. For instance, changes in income tax rates will affect Scottish households' net incomes, which may affect the amount of Universal Credit to which they are eligible (which is assessed on the basis of net incomes). The Smith Commission argued that these budgetary interdependencies should be compensated in some way through the principle of 'no detriment in respect of government budgets'. We discuss the difficulties of this element of the second 'no detriment' principle in Sections 6 and 7 of this paper.

Under the methods outlined above, however, changes to tax rates in rUK will have an effect on the BGA, as they will increase or decrease the revenues in rUK, which implies a corresponding increase or decrease in the BGA. Doesn't this breach the 'taxpayer fairness' element of the second 'no detriment' principle though? Changes in devolved taxes in the rest of the UK are affecting the block grant given to the Scottish Government

To see this is not the case, one must recognise that if the UK Government were to raise income tax rates and thereby increase tax raise tax revenues, Scottish taxpayers would benefit, however those additional revenues were spent:

- If the additional UK revenues were used to fund higher spending on comparable services in England, the Scottish Government would receive an increase in the underlying block grant (i.e. before the BGA) through the Barnett Formula;
- If the additional UK revenues were used to fund higher spending on 'reserved' functions such as defence or welfare, Scottish taxpayers would benefit in the same way that rUK taxpayers do;
- If the additional revenues are used to reduce the UK debt more quickly, Scottish taxpayers would benefit from the improvement to the overall UK Government finances through having less debt to service in future.

These effects offset the changes to the BGA, meaning the amount spent for the benefit of Scotland will likely be *broadly* the same as before the tax change. If the tax increases in rUK funded comparable spending that is devolved to Scotland, then the Scottish Government's budget would be broadly the same as before with a bigger underlying block grant and a bigger BGA. If, on the other hand, higher taxes in RUK were used to fund increases in reserved spending or deficit

reduction, the 'taxpayer fairness' principle set out in the SCA would imply that Scotland should either make an equivalent contribution in terms of raising its tax, or receive an equivalent cut in the Scottish Government block grant. Otherwise Scotland would be benefitting from the same level of spending on devolved services and higher spending on reserved services or deficit reduction, without any change in its tax effort.³⁴

Similarly, reductions in tax rates in rUK – which would necessarily feed through to Scottish people either via a reduced underlying block grant, lower spending on reserved functions, or higher debt interest payments – should be compensated for by a smaller BGA, to ensure that Scots continue to benefit from the same level of overall government spending should the Scottish Government leave its tax effort unchanged.

However, *broadly* the same does not necessarily mean *exactly* the same for two reasons.

The first relates to the precise way in which the BGA is indexed against rUK revenues. To see this consider a scenario where the UK Government has raised income tax rates in such a way that its revenues have increased by £10bn, which billion is equivalent to around £170 per person in rUK, or around 6.8% of tax revenues.³⁵ Assume further that it spends revenue on comparable/devolved services in England, generating additional money under the Barnett formula for Wales and Northern Ireland *and* Scotland. Scotland's share of this Barnett consequential is around £903 million, which is a population share of the equivalent increase in spending in England of £9.16 billion.³⁶

What about the BGA? Under the ID approach Scotland's BGA would rise by 6.8%. But because Scotland's tax revenues per capita are lower than in rUK, a 6.8% increase in Scotland's BGA would be less, in cash terms, than the additional block grant Scotland receives under the Barnett formula (which is based on Scotland's population share of 7% revenue growth in rUK). The Scottish Government's budget therefore increases a little (in this illustrative scenario by around £111million).

These issues arise because of the way that the ID method combines a change in the BGA that is based on *percentage* changes in revenues on one side, with a change in underlying block grant via the Barnett formula that is based on *cash-terms* changes in spending on the other. These do not necessarily exactly balance. Scotland benefits from an increase in spending elsewhere in the UK if this is funded through a devolved tax which raises less per person in Scotland than in rUK (such as income tax). Conversely, the Scottish Government's budget would fall a little if there were reductions in spending and the same tax in rUK. If,

³⁴ UK Government (2015) describes the rationale for these offsetting effect as follows: 'The tax deduction element of the funding model [...] needs to work alongside the Barnett Formula to ensure that increases in "rest of UK" tax do not fund higher spending in Scotland'.

³⁵ Based on 2013–14 values as estimated by HMRC.

