Starting From Scratch?
A New Approach to Subnational Public Finance
with Application to the UK

David Bell*
University of Stirling

Willem Sas†
University of Stirling & KU Leuven

Preliminary version - Do not cite

Abstract

The systems in place in most countries to finance public goods and services provided by lower-level governments, are flawed. They are often perceived as unfair, are opaque and lack legitimacy. Their attempts to promote local accountability through the transfer of tax-raising powers are insufficient, or ineffective. This paper therefore explores a set of equitable, transparent and intuitive guidelines to design a new fiscal framework for regional public finance. It builds these from a set of principles rather than an ad hoc process dictated by political expediency, recognising that any fiscal arrangement will be subject to a participation constraint. Doing so, we account for the possibility that one or more regions is willing to leave the federation or call for additional phases of devolution, if this has sufficient popular support. To "soften" this participation constraint, and within certain bounds, our framework allows regional jurisdictions to unilaterally decide how much tax autonomy and fiscal responsibility to take on, which offers a fresh perspective on regional public finance.

This draft: 20th November 2018.

---

*Economics Division, University of Stirling. E-mail: d.n.f.bell@stir.ac.uk

†Department of Economics, KU Leuven, Belgium & Economics Division, University of Stirling, UK. E-mail: willem.sas@stir.ac.uk

1   Introduction

The systems in place in most countries to finance public goods and services provided by lower-level governments are frequently flawed. They are often perceived as unfair, are opaque and lack legitimacy. Their attempts to promote local accountability through the transfer of tax-raising powers are insufficient or ineffective, and their asymmetry is more often than not defined by political expediency alone.

In the UK, for example, the main element of the financing scheme has been the Barnett formula, which was introduced in 1979 as a stopgap measure. It has since been used to determine funding allocations to Scotland, Wales and Northern Ireland. After the extension of political devolution following the independence referendum of 2014, the Scottish Parliament has acquired new tax and welfare powers. Some more limited tax devolution is planned for Wales and Northern Ireland. When implemented, different taxes will be devolved to different parts of the UK to differing extents, and different versions of the Barnett formula will apply to allocate grants to the devolved institutions. Recent fiscal innovation also provides for tax sharing for the first time, though this mechanism is intended only for Scotland in the first instance. Although these arrangements may solve some short-term political problems, they cannot form a stable long term solution to the problem of the distribution of fiscal powers and equalisation payments: their lack of transparency and of equity will eventually prove their undoing.

Zooming in on 24 OECD countries in Fig. 1 below, it becomes clear why having such a defunct fiscal framework in place to organise regional finance can be detrimental to overall welfare. For starters, the revenue streams needed to finance devolved spending are substantial, as lower-level governments take on around 30% of total public spending on average. Moreover, lower-level governments only partially finance their expenditures with their own revenue-raising instruments such as taxes or user fees. Indeed, own revenues cover about half of local and regional spending, of which taxation only picks up two thirds. Finding the appropriate split between grant finance and lower-level revenue raising therefore becomes crucial as well.

This paper therefore explores a set of more equitable, transparent and intuitive guidelines to design a new fiscal framework for regional finance. We discuss how it might be applied to the UK nations. Our framework is built from a set of principles rather than an ad hoc process dictated by political expediency. It also recognises that any fiscal arrangement will be subject to a "participation constraint", which will be perceived as more or less stringent depending on the regional entity in question. Hence, and unlike most of the fiscal federalism literature, our paper explicitly recognises that continuation of the "federation" – in this case, the UK Union – cannot be taken for granted. Disintegration of the federation may

Figure 1: Regional expenditures and (tax) revenues in % of total public spending
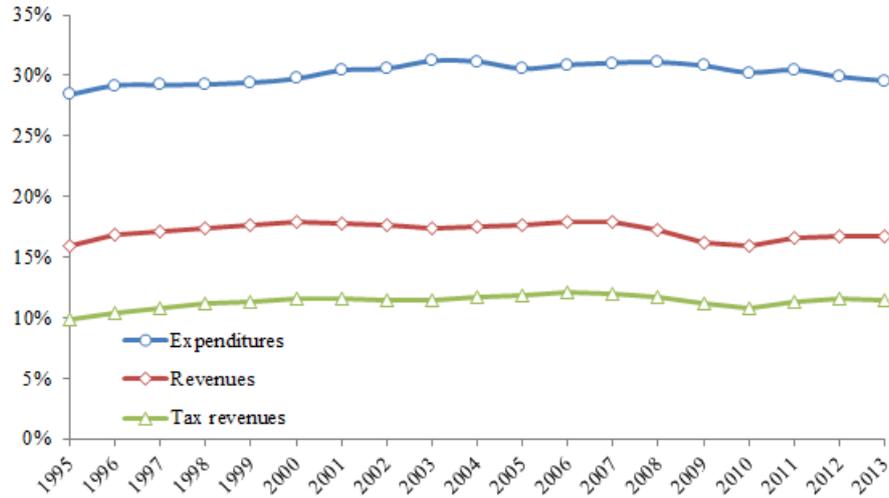
occur due to dissatisfaction with the fiscal framework itself, whose role in fundamental to determining regional levels of taxation and the provision of public services. Indeed, real or perceived unfairness in the distribution of public resources across subnational jurisdictions undermines confidence in any federation. Aside from unfairness in resource distribution, willingness to dissolve the federation may also reflect a trade-off between economic advantage and the perceived political benefits of autonomy for the subnational unit. We therefore examine the process and institutional arrangements needed to support transition to a new fiscal framework, as well as to sustain it. Doing so, our framework allows regional jurisdictions to unilaterally decide how much tax autonomy and fiscal responsibility to take on, which, as a blueprint for regional public finance, is unprecedented. Of course, to guarantee the stability of national public finances as a whole, such unilateral regional discretion will inevitably have to be bound by nationally agreed upon thresholds. In that sense we rather propose to "soften" the participation constraint, rather than doing away with it altogether, to make the federation more viable and resilient to shocks.

Importantly, and although we apply our approach to the UK setting in what follows, the principles-based nature of our approach makes it particularly flexible and generalisable. In this sense it can easily provide guidelines for reform in other federations as well.

