David Heald's responses to the HM Treasury/Scottish Government consultation on Block Grant Adjustments as part of the Fiscal Framework Review

a. What do you consider to be the strengths and weaknesses of the current approach to calculating block grant adjustments for devolved taxes and social security spending for Scotland?

Since the 1999 re-establishment of the Scottish Parliament, it has been on a shaky journey towards funding more of its spending from internal Scottish sources. Particularly due to the asymmetric nature of UK devolution, there has had to be learning-on-the-job. In recent years substantial improvement in the transparency of funding has occurred, for example through the Treasury's Block Grant Transparency and through the work of the Scottish Fiscal Commission.

Contrary to some media and political expectations, when tax powers are transferred from the UK Government to the Scottish Government, that does not, of itself, bring "more money" but, in principle, provides more financial discretion to the Scottish Parliament. The BGA for each tax represents the loss of revenue to the UK Government in the base year, with counterfactuals being calculated for future years. A parallel process applies to devolved social security spending. The aim is to estimate in each year the revenue foregone by the UK Government from taxes that have been devolved; this is easier to grasp intuitively than it is to measure. Over time, especially if policies significantly diverge, the scope for disputes grows.

Viewed from outside governments, the big issue when the 2016 Fiscal Framework was agreed shortly before the Brexit Referendum seemed to relate to the choice between the Indexed Per Capita method (favoured by the Scottish Government) and the Comparable Method (favoured by the UK Government). Although with uncertain long-term status, the Indexed Per Capita method has been used during this so-called transitional period between 2016 and the forthcoming Fiscal Framework Review. This has protected the Scottish Government against population risk.

What has emerged in the public domain since implementation of BGAs is the importance of the differential composition of the UK and rUK tax bases, particularly for income tax. The complexities of this issue have been explored in Annex A of the Scottish Government's December 2021 Medium-Term Fiscal Strategy. There are several dimensions, including: earnings profile; demographic structure; and differences in the sectoral composition of income tax revenues (e.g. finance employment in London and the wider south east, and oil-related employment in North East Scotland). Even if identical tax policies were followed, there could be divergences between devolved revenues and BGAs, most likely to the disadvantage of Scotland. This is a topic that should be investigated further in the Independent Report.

I support the continuation of the Indexed Per Capita method. The Scottish Government does not have powers to manage population risk and could not have those within a continuing United Kingdom. Scotland's share of UK population has fallen steadily since the first modern census in 1801 and probably ever since the Act