³⁶ The other £840 million goes to Wales and Northern Ireland.

revenues per person from the devolved tax were higher in Scotland than in rUK, then these effects are reversed (Scotland losing from tax/spending increases, and gaining from cuts).

This means that the ID method does not fully satisfy the 'taxpayer fairness' element of the second 'no detriment' principle. Neither does the PCID method, although the situations under which Scotland ends up winning or losing from changes in taxes and spending in rUK are also affected by relative population change in this case. This is because relative population change feeds into the BGA under the PCID method but is not taken into account when determining the change in the underlying block by the Barnett formula.

Consider instead the LD method of indexing the BGA. If the UK Government raises an additional £10bn in revenue from a devolved tax to fund £10bn of additional comparable/spending in rUK, Scotland again receives a population share of this spending increase under the Barnett formula. But the BGA is also increased by a population share of the revenue increase in rUK. Because the increases in spending and revenues in rUK are the same size, the increases in the underlying block grant and the BGA are therefore identical – and the Scottish Government sees no change in its net grant from the UK Government. Thus, the LD approach *does* satisfy the 'taxpayer fairness' principle when changes in devolved taxes in rUK fund changes in devolved/comparable spending.

However, our analysis in Section 4 showed that LD indexation means the rate of growth of revenues in Scotland needs to be greater than in rUK if Scotland is to avoid losing out from tax devolution. We suggested that this seems inconsistent with the spirit of the 1st no detriment principle that neither government should lose from the 'decision to devolve a power'. In contrast, the PCID means that the Scottish Government's budget evolves as it would have in the absence of tax devolution of tax revenues per person grow at the same rate in Scotland as in England, which seems more consistent with this 1st no detriment principle.

Therefore the 1st no detriment principle ('no detriment from the decision to devolve') and the 2nd no detriment principle ('no detriment from subsequent policy decisions of the other government') are somewhat in conflict. BGA indexation based on percentage changes in revenue, and in particular the PCID method seem more consistent with the 1st, but do not fully satisfy the 2nd as long as the Barnett Formula remains in place. On the other hand, the LD approach satisfies the 2nd, but seems less consistent with the spirit of the 1st no detriment principle. How the UK and Scottish Governments might choose between these two approaches given this conflict is discussed in Section 7 of the paper.

The second instance where the 'taxpayer fairness' principle is unlikely to be exactly satisfied is when changes in taxes in rUK are linked to changes in spending on areas that are reserved – such as defence, or most areas of social security spending. In this instance, no simple method of indexing the BGA will necessarily exactly offset other changes in spending for the benefit of Scotland. To see this, take the example of the State Pension. Suppose the UK government

decided that it wanted to increase the State Pension, and wanted to increase income tax to do so. Because a slightly higher proportion of the population in Scotland are pensioners than in the rest of the UK, spending on the State Pension is higher per person than in the rest of the UK. Scots are therefore likely to benefit from a greater-than-population share of the gains from increases in the State Pension. On the other hand, they would contribute (in the form of a larger BGA and therefore lower net block grant) either a population share of the costs under the LD method, or a less-than-population share of costs under the ID method. Scotland would therefore enjoy higher overall levels of public spending even though its citizens are paying the same tax as before under either of these methods. The reverse would be true for increases in spending on reserved social security for which Scotland would receive less than a population share (such as Child Benefit – due to a lower fraction of the population being children –, or Housing Benefit – due to lower rents).

This means that one cannot expect to fully satisfy the ‘taxpayer fairness’ principle in all circumstances. Sometimes changes in taxes in the rest of the UK will have some knock-on effects to total public spending in Scotland (either devolved, or reserved). Scotland may gain or lose depending on which taxes and spending are changing, and whether they are being increased or decreased. Such effects are inevitable where some functions remain reserved to the UK Government.

6. Additional considerations for welfare devolution

As already mentioned, in addition to new tax powers and revenues, the Scottish Government will also become responsible for spending and policy in some areas of welfare (see Table 2, above). The employment programmes being delivered by DWP, mainly through the Work Programme and Work Choice, will also be devolved to the Scottish Parliament.

The Scottish Government's block grant will need to be increased to account for these additional devolved spending responsibilities. As is the case for devolved taxes, these initial BGAs will need to be indexed in some way to reflect the fact that spending on these benefits will likely vary over time (because of inflation, changing benefit caseloads, etc).