2 Principles

Economists generally agree on which functions of government to devolve to lower levels of government, and which competences to keep at the central level. As a general rule, powers that can be tailored to local needs or preferences are best organised on the regional or municipal level (Oates, 1999). Discretion over education, welfare, health care, infrastructure, transport, or labour market policies is therefore often assigned to substate governments. Conversely, whenever devolution leads to large losses in scale economies or coordination failures between local governments, policy levers are best kept in the hands of the highest level of government (Oates, 1972). As a result, areas such as defence, macro-economic stabilisation, social security, trade or foreign policy tend to remain the prerogative of central governments (Boadway and Shah, 2009).

Unfortunately, there is no similar consensus regarding the revenues that pay for devolved functions. Finding the optimal arrangement almost always boils down to striking the right balance between a variety of interrelated principles, which vary in importance according to local conditions. We describe the most important principles below.

2.1 Efficiency & Accountability

Too much tax autonomy among substate governments can lead to tax competition which reduces the gains from local control over spending (Wilson, 1986; Keen and Kotsogiannis, 2002; Wilson and Wildasin, 2004). Differential tax rates cause resource misallocation because the market, in the case of our example, the UK market, becomes distorted. For example, if Northern Ireland was given the power to reduce corporation tax to enable it to compete with the Republic of Ireland, businesses from the rest of the UK might move to Northern Ireland simply to take advantage of a lower corporation tax rate. And if the rest of the UK reduces corporation tax rates in response inefficient relocations may be avoided, but revenue reductions resulting from such tax competition may reduce public service provision below welfare maximizing levels.

In the general case, central government can simply provide cash transfers to substate governments instead of granting them tax autonomy. This avoids efficiency losses. The downside of this strategy is that voters will be less able to hold local politicians to account because the link between spending and revenue raising is absent (Boadway and Tremblay, 2012). Indeed, the transparency of the fiscal framework is undermined in this case, so that voters have diminished incentive to understand the connection between revenue raising and public spending (Lockwood, 2006). Poorly informed voters are also more likely to acquiesce to excessive regional borrowing if substate governments are allowed to raise debt on financial

markets. This may lead to defaults if substate governments expect to be bailed out by higher levels of government, which in itself is more likely the more they are dependent on grant-based finance.[1]

2.2 Responsibility

There is clearly a trade-off between the efficiency losses of tax competition and the benefits of accountability in relation to tax powers.[2] But there is also the 'responsibility' effect of fiscal incentives, which relies on a positive feedback loop between 'good' policies and consequent financial rewards (Weingast, 1995; Jin et al., 2005; Weingast, 2009). The idea is intuitive: if substate government's policies serve to boost locally-raised revenue, the incentive is there to continue or extend these policies as local growth generates ever more tax revenues. And since local growth often promotes local welfare as well, politicians are held 'responsible' for their good policies through the tax revenues they generate. This is important in settings where democratic accountability is weak, and politicians are corrupt or inept. The degree to which substate governments retain the financial rewards from their policies – what we will call the 'retention rate' – will be vital here. Importantly, grants can also be made 'responsibility-sensitive', linking the retention rate to regional growth or other performance-based measures such as unemployment rates or schooling levels.

2.3 Horizontal Equity

Crucially, even if local politicians are incentivised to take local welfare into account, the nationwide welfare effects of such a 'responsibility-sensitive' fiscal framework may be less beneficial. When education or welfare policies are devolved for example, and partially financed on the basis of (economic) performance as suggested above, underperforming regions will have fewer funds at their disposal to keep schools and welfare programs equivalent to those in more prosperous regions. The notion of 'horizontal equity' — generally understood as the equal treatment of otherwise equal citizens by the government, regardless of their locality — is violated as a result. To compensate for such imbalances, an equalisation mechanism is usually established to realign interregional differences in fiscal capacities (Boadway and Shah,

---

[1]Aside from a more detailed description of both theoretical arguments regarding regional borrowing, an empirical validation is given by Rodden (2002), Rodden et al. (2003), Blankart and Klaiber (2006), Bordignon and Turati (2009) and Baskaran (2012). See also Foremny (2014) for evidence on the positive effect of tax autonomy on regional fiscal deficits. Note that some degree of regional borrowing discretion should be guaranteed in any case when taxation is devolved, to smooth regional-specific shocks to the fiscal base. Especially when equalisation mechanisms are not in place to provide interregional insurance.

[2]Other arguments, which rely more on the benefits from economies of scale, support the case for centralised taxation. Harmonising tax base and rate structures at the central level economises on administrative costs, as compliance and collection costs are reduced (Oates, 2005).

2007).[3]

However, such equalisation payments reduce the 'retention rate' and therefore erode the fiscal independence and responsibility of local politicians. Also, equalisation transfers between substate jurisdictions are often politically controversial, especially if interregional inequality (e.g. in income per capita) is persistent.[4] In what follows, we therefore aim to design funding instruments that can uphold fairness in a less blunt and contentious way, without eroding political responsibility too much. Doing so, we trace the issue of horizontal equity back to more normative grounds and specifically to the notion of 'equality of opportunity'. This notion hinges around consideration of those aspects of existence that can be considered beyond an individual's control, and which therefore might be compensated for by public insurance with the objective of maintaining equal opportunities. In contrast, one might expect that individuals should experience the gains (or losses) of actions for which they can justifiably be held responsible (Cohen, 1989; Roemer, 1998). Translated to our setting, our argument is that if local electors cannot reasonably be held responsible for adverse (or beneficial) local outcomes, then it is justifiable for national government to offer some compensation (or to extract rent). In contrast, no compensation should be paid for outcomes that are a clear consequence of local policy-making.

2.4   Continuity & Predictability

The longevity of inter-governmental fiscal arrangements is contingent on continued participation by central and substate governments. New arrangements will be required if national or substate governments fail to adhere to fiscal rules or if external shocks undermine compliance. substate governments may quit the polity if their electorate takes the view that autonomy is preferable to compliance with fiscal agreements. Obviously, such a decision would also have monetary implications, but we solely focus on the fiscal issues in this paper. Also, if substate governments anticipate eventual breakdown, their incentive for entering a permanent fiscal agreement will be substantially weakened. However, if central government can credibly commit to compliance, substate governments will be more willing to join. In turn, this implies a need for central government to behave in a predictable fashion, as set out in the agreement, and for transparency around the information required for its implementation, given that substate governments will continually assess compliance. Most countries address these issues through an institutional mechanism that is independent of central government, such as

---

[3]Much like a centrally organised tax-benefit system, equalisation mechanisms also provide insurance against region-specific, cyclical shocks which would otherwise have increased local costs or reduced revenues. Empirical evidence of interregional risk-sharing is e.g. given by Asdrubali et al. (1996) for the US, in Andersson (2008) for Sweden, and Arachi et al. (2010) for Italy.

[4]Sas (2017) lastly, shows that standard equalisation schemes also fail to mitigate tax competition in realistic settings where taxes are raised on an ad valorem basis.

a fiscal council.

3 A new fiscal framework

Combining the principles touched upon above, an 'ideal' framework of regional public finance would implement a positive feedback loop between political performance and (tax) revenues, whilst remaining sufficiently transparent to voters and robust to competitive pressures between subnational governments as well as credibility challenges from all actors.

3.1 Tax autonomy

A first option would be to devolve those tax instruments which are to some measure linked to economic activity, and are raised on a relatively inelastic – yet salient – tax base. Decentralised arrangements for inheritance tax and property taxes would be obvious candidates for the UK, in the sense that these automatically funnel the proceeds of growth-enhancing policies back into local revenues, are highly visible, and are raised on less mobile tax bases. Thus, if substate governments successfully promote local growth, they will reap the rewards through higher taxes, which can then be used to increase re-election chances. The increased visibility of policy interventions will also improve electoral outcomes, if the policies being pursued are sound. As a result, because economic activity can often be linked to higher levels of welfare, local politicians are held accountable and responsible for their actions, to the benefit of their electorate.

Note that tax sharing, where revenues from a specific tax base are shared between national and substate governments, apparently provide an alternative mechanism for increasing regional tax autonomy. However, tax sharing can be implemented in many ways and these vary considerably in the extent to which they genuinely extend substate fiscal autonomy (Blöchliger & King, 2006). For example, the UK government has agreed to share half of VAT revenue raised in Scotland with the Scottish Government from 2019. This will significantly increase tax revenues retained in Scotland, but will have little effect in extending the fiscal powers of the Scottish Government, given that it will control neither the VAT tax base, nor the rates at which is applied. Moreover, as it is difficult to link measures of local social and economic performance to local VAT revenues, variations in these revenues will not incentivise local politicians to perform better. The 'responsibility feedback loop' as described above, will fail.

3.2 The need for a new grant mechanism

As discussed above, devolving tax powers to lower levels of government will almost always reduce horizontal equity between regions. The more a substate government has to rely on its own fiscal base, the more interregional differences in fiscal capacity – due to a variety of factors, often beyond the control of politicians – come to the surface. This then trickles down into the quality of local public services, leading to different treatments of otherwise equal citizens across localities. Equalisation mechanisms to correct for such imbalances are generally imperfect, in the sense that they are usually rather blunt instruments of explicit interregional redistribution. They often foment political backlash, but also undercut the retention rate of public finance which is important in establishing the responsibility feedback loop.

Moreover, the responsibility effect of tax autonomy tends to be rather blunt. The point of the responsibility feedback loop is to reward local politicians for policies that increase local welfare. But what are such policies, and how do they interact with the evolution of tax revenues? Boosting economic growth is a sensible policy objective for substate governments with tax autonomy: it expands the fiscal base, which in turn provides local policy makers with additional tax revenues. But if economic growth is based on the attraction of foreign investment in polluting industries or on developers who run down natural resources, welfare may not be increased even though tax revenues are. Timing matters too, since investments in, for example, education, skills and infrastructure are likely to have medium to long term effects on local growth and prosperity, but may be less attractive to current politicians because they do not offer an immediate political return via the fiscal base. Although economic growth may seem an obvious objective of responsible policy-making, it may be necessary to consider more refined and desirable criteria to direct local fiscal incentives towards longer term objectives that enhance welfare.

To remedy both shortcomings of full tax autonomy — in that it erodes horizontal equity and offers a blunt measure to boost local welfare — without losing too much of its advantages, we propose a new grant mechanism. It allows for a more subtle, flexible and intuitive interpretation of horizontal equity, and also takes into account various responsibility criteria to nudge local politicians in the right direction. It would, however, always operate alongside a substantial degree of devolved tax autonomy to preserve the transparency benefits of the latter, but offsetting its deficiencies. Subnational governments should therefore also be given some discretion over borrowing decisions to smooth region-specific revenue shocks. In the UK, these might extend beyond the provisions available through the current fiscal framework

agreements in place for Scotland and Wales.[5]

3.3 Designing a new grant mechanism

In designing a new grant mechanism, we separate outcomes for which local politicians can and should be held responsible, from outcomes which are beyond their control because of geographic and demographic or other circumstances, often historically defined. Hence, we propose that regional public finance combines both 'responsibility' and 'compensation'. This concept draws on the axiomatic formulation of both principles developed by Flearbaey (2008), and is founded on an equality of opportunity ethical framework. Our proposal is therefore that the fiscal framework for substate governments should explicitly and transparently combine a compensation and a responsibility element, in a relatively simple and understandable formula. The compensation element will reflect equity considerations by reducing the disadvantages that citizens face due to the circumstances of the area in which they happen to live. The responsibility mechanism rewards authorities that drive favourable outcomes in policy areas over which they have discretion. Lastly, subnational governments have the authority to decide on the relative importance of responsibility and compensation elements in their revenues.

3.3.1 The formula

In more concrete terms, the grant scheme will be linked to performance-based factors relating to local welfare such as mean per capita income, measures of inequality, labour market outcomes or capital formation (tangible and intangible, social and natural). This comprises the responsibility mechanism. The compensation mechanism will be determined by needs-based factors comprising demographic, geographic and other indicators.