The issues involved in indexing these welfare BGAs are much the same as for taxes: one needs to think about the type of risks and incentives the Scottish Government should face. The same sorts of options as for the tax side are available: ID, PCID and LD could all be used. The same sorts of effects as discussed above are also evidence – although the signs of them are sometimes reversed. For instance, under PCID, Scotland would lose relative to basic ID if Scotland's population grew relatively more slowly (whereas it would gain for tax devolution). This is because the additional money Scotland gets with the welfare BGA would go up less quickly under PCID than under ID in such circumstances.

Across the ten benefits being devolved, per capita expenditure is significantly higher in Scotland than in rUK. This is partly because of Scotland's slightly older population structure (most of the devolved benefits are targeted at the older population), but it is mainly because of higher disability claimant rates in Scotland. Hence, when spending on these benefits *increases* in rUK, Scotland would do less well under LD than ID. This is because a population-based share of any increase in spending in rUK will be smaller than the same percentage increase in spending in Scotland. Over time, the LD would therefore lead to convergence between the amount given to Scotland per person to pay for disability benefits (the BGA) with the amount spent on such benefits per person in rUK – even if there was no convergence in the relative incidence of disability.

Conversely, if rUK spending on the devolved benefits was to *decline* over time, then the levels approach results in divergence between rUK spend per person and the Scottish BGA allocation per person. This is akin to the situation for public service spending under the Barnett Formula at present – and as with that these convergence or divergences could be accentuated or mitigated by relative population change (see Box 1, Section 2).

Other, more complex options have also been discussed in the context of welfare devolution. Rather than indexing relative to spending or spending per person in

the rest of the UK, could one index relative to spending per *person at risk*. This would mean account for differential demographic or social-economic change – such as the age structure of the population – which might affect the ‘need’ for spending on such benefits. Of course, including additional factors in the determination of the BGA, means that while the Scottish Government is insulated from spending risks associated with these factors, it also no longer faces financial incentives to address these factors. For instance, if the BGA for disability benefits is adjusted to account for the proportion of the population reporting ill-health, then the Scottish Government would not have the financial incentive to reduce disability spending by reducing the prevalence of ill-health (as doing so would lead to offsetting changes in the BGA). As always there is a trade-off between insurance against risks and incentives.

In addition, simpler definitions of the population ‘at risk’ would run into problems. Consider the case of linking the BGA to changes in the older (say 50+) population, on whom a large fraction of the spending to be devolved is targeted (due to higher rates of disability). First, although the benefits being devolved are broadly targeted at older people, some spending goes towards younger adults and children. Second, the focus of benefit spending could theoretically shift in the future: recent years have seen disability spending become less focused on older people, for instance. Third, it is not clear what the rationale would be for introducing an element of spending need assessment into the allocation process for welfare spending, when this is absent from other aspects of the Scottish Government’s block grant.

It therefore seems sensible to restrict attention to the simpler sort of adjustment methods based on overall spending changes, or spending changes per person, that we focus on in this paper. In doing this, there may be an a priori case for consistency between the approach chosen to index the BGA for devolved welfare as there is for devolved taxes. If the Scottish budget is exposed to the risk/reward of differential population growth on the tax revenue side, then it seems natural that it also be exposed to the risk/reward on the welfare expenditure side. Departures from this symmetry could easily become attempts to game the system (each Government choosing the option that benefits them the most in each case). However, as discussed in Section 7, there may be a rationale for different approaches being taken if the UK and Scottish (and other devolved governments) undertake a more comprehensive reassessment of the kinds of risks and incentives that should be pooled across the UK and those that should not be.

In addition to areas that will be fully devolved, the Scottish Government will also gain powers in areas where funding is not being devolved and no BGA will be given. For instance, it will also have the ability to vary elements of Universal Credit (UC), notably the under-occupancy charge, the local housing allowance rates, and eligible rent (as well as administrative issues like the frequency of payment). It will also have the power to top-up benefits, and create new benefits in certain areas.

Scotland will continue to be insulated from any changes in relative spending on these benefits in Scotland if it chooses to make no changes. But the Scottish Government will have to bear the budgetary effects if it utilises these powers. If it were to increase the generosity of UC housing elements, say, it would have to make a transfer to the UK Government to pay for this (because the UK Government will continue to administer UC).³⁷

To see the difficulties this might cause, consider an example policy: suppose the Scottish Government said that rather than setting housing allowance rates to cover the cheapest 30% of properties, they wanted to cover the cheapest 40% of properties. Calculating the mechanical cost of this – holding the behaviour of the tenants and landlords fixed – should, in principle be straightforward. The Scottish Government could transfer an estimate of this amount to the UK government to pay for this policy.

But such a change may also have behavioural effects and the Smith Commission's principles suggest the Scottish Government should also bear these too.³⁸ For instance, tenants may decide to rent more expensive properties as a result, or landlords may decide to increase rents up to the new threshold. The higher levels of UC housing support many tenants would be entitled to would be tapered away over a bigger range of income, reducing work incentives, and potentially reducing the amount tenants work. This would have knock on effects for benefit spending and both devolved and non-devolved tax revenues. Each of these behavioural could push up the cost of the policy beyond the mechanical (no-behavioural response) estimate. But evidence on how big these responses might be is limited. There would therefore be wide margins of error around any estimates of the full costs of such a reform that attempt to account for these behavioural effects. Things would not improve much even after the policy was introduced. If rents or working behaviour changed, is this the result of the policy change, or a result of other underlying change in the Scottish housing or labour markets?

There would therefore be significant scope for (reasonable) disagreement between the two governments about just what the costs of such a policy would be, and what the Scottish Government should transfer to the UK government to pay for the policy. This might mean the system becomes bogged down in difficult negotiations and political conflicts that could undermine its stability. Resolving such issues would seem to inherently require ongoing ad-hoc adjustments and negotiation to implement, breaching one of the Smith Commission's principles.

Other tricky issues may arise in other areas of welfare spending due to interactions between devolved and non-devolved parts of the social security system. For instance, eligibility for Personal Independence Payments (PIP) – one

³⁷ Conversely, the UK government would have to make a transfer to the Scottish Government, if the Scottish Government decided to reduce the generosity of these elements.

³⁸ The principle of 'economic responsibility' says that the Scottish budget should bear in full any costs or gains associated with Scottish Government policy. Furthermore, the second 'no detriment' principle says that any knock on effects for the revenues or spending of the other government by the government making a policy change should be compensated for.

of the benefits which will be devolved to Scotland – are used to determine whether people are also eligible for a range of things that will not be devolved – such as disability premiums in Housing Benefits, Tax Credits and Universal Credit, and exemptions or deductions from Vehicle Excise Duty, for instance. Any change in the eligibility criteria for PIP after devolution, or a change in the nature of the benefit, could therefore have knock-on effects for UK Government spending on these benefits. But how much of an effect? If the medical tests were changed in such a way that meant more people receive PIP or its Scottish replacement we would need an estimate of how many more. But if the old tests were no longer being conducted, we would not know how many would have satisfied those, and therefore how many additional people passed the new tests. Again, there seems scope for significant (and reasonable) disagreement in this context.

There will also be other types of interactions between devolved benefits and non-devolved benefits. For instance, Carers Allowance reduces eligibility for other out-of-work benefits such as Income Support or Universal Credit. Changes to Carers Allowance by the Scottish Government may therefore have knock-on effects for UK government spending on other benefits. As above, taking into account the mechanical interaction effects is likely to be more straightforward and practical than attempting to incorporate behavioural effects too.

7. Difficulties with Smith's 'compensation principle'

Section 6 shows that the idea that 'compensation payments' should be paid for any knock on effects on the other government's revenues or spending (required by the Smith Commission's principles of 'economic responsibility' and 'no detriment') runs into difficulties on assessing the size and nature of behavioural response to policy changes in the context of welfare devolution. But these principles and associated difficulties also affect tax devolution (and, could be applied to existing devolved public services too).

For instance, because UC (a non-devolved benefit) is assessed on the basis of income *net of tax*, changes in Scottish income tax rates or thresholds could impact on UK government UC spending in Scotland. Changes to Scotland's income tax, especially to the top rates, might also have large and complex behavioural effects that impact on UK government tax revenues. For example, individuals might respond to an increase in Scottish income tax in a number of ways:

- They might reduce the amount they work, in which case the amount of National Insurance (NI) they pay would also be reduced. Because NI is not being devolved, the UK government would see a fall in its revenues.
- They might shift their income from earnings into dividends (on which the lower UK rate will continue to apply). This would reduce Scottish Government revenues but boost UK government income tax revenues.
- They may also move to the rest of the UK, again reducing Scottish Government revenues but boosting UK government revenues.