Both mechanisms are then rolled into the same formula expressed by Eq. (1) in Section A of the appendix, spelling out the way in which an annual lump sum G^t assigned for regional spending is divided across substate governments in a given year t. To illustrate, suppose the grant mechanism would only finance education. The lump sum G^t would then cover the total amount deemed sufficient for spending on education nationwide. This could be actual total spending at the time of introduction of the new grant scheme, or some other agreed figure. Following its introduction, the lump sum would increase over time based on pre-defined indices such as inflation and/or economic growth. However, this only keeps the lump sum G^t from eroding in real or nominal terms, but does not subdivide it across the substate governments. This is where the compensation mechanism comes in, since needs-based indicators will be

---

[5]Scotland for example, is allowed to borrow to make up for a forecast shortfall in revenues – when Scottish GDP growth is below 1% and at least 1 percentage point less than UK GDP growth. Since the correlation between devolved revenues and Scottish GDP is far from perfect, this is often argued to be constraining. See also Bell et al. (2016) for more details on the capital and resource borrowing limits in numbers.

used for this allocation. Sticking to our example of education, the lump sum could thus be split across regions based on a simple measure of need – e.g. annual population of school age.

Next, and to reward regional politicians for welfare-enhancing policies via their incoming revenues, the responsibility mechanism introduces performance-based indicators to the grant scheme. These indicators determine the growth of the sub-divided lump sum, and thus provide regions with a bonus if they perform better than average. Keeping to our example of education, this could imply that regions where the growth rate of pupil numbers is higher than the national average in a given year – or drop-out rates are lower – receive a bigger grant in that year. Crucially, the potential size of this bonus can be weighted using unilaterally defined measures which – within certain nationally agreed-upon bounds – can vary through time according to regional preference.

Our grant scheme has a number of attractive features in the pragmatic sense, along with the more conceptual advantages discussed in Section 3.2. Since it simultaneously combines responsibility and horizontal equity elements, it is more intuitive and politically operational than standard equalisation grants which are only equity-oriented and far less underpinned by specific compensation arguments. In addition, the formula is fixed and transparent, yet automatically adapts to evolutions in needs and performance through time, which sets it apart from the discretionary regional finance prone to political meddling and resulting inefficiencies.[6] The grant scheme is also highly flexible, in the sense that it can be separately applied to specific policy areas — as in our education example above — or to the sum total of regional spending to be financed by grants under the new fiscal framework.[7] Since we can also attribute weights to the responsibility and compensation components of the grant formula, the scheme not only allows for flexibility in the choice of needs-based and performance-based criteria, but also regarding their relative importance. It would thus be possible to allow the relative weight of responsibility to vary across lower-level jurisdictions, as the latter can decide separately and unilaterally on the likelihood or appropriateness of being rewarded.

These degrees of freedom with respect to interregional fairness, responsibility and local needs can then be expected to facilitate the democratic debate that would precede the install-ation of a new fiscal framework. In other words, the grant scheme can easily be calibrated to the demands of any real-world setting elicited through the electoral process.

---

[6]Since the compensation component of the grant scheme is largely shock-proof for the same reason, the grants will implicitly maintain a lower bound of interregional insurance in each year.

[7]In this latter case the lump sum G_t would simply capture overall spending, and the aggregate needs-based and performance-based indicators could be a weighted average of various specific and relevant indicators. If applied to one policy area furthermore, the grant scheme also allows for the grant to be 'earmarked', i.e. subject to the precondition that the funds have to be spent in the specific area in question.

3.3.2 How to measure compensation and responsibility?

Choosing which criteria are relevant to introduce as needs- or performance-based indicators will partially follow from the democratic process, which logically depends on the kinds of policy domains that are devolved in the first place. The 'compensation' mechanism would rely on indicators and measures of need, on which agreement is often problematic. The Independent Commission on Funding and Finance for Wales for example, selected a group of indicators and then attempted to determine equalisation payments for Wales based on identifying comparable levels of need and associated levels of financial support in England. Clearly the new fiscal framework would have to include an institutional process to identify appropriate indicators, their weighting, and how they might be adjusted over time. There are international exemplars of such institutions, including the Australian Grants Commission. To ensure trust, this institution would have to be transparent and independent from government at all levels.

Focusing on the UK in what follows, and based on the functions of government currently devolved to the UK nations, we would propose the following list of easily accessible indicators to underpin the compensation mechanism. All of these are based on population numbers, and are therefore highly transparent, and for this reason applicable to other developed countries with similar devolved powers. They will always be expressed as regional population shares in the relevant nationwide population, before being inserted into the grant formula:

- Number of older people (+80)

- Number of dependent children (-18)

- Number of people claiming income-related benefits

- Number of people with a long-term limiting illness

- Number of people from a minority ethnic group

- Number of people living outside communities of 10,000 people or more

- Number of people living on remote islands

The responsibility mechanism secondly, would involve the retention of revenues based on achievement of particular objectives. Such mechanisms already exist for the Scottish and Welsh fiscal frameworks, but are Byzantine in their complexity. For example, the Scottish Government gains (or loses) the revenue associated with increasing (decreasing) its income

tax per head at a faster rate than in the rest of the UK.[8] It can use this additional revenue for its electoral advantage as it sees fit.

For reasons set out in Section 3.2, responsibility incentives need not be limited to the retention of additional tax revenues. They could also include targets for the outcomes of spending programs such as increases in new firm formation or participation rates, improvements in life expectancy or schooling levels, reductions in inequality or increases in skills formation. This would be controversial in a UK context since each government jealously guards its ability to set priorities. However, there are perhaps some broad objectives such as increased life expectancy which might be agreed without interfering with governmental strategy as to how they should be achieved. Nevertheless, unless the process for determining those outcomes that would be subject to these incentives – and how good performance would be rewarded – was agreed by all governments, progress on this form of responsibility payment might be difficult. On the other hand, the fact that our mechanism allows each government to determine the relative importance of the responsibility bonus, can be expected to reduce frictions here.

3.4 Towards a softer participation constraint

A 'fair' combination of responsibility and compensation may be normatively sound and intuitive, but will not necessarily be perceived as such. Preferences regarding horizontal equity may differ across regional jurisdictions for all sorts of reasons, and more importantly, are susceptible to change. Preference instability of this kind is then liable to influence the process of devolution itself. If for example a compensation-inspired grant system is under pressure in a certain region, the fact that it is normatively fair from a nationwide perspective will matter less. In this case, and as resistance to interregional fairness mounts, calls for more responsibility in the form of (tax) autonomy or even full independence can increasingly be expected. To different degrees this is happening in Spain, Italy, Belgium, the UK, Canada, Germany and other countries. Our proposed mechanism would not prevent substate governments seeking more fiscal responsibility: it would simply make more transparent the fiscal implications of such a choice, and would give each subnational government more leeway to make it.