Working out how much the Scottish Government had to compensate the UK government or vice versa for these various effects would be inherently difficult, with much room for disagreement over the methods and assumptions used. Estimating models of how people respond to tax changes is notoriously difficult and even once an agreed method is chosen, subject to wide margins of error: for instance, the 'statistical confidence intervals' around HMRC's estimate of how much the 50p tax rate raised in the UK go from less than nothing (i.e. costing the Government money) to over £4 billion a year.³⁹

Similarly UK government decisions may affect Scottish revenues. Suppose the UK increased the standard rate of VAT, leading people to shift their spending from goods subject to the standard rate of VAT (say, hot take-away food) to goods not subject to VAT (say, cold take-away food). Scotland's 10 percentage points of assigned VAT revenues would therefore be lower. But calculating how much lower – and therefore how much compensation should be paid by the UK government – would again be difficult.

³⁹ HMRC (2012).

Looking ahead, it will therefore be important to recognise that such compensating transfers will only be practical if, in general, attention is restricted to direct, mechanical revenue effects, which can be estimated more easily. This would not be out of line with international practise. For instance, the United States does not require transfers between the States and Federal Government when either changes income tax, despite the potential for behavioural knock on effects on the other's revenues.

If a 'shared understanding' of the type and scale of behavioural factors can be agreed in particular instances, especially where such behavioural effects seem likely to be large, then it would be worth trying to take these into account. Because if such behavioural effects are not taken into account, then the governments do not bear the full cost of their policies themselves, and may therefore face distorted policy incentives. Of course, the cases where such behavioural effects are most significant, are also likely to be the cases where the amount of cash at stake from the precise choice of assumptions to make are also greatest, which may not be conducive to developing such a 'shared understanding'. This means it will not be surprising if this element of the fiscal framework leads to difficult disagreements. The challenge will be to prevent these disagreements from undermining the working relationship between the UK and Scottish governments and undermining the stability of the new fiscal framework.

8. Broader discussion and conclusions

This paper has focused on one of the key issues in implementing further devolution of tax and welfare to Scotland and an integral part of the new fiscal framework being negotiated: how to calculate the resulting block grant adjustments (BGAs).

In doing this it has emphasised the different risks and incentives that different BGA options would entail, and has assessed to what extent they satisfy the principles the Smith Commission said Scotland's new fiscal framework to adhere to. We have found that the precise way in which the BGA is indexed can have quite significant effects on the Scottish Government's budget – and on the type of risks Scotland faces. We have also argued that while Smith's principles may seem reasonable in the abstract and can play a role in guiding the new fiscal framework, they cannot be implemented fully in practise. Several issues seem worth highlighting as we conclude.

First, we have seen two of the 'no detriment' principles seem to conflict with each other given the current way of calculating the underlying block grant using the Barnett Formula. The Levels Deduction (LD) approach seems to cope best with the 'taxpayer fairness' principle that public spending in Scotland should neither rise nor fall when taxes change in the rest of the UK (rUK). But because it would require a faster percentage rate of growth in Scotland than in rUK to avoid the Scottish Government's budget falling below what it would be in the absence of devolution, it seems less in accord with the spirit of the principle that there should be 'no detriment from the decision to devolve'. In contrast, the Per Capita Indexed Deduction (PCID) method has the opposite properties. In particular, because this method does not fully align with the way the Barnett Formula works on the spending side, public spending in Scotland will almost inevitable increase or decrease a bit when taxes change in rUK.

If the different approaches satisfy different Smith principles, which should be chosen? That would require judgement on what is the most important. In making these judgements one might note that because tax rates increase and decrease, one might expect the gains and losses Scotland faces under the PCID method when taxes change in rUK to broadly offset each other over the long run.

Recognition of this is important for another reason: When a change to a tax in rUK that is devolved to Scotland is used to fund changes in reserved areas of spending like defence or pensions, no system of BGAs will fully satisfy the 'taxpayer fairness' principle. The only sure fire way to resolve this problem would be to say that the UK government cannot vary taxes that are devolved to Scotland to fund (or in response to) changes in spending on reserved matters. But this does not seem practical given that there is no direct link between particular tax changes and particular spending changes. Nor does it seem sensible to rule out the UK government from varying its single largest source of

revenue – Income Tax in rUK – to fund changes in such key reserved areas as defence or welfare.

Second, in deciding which method to implement, one should think about the potential long run effects of these methods in the context of a history of long-term relative population decline in Scotland. Our analysis has shown just how big a difference there can be.

For instance, we have estimated what would have happened if each of these approaches had been in place between 1999–00 and 2013–14 to get an idea of just how big a difference they could make. While these figures are approximate and refer to the past rather than the future, they show these differences can be substantial. Relative to the LD method, the PCID method could have resulted in the Scottish Government’s budget being more than £1.5 billion higher a year after 14 years. This is quite sizeable numbers in the context of a block grant to Scotland equal to around £30 billion a year in 2013–14.

In the medium term, the ID method – whereby the BGA increases in line with the rate of growth in total comparable revenues per person in rUK (rather than revenues per person) – would give Scotland a budget somewhere between what it would get under PCID approach and the LD approach. It is less generous than the PCID approach because it does not take account of the fact that Scotland’s population typically grows less quickly than that of rUK – which would act as a drag on overall revenue growth. But it is more generous than the LD approach.

Eventually though, if relative population decline continues, Scotland would start to do less well under the ID method than the LD method. Indeed, because the ID method *never* gets updated to reflect the fall in Scotland’s relative population, it can eventually imply a *negative* budget for Scotland if one far enough in the future. Clearly such an outcome would never be allowed to come to pass. But it illustrates that the ID method would not represent a sustainable long-term compromise between the PCID and LD methods that, in the short term, would be most beneficial to the Scottish Government and UK Treasury, respectively.

A third issue to highlight is a lack of consensus, or even debate, about the type of fiscal risks and incentives Scotland (and other devolved governments) should face. Should Scotland face risks associated with relative population change? Should it face risks associated with differential demographic or economic change, over which the Scottish Government might have only limited control? Without such a debate it is difficult to recommend a particular form of BGA as these are intimately linked to the type of fiscal risks and incentives a Scottish Government will face.

The Smith Commission said that the UK government should bear ‘economic responsibility’ for its own policies, and the risk of any shocks that affect the whole of the UK, and the Scottish Government should bear the ‘economic responsibility’ for its policies, it said nothing about who should bear the underlying revenue or spending risks in Scotland. The methods we have outlined

in our paper – which are perhaps the most widely discussed methods –, would mean that the Scottish Government bore all such risks associated with devolved revenues and spending. But one might want to pool at least some of these risks across the UK's fiscal union.

The key difficulty in doing this though is isolating the effect of government policies – the consequences of which, as far as possible, you would want the government who made them to face – and other underlying risks that you might want to pool. But modelling the behavioural impact of policies is notoriously difficult, so separating out these effects exactly is nigh on impossible. Instead, simplifying assumptions have to be made – such as ignoring most potential behavioural effects of policies. Making such calculations on a case by case basis might create the opportunity for repeated argument. Instead, a more systematic approach might be more sustainable. Other countries – such as Canada, Germany and Australia - do this by trying to assess the spending needs or the revenue-raising capacity of different parts of their country, and then calculating transfers to offset some or all of the differences. A similar approach was traditionally used for local government grants in the UK. Implementing such an approach for the financing of devolved governments would require more significant changes to the fiscal framework than the Smith Commission advocated though.

This brings us to the issue of whether more fundamental changes to the devolved fiscal framework than advocated by the Smith Commission would actually be desirable. Could it address some of the tricky issues we have identified here?

For instance, the LD method for indexing BGAs ensures consistency with the Barnett Formula but seems likely to lead to a more severe Barnett Squeeze that seems incompatible with the 'no detriment' principle. On the other hand the PCID is inconsistent with the Barnett Formula and means the 'taxpayer fairness principle' is not fully satisfied.

If the PCID method were chosen, it could also be argued that Scotland's fiscal framework was asymmetric. It would be fully insulated from relative population change in relation to devolved tax revenues, where relative population decline would otherwise mean it would lose. But because of the way the Barnett Formula works, it would be exposed to such risks on the public spending side, where relative population decline is actually of benefit to Scotland (because public spending has to be shared among fewer people). This could be seen as somewhat unfair.

One solution would be to reform the Barnett Formula so that rather than allocating Scotland and the other devolved nations a population-based share of the cash terms change in spending in England as is the case now, it allocated them the same percentage change in spending per person. This would make the formula on the spending side consistent with the PCID approach to indexing the

BGA. It would therefore be possible to satisfy both the 'no detriment' and 'taxpayer fairness' principles.