By allowing each of the substate governments to modulate the relative weights of compensation and responsibility directly and unilaterally, our proposed framework thus resolves some of these issues from the outset of the newly introduced system. Moreover, and from a dynamic point of view, the resulting flexibility of the participation constraint is guaranteed across time, as each new regional government can unilaterally decide to refashion its desired

---

[8]This is due to the adjustments made to the Barnett formula, introduced to compensate for the decentralisation of the income tax after 2016. See Bell et al. (2016) for a detailed account on these reforms, as well as their expected (budgetary) effects.

trade-off at any point – ideally after periods of a pre-defined amount of years. The price of this stability will largely be paid by central government, since any shifts in subnational support paid through the grant scheme will ultimately be financed by the national budget, and/or public borrowing. The game theoretical implications of a softer participation constraint are interesting by itself, but also crucial to understand in order to keep the framework stable. On the one hand regions performing above average at a consistent rate would be driven to set their responsibility weights at the maximum level as a best response in equilibrium, whilst below-average regions would do the opposite. However, as economic fortunes and policies may endogenously vary over time, below-average regions would still be tempted to outperform the average in the medium term, after which they would re-adjust their weights. The incentive to do better is therefore present even for underperforming regional governments, even when a certain lower bound of compensation is always guaranteed. Nonetheless, the responsibility bonus will need to bounded by certain thresholds, framed by a nationally agreed upon consensus.

Additionally, we also foresee the need to introduce a transition mechanism to facilitate political compliance, which would compensate for the differences between the old and the new system. Over a sufficiently wide – and to be agreed upon – time frame, the net budgetary effects of introducing the new grant system for each substate government would be neutralised. Such a correction could then be phased out gradually as the envisaged transition period draws to an end.

3.5 Bridging tax autonomy and grant finance

In most countries the degree of devolved taxation usually follows from consensus-driven deliberation on a national scale. However, as was the case when determining the responsibility bonus in our grant mechanism set out in Section 3.3 above, unilateral authority could also be granted with regards to the decision on regional tax autonomy. This implies a transfer of risk which the devolved governments will embrace with varying degrees of enthusiasm. They could opt for a high degree of compensation via the grant mechanism and relatively little fiscal responsibility, or they could opt for full tax autonomy, which implies a complete transfer of fiscal risk. If local borrowing limits are fully relaxed in the process, this would differ from complete independence only insofar as control of monetary policy would still be the responsibility of the central government, and there would be a requirement to pay for those services financed by the central government that are necessary to support a sovereign state. To guarantee the stability and sustainability of the federation, as well as the public finances underpinning the union, nationally agreed-upon bounds to regional fiscal autonomy will consequently be necessary to streamline the fiscal framework. Similarly, the feedback

loop from fiscal to monetary policy through bond yields and interest rates would likely require agreement on deficits and debt between the different levels of government. The extent to which the unilateral discretion is bound by nationally agreed upon thresholds will also serve as a check on interregional spill-overs.

The unilaterally chosen percentage of tax autonomy would be applied to the nationwide budget X^t – which captures total spending on devolved powers – to calibrate the revenues from tax autonomy in the starting year t of the new system for each region i, as explained in Section B of the appendix. These revenues would subsequently evolve through time following the growth rate of the various tax bases, devolved or shared for the purpose of tax autonomy, as well as the evolution of chosen tax levels. If a counterfactual evolution of the nationwide budget X^t is maintained moreover, for example using the indexation method used in Eq. (3), the unilateral choice for more or less tax autonomy can be remade after a pre-defined but given amount of years, as was the case regarding the responsibility bonus inherent to the grant scheme suggested above.[9] Since regionally devolved tax autonomy always functions alongside the grant scheme in our framework, the latter residually covers the remainder of the regionally allocated budget. The conceptualisation of the grant mechanism is exactly the same as before, in Section 3.3.

Bridging tax autonomy and grant finance as we suggest here, has two main advantages. First, we use needs-based indicators – such as regional population shares given by Eq. (3) in appendix – to define how much of the national budget X^t is to be financed via regional tax revenues. This links up tax autonomy with the compensation-inspired approach set out earlier, independent of the unilateral choice of how much tax autonomy to take on as a region. Second, by allowing regions to unilaterally decide on how much of this allocated budget they want to raise themselves by means of devolved tax authority, and how much will be financed through grants, the participation constraint – as discussed in the previous section – will be "softened" even more. Of course, the main implication of such flexibility will be asymmetric degrees of tax authority between the national and regional levels of government, yet often with regard to the same tax bases.[10] Some regions may opt for a higher degree of autonomy over e.g. a shared labour income tax base, and others for a relatively lower degree.[11]

---

[9]The same game theoretical caveats discussed in Section 3.4 are valid here. Depending on the evolution of the relevant devolved tax bases, we would therefore expect the bounds of unilateral discretion to offer less flexibility to the regions, given the potentially more pronounced strains on the federal/national purse following these unilateral decisions.

[10]Formulas of tax sharing between the central and lower levels of government could be based on the "split rate" approach, where e.g. two entirely different systems of tax brackets would be applied on the same, shared labour income tax base. Alternatively a regional surcharge – or "piggy-back"– tax rate could be levied on central tax revenues, which are then lowered for the purpose.

[11]Note that our proposal would most likely also imply that the full tax autonomy currently enjoyed by the Scottish government in terms of the personal income tax, will have to be reversed to some extent.

3.6 Comparison to current UK framework

It is clear that the current UK fiscal framework fails on most fronts set out in Section 2 above. UK Government reaction to the possibility of Scotland leaving the Union has resulted in the transfer of tax powers. When complete, this will give Scotland a retention ratio of around 40%, one of the highest for a substate government in the developed world and significantly higher than the other UK nations. This process of extending fiscal powers was based on very short-run political considerations: the outcome is a set of complex arrangements, which are poorly understood by the electorate. Compounding this complexity, these arrangements are overlaid by the Barnett formula, which itself is not well understood and provides little (fiscal) incentive for politicians to perform well.

Indeed, the Barnett-based block grant offers no real accountability in the sense that it is not transparent, and is not even linked to current spending in each respective nation but rather to spending in England. It fails to hold local politicians responsible as well for much the same reasons. Since there is no clear link between regional revenues and sensible regional policies there is no responsibility feedback loop underlying the Barnett formula. The current arrangement is also more unpredictable than a block grant linked to regional performance indicators – which are partially under a local government's effective control – as well as to relatively stable and predictable needs-based indicators, which is exactly the grant scheme we proposed above in Section 3.3. The Barnett formula also fails in terms of horizontal equity as its per capita endowments differ widely across UK nations. In addition, there is wide suspicion of the formula due to the role of HM Treasury which both decides the scope of the formula and how much will be spent on it.

The new arrangements regarding Scottish tax autonomy also fail to introduce the principles of regional public finance developed in Section 2. Tax autonomy ensures some degree of transparency and responsibility, yet the grant adjustment mechanism undercuts transparency and the feedback loop working through tax revenues.[12] Furthermore, there are no horizontal equity guarantees beyond the grant adjustments itself. The VAT tax sharing arrangement lastly, hardly ticks any of the boxes put forward in Section 3. In contrast, our proposal is to have a transparent process which exposes how far the devolved nations' resources derive separately from compensation and responsibility elements, letting the public choose the balance between these different approaches to funding. Independence and transparency of process is a necessary condition for building confidence in a new fiscal framework. These attributes would necessarily apply to data collection, interpretation and associated decision-making.

Risk transfer is embodied within the recently agreed Scottish and Welsh fiscal frameworks,

---

[12]For a detailed account of these grant mechanism adjustments, compensating for most of the budgetary consequences of devolving the income tax, see Bell et al. (2016).

but in a convoluted and opaque fashion which derives from the politically expedient decision to continue with the Barnett Formula. Note as well that significant moves towards full fiscal autonomy are only consistent with the UK government holding responsibility for the UK fiscal stance if the extent of this autonomy does not have a destabilising effect on that stance. This is only feasible because the devolved nations are relatively small compared with the UK as a whole. But it does create a difficulty if England had similar powers to design its tax structures and retain the associated revenues, in which case a shared income tax arrangement between the central and lower levels of government is advised. Of course, this issue would pose less of a problem the more England itself would in turn be internally devolved, and tax competition can be avoided. Our proposed fiscal framework circumvents these issues by providing an alternative to full tax autonomy, whilst keeping much of its benefits in terms of responsibility and transparency.

Subject to the constraint that central government retains control over the UK fiscal stance and stabilisation policy, the issue of risk is also relevant to the choice of tax instruments allocated to different levels of government. Increased fiscal autonomy will reduce the efficiency of the UK single market if it distorts resource allocation. This effect will be greater if taxes are applied to more mobile factors, as already mentioned in Section 2. Taxes on property, profits or income each have their own set of risks and rewards for government revenues which depend on the reactions of individuals and firms to their imposition. These arguments are well worn also in the UK context, and relate to issues such as the stability of the tax base, costs of collection and effects on economic efficiency. These are ably summarised in the Mirrlees Review (2010), while the arguments in relation to substate tax policy are reviewed in the first report of the Independent Expert Group to the Calman Commission (2008) and in the final report of the Independent Commission on Funding and Finance for Wales (2010). We do not repeat these arguments here.

4   Process and institutions

This section deals with the processes and institutions that might underpin the fiscal framework proposed for the UK nations in previous sections. First, we outline the macroeconomic framework behind current arrangements, arguing that these should remain largely intact. Secondly, we discuss how a new institution tasked with regulating the fiscal framework might be positioned alongside current structures. Currently, UK macroeconomic policy comprises three main elements:

- HM Treasury, part of the UK government, with oversight and control of the public finances

- the Office of Budget Responsibility (OBR) whose role is to construct independent economic forecasts and to analyse the public finances

- the Bank of England (BoE), an independent central bank

Underpinning these arrangements is an acceptance that the UK government has the authority to set its own fiscal stance and, in particular, to control levels of deficit and debt. If it so wishes, it can choose to exercise these fiscal powers in a counter-cyclical fashion and so protect citizens from the adverse effects of variations in aggregate demand. However, the exercise of these fiscal powers is constrained by the transparency accorded to the independent forecasts produced by the OBR. Such constraints have been accepted so that current government action should not have a radically negative effect on the opportunities available to future governments. However, the OBR has no protection should future governments wish to repeal the legislation which established it. Finally, the main role of the Bank of England is to maintain the rate of price inflation close to its target of 2%. It uses various forms of monetary policy to achieve this goal. Like the OBR, its independence is not constitutionally guaranteed.

At present, control of fiscal and monetary policy is reserved to the UK Government. Fiscal policy is exclusively reserved to central governments in most, but not all, developed countries, while monetary policy is exclusively a responsibility of central governments. Within the framework of UK legislation, monetary and fiscal policy are reserved powers as set out in the Scotland Act 1998, Schedule 5, Part II, Para A1, though this was amended in 2016 to make provision for the new tax powers assigned to Scotland under the Scotland Act 2016. We do not expect any change in this framework, though a case can certainly be made for a more consultative inter-governmental approach to planning taxation and public spending. Our proposal for a new fiscal framework for the UK nations would sit alongside these existing arrangements for fiscal and monetary policy.

While these institutions have some stern critics, their radical rearrangement to accommodate a new fiscal framework for the UK nations might be seen as unnecessary and disruptive. Nevertheless, some aspects of their structure provide useful insights into the design of this framework. In particular, just as with the BoE and the OBR, independence and transparency are key attributes of institutions that are necessary for there to be confidence in the new fiscal framework. These attributes are relevant for the collection of data, for its interpretation, and for the consequent decisions. The present arrangements under the Barnett formula do not have these attributes, which partly explains their ad hoc and inconsistent characteristics.