Of course, the Smith Commission also said that the Barnett Formula should be retained, perhaps because Scotland is perceived to do relatively well under it at present.⁴⁰ In the long term though, the Barnett Formula implies some degree of convergence in spending per person in the devolved nations with the level in England. Scotland might not continue to do as well under the Barnett Formula in future then. Could this give an opportunity for more fundamental reform?

Replacing the Barnett Formula with a formula akin to the PCID approach would stop further convergence if Scotland's revenues per person kept pace with those in rUK. This Percentage Per Capita (PPC) formula could be implemented with or without an initial needs or revenue capacity assessment to determine the appropriate level of block grant to give to Scotland and the other devolved nations.

However, it is important to note that if, after an initial assessment, ongoing needs or revenue capacity based assessments were used to update the underlying block grant over time, that this would necessarily blunt the fiscal incentives that tax and welfare devolution otherwise entails. If Scottish Government policy led to relatively poor economic performance, for instance, the falls in devolved tax revenues associated with this might be offset by a bigger block grant following a higher needs assessment (due to higher poverty, say). There is inherently a trade off between insurance and equalisation on one side and incentives on the other, which needs to be recognised in any discussion about replacing the Barnett Formula with something more needs-based. Recent devolution debates, at least in the case of Scotland, have been almost all about providing stronger fiscal incentives. Moves towards *ongoing* needs assessments would be a move in the opposite direction.

That is not say such a move would be wrong. That depends on ones view about what the UK's fiscal union is for. What risks should be pooled and what ones should not. And where the balance between incentives, insurance, and equalisation in devolved government finance should be. There would be real merit in a proper debate about these issues.

Finally we should add that this paper should not be seen as a criticism of the work of the Smith Commission. Its remit in relation to the fiscal framework was actually rather limited and it faced the difficult task of coming up with a package that political parties starting from very different positions could agree on in just a couple of months. Resolving all the issues we have identified in this paper will require a more fundamental reassessment of the funding regime for the UK's devolved governments than the Smith Commission was tasked with.

⁴⁰ Holtham (2010).

References

- Bell, D. and D. Eiser (2014), *How should the Barnett formula be adjusted to reflect devolved taxes?* (available at: <http://scotfes.com/2014/11/20/how-should-the-barnett-formula-be-adjusted-to-reflect-devolved-taxes-a-question-of-allocating-risk-wonkish/#more-890>).
- Commission on Scottish Devolution (2009), *Serving Scotland better: Scotland and the United Kingdom in the 21st century*
- Cuthbert, J. (2001), *The Effects of Relative Population Growth on the Barnett Squeeze*, Fraser of Allander Institute Quarterly Economic Commentary, Vol. 26, No. 2
(<http://www.cuthbert1.pwp.blueyonder.co.uk/papers%201/Barnett%20Rel%20Pop%20Note%2016%2012%2000.doc>)
- Cuthbert, J. (2015), *Written Evidence to the House of Lords Economic Affairs Committee Enquiry on the Devolution of Public Finances* (available at: <http://data.parliament.uk/writtenevidence/committeeevidence.svc/evidencedocument/economic-affairs-committee/the-devolution-of-public-finances-in-the-united-kingdom/written/21826.pdf>)
- HM Government (2015), *Scotland in the United Kingdom: an enduring settlement*, (available at: <https://www.gov.uk/government/publications/scotland-in-the-united-kingdom-an-enduring-settlement>)
- HMRC (2012), *The Exchequer effect of the 50 per cent additional rate of income tax*
- Independent Commission on Funding & Finance for Wales (2010), *Fairness and accountability: a new funding settlement* (available at: <http://gov.wales/funding/financereform/reports/fairness/?lang=en>)
- Phillips, D. (2014a), *Business as usual? The Barnett formula, business rates, and further tax devolution*, IFS Briefing Note No. 155, Institute for Fiscal Studies (available at <http://www.ifs.org.uk/publications/7442>)
- Phillips, D. (2014b), *The Smith Commission's proposals – how big a change to they represent? And what questions remain to be addressed?*, IFS Briefing Note No. 157, Institute for Fiscal Studies (available at <http://www.ifs.org.uk/publications/7484>)
- Scottish Government (2015), *Government Expenditure and Revenue Scotland 2013–14*, (available at <http://www.scotland.gov.uk/Topics/Statistics/Browse/Economy/GERS>)
- Smith Commission (2014), *Report of the Smith Commission for further devolution of powers to the Scottish Parliament*, (available at <https://www.smith-commission.scot/smith-commission-report/>)