The new fiscal framework for Scotland, which includes new tax and welfare powers, also provides for the establishment of a Scottish Fiscal Commission, whose role is to construct economic and fiscal forecasts for Scotland which the Scottish Government must take account

of in its budgetary decisions. The Scottish Fiscal Commission is independent of the UK and Scottish governments and works closely with the Office of Budgetary Responsibility. The Scottish Fiscal Commission and the OBR have a statutory duty to cooperate in ways set out in a Memorandum of Understanding between the two bodies. One key element of this is the coordination of their forecasting processes. They do not have to produce identical forecasts, but they do have to explain any differences that do arise. Similar institutions might emerge for Wales and Northern Ireland should they opt to significantly increase their fiscal responsibilities.

In the light of these developments, we propose the establishment of a Responsibility and Compensation Commission (RCC) whose task would be to:

- determine compensation/responsibility payments for UK nations by applying the approach outlined in this paper. It would carry out this task in conjunction with the UK and Scottish governments and the Welsh and Northern Irish assemblies. It would liaise with the OBR and with the ONS, which would be responsible for collecting comparable data across the four nations.

- work with HMRC and the devolved authorities to ensure that the responsibility element of the devolved governments' funding is transparently and fairly administered

- conduct decennial review of the formula, its scope, its efficacy in promoting responsibility and its fairness as between the nations, also depending on electoral outcomes.

Like the OBR, the RCC would be an advisory non-departmental public body. To retain the confidence of the devolved authorities, it should not be based in London.

As argued in Section 3.4, there would also have to be a transitional period of adjustment to the new arrangements. These would have to deal with the political sensitivities relevant to the time of introduction. Thus, for example, devolved governments could be reassured that the new arrangements would be put in place over a period of years during which they would be guaranteed no reduction in funding.

5 Conclusion

We suggest that the complexity, lack of transparency and independence of existing arrangements for distributing funds to substate governments undermine both their sustainability and effectiveness. We propose that a new mechanism, which makes explicit the "compensation" and "responsibility" elements of their funding, should be established. This mechanism would permit substate governments to determine the extent of responsibility and autonomy (and the

level of risk) that they wished to take on, relative to a guaranteed amount of compensation for regional needs.

Applied to the UK setting, this would mean the Celtic nations could benefit or lose relative to the rest of the UK, depending on their decisions, and would potentially have to accept unfavourable outcomes. Just as the rest of the UK would have to accept outcomes where they fared better. As a result, our framework would necessarily involve the establishment of processes to identify relative need across different parts of the UK, and to establish a transition mechanism to adjust toward the implied distribution of funding support. It would also be required to monitor spill-over effects from idiosyncratic fiscal decisions. This would ideally require an institutional base: our proposal is to establish a Responsibility and Compensation Commission (RCC). It could be expected to be scrutinised by each Parliament.

Appendix A Grant mechanism: Formula

The annual lump sum G^t – assigned for regional spending – is divided across substate governments in a given year tEq. (1) as follows

$$G_i^t = (\overbrace{\beta_i^t \times G^t}^{Compensation}) \times [1 + \underbrace{(\mathbb{1}_E \times \omega_i \times (\alpha_i^t - \bar{\alpha}^t))}_{Responsibility}], \tag{1}$$

which defines the annual grant G_i^t going to each substate government i in each year t. To illustrate Eq. (1) further, suppose the grant mechanism would only finance education. The lump sum G^t in Eq. (1) would then cover the total amount deemed sufficient for spending on education nationwide. This could be actual total spending at the time of introduction of the new grant scheme, or some other agreed figure. Following its introduction, the lump sum would increase over time based on pre-defined indices such as inflation π^t and/or economic growth $\bar{\alpha}^t$ as specified below, so that

$$G^t = G^{t-1} \times (1 + \bar{\alpha}^t) \times (1 + \pi^t). \tag{2}$$

However, this only keeps the lump sum G^t from eroding in real or nominal terms, but does not subdivide it across the substate governments. This is where the compensation mechanism on the right hand side of Eq. (1) comes in, since needs-based indicators will be used for this allocation. Sticking to our example of education, the lump sum could thus be split across regions based on a simple measure of need - the annual population N_i^t of school age. This would set the regional share β_i^t, used in Eq. (1) to allocate the lump sum G^t, equal to

$$\beta_i^t = \frac{N_i^t}{N^t}. \tag{3}$$

Next, and to reward regional politicians for welfare-enhancing policies via their incoming revenues, the responsibility mechanism introduces performance-based indicators α_i to the grant scheme, as also marked on the right hand side of Eq. (1). These indicators determine the growth of the sub-divided lump sum, and thus provide regions with a bonus if they perform better than the average value $\bar{\alpha}^t$, given by

$$\bar{\alpha}^t = \frac{\sum_{i=1}^n \alpha_i^t}{n}, \tag{4}$$

so that we can define the indicator function in Eq. (1) as follows

$$\mathbb{1}_E(\alpha_i^t, \bar{\alpha}^t) = \begin{cases} 1, & \text{if } \alpha_i^t > \bar{\alpha}^t \\ 0, & \text{otherwise} \end{cases}, \tag{5}$$

which then sets out the conditions for receiving the bonus. The reward can be weighted using the relative measure $\omega_i \in [\underline{x}, \bar{x}]$, with x setting the upper and lower bounds on regional

discretion. Keeping to our example of education, this could imply that regions where the growth rate α_i^t of growth in pupil numbers is higher than the national average $\bar{\alpha}^t$ in a given year, receive a bigger grant in that year.

Appendix B   Tax autonomy and grant finance

The grant scheme proposed in Section 3.3 will always function alongside a certain amount of regionally devolved tax autonomy. Denoting regional revenues from devolved taxation by T_i, we can write overall regional revenues R_i of region i as

$$R_i^t = T_i^t + G_i^t, \tag{6}$$

which is nothing more than the sum of grant funding G_i^t and tax revenues T_i^t. Now, to define the amount T_i^t reserved for devolved taxation and the amount G_i^t to be financed via the grant mechanism in Eq. (6), we first define a measure X^t comprising nationwide overall desired spending on devolved powers. If we follow a compensation-based approach to compute how much of this nationwide budget X^t should be allocated to each region i, regional revenues spelled out in Eq. (6) become

$$R_i^t = T_i^t + G_i^t \equiv \overbrace{\beta_i^t \times X^t}^{Compensation}, \tag{7}$$

with β_i^t a – possibly weighted – average of all selected compensation criteria, as described in Section 3.3.2. Now that we know the total size of the pie per region, the question is how much of it will be raised via devolved taxation.

In most countries the degree of devolved taxation usually follows from consensus-driven deliberation on a national scale. However, as was the case when determining the responsibility bonus in our grant mechanism set out in Section 3.3 above, the regions could also be allowed to unilaterally decide on tax autonomy here. The extent of this kind of unilateral discretion, which we denote as γ_i in Eq. (8), would necessarily have to be bound by nationally agreed upon thresholds $[\underline{y}, \overline{y}]$, where we set $0 < \underline{y} < \gamma_i < \overline{y} < 1$. This unilaterally chosen percentage γ_i is then applied to the regionally allocated nationwide budget R_i^t, so that the revenues from tax autonomy in the starting year t of the new system in region i would be calibrated to equal

$$T_i^t = \gamma_i \times \overbrace{(\beta_i^t \times X^t)}^{Compensation} \equiv \tau_i^t B_i^t, \tag{8}$$

which would subsequently evolve through time following the growth rate of the various tax bases devolved – or shared – for the purpose of tax autonomy, here captured by B_i^t, as well as the evolution of regionally chosen tax levels τ_i^t. If a counterfactual evolution of the nationwide

budget X_t is maintained moreover, for example using the indexation method used in Eq. (3), the unilateral choice defined by γ_i in Eq. (8) can be remade after a pre-defined but given amount of years, as was the case regarding the responsibility bonus in Eq. (1).[13] Next, and logically, the grant scheme would cover the residual of the regionally allocated budget as follows

$$G_i^t = (1 - \gamma_i) \times \overbrace{(\beta_i^t \times X^t)}^{Compensation} \times [1 + \underbrace{(\mathbb{1}_E \times \omega_i \times (\alpha_i^t - \bar{\alpha}^t))}_{Responsibility}], \tag{9}$$

where the conceptualisation of the grant mechanism is exactly the same as before, in Eq. (1).

---

[13]The same game theoretical caveats discussed in Section 3.4 are valid here. Depending on the evolution of the relevant devolved tax bases, we would therefore expect the bounds of unilateral discretion $\gamma_i \in [\underline{y}, \overline{y}]$ to offer less flexibility to the regions, given the potentially more pronounced strains on the federal/national purse following these unilateral decisions.

References

Andersson, L., 2008. Fiscal flows and financial markets: To what extent do they provide risk sharing within sweden? Regional Studies 42, 1003–1011.

Arachi, G., Ferrario, C., Zanardi, A., 2010. Regional redistribution and risk sharing in italy: the role of different tiers of government. Regional Studies 44, 55–69.

Asdrubali, P., Sorensen, B.E., Yosha, O., 1996. Channels of interstate risk sharing: United states 1963-1990. The Quarterly Journal of Economics , 1081–1110.

Baskaran, T., 2012. Soft budget constraints and strategic interactions in subnational borrowing: Evidence from the german states, 1975–2005. Journal of Urban Economics 71, 114–127.

Bell, D.N., Eiser, D., Phillips, D., 2016. Scotland's fiscal framework: Assessing the agreement. Technical Report. IFS Working Papers.

Blankart, C.B., Klaiber, A., 2006. Subnational government organisation and public debt crises. Economic Affairs 26, 48–54.

Boadway, R., Shah, A., 2009. Fiscal federalism: principles and practice of multiorder governance. Cambridge University Press.

Boadway, R., Tremblay, J.F., 2012. Reassessment of the tiebout model. Journal of public economics 96, 1063–1078.

Boadway, R.W., Shah, A., 2007. Intergovernmental fiscal transfers: principles and practices. World Bank Publications.

Bordignon, M., Turati, G., 2009. Bailing out expectations and public health expenditure. Journal of Health Economics 28, 305–321.

Cohen, G.A., 1989. On the currency of egalitarian justice. Ethics 99, 906–944.

Foremny, D., 2014. Sub-national deficits in european countries: The impact of fiscal rules and tax autonomy. European Journal of Political Economy 34, 86–110.

Jin, H., Qian, Y., Weingast, B.R., 2005. Regional decentralization and fiscal incentives: Federalism, Chinese style. Journal of Public Economics 89, 1719–1742.

Keen, M.J., Kotsogiannis, C., 2002. Does federalism lead to excessively high taxes? The American Economic Review 92, 363–370.

Lockwood, B., 2006. The political economy of decentralization, in: Handbook of multi-level finance. Ahmad, Ehhisham and Brosio, Giorgio. chapter 2.

Oates, W., 1972. Fiscal federalism. Harcourt Brace Jovanovich (N.Y.).

Oates, W.E., 1999. An essay on fiscal federalism. Journal of Economic Literature 37, 1120–1149.

Oates, W.E., 2005. Toward a second-generation theory of fiscal federalism. International tax and public finance 12, 349–373.

Rodden, J., 2002. The dilemma of fiscal federalism: grants and fiscal performance around the world. American Journal of Political Science , 670–687.

Rodden, J., Eskeland, G.S., Litvack, J.I., 2003. Fiscal decentralization and the challenge of hard budget constraints. MIT press.

Roemer, J.E., 1998. Why the poor do not expropriate the rich: an old argument in new garb. Journal of Public Economics 70, 399–424.

Sas, W., 2017. Can fiscal equalisation mitigate tax competition? ad valorem and residence-based taxation in a federation. International Tax and Public Finance 24, 817–853.

Weingast, B.R., 1995. The Economic Role of Political Institutions: Market-Preserving Federalism and Economic Development. Journal of Law, Economics and Organization 11, 1–31.

Weingast, B.R., 2009. Second generation fiscal federalism: The implications of fiscal incentives. Journal of Urban Economics 65, 279–293.

Wilson, J.D., 1986. A theory of interregional tax competition. Journal of Urban Economics 19, 296–315.

Wilson, J.D., Wildasin, D.E., 2004. Capital tax competition: bane or boon. Journal of Public Economics 88, 1065–1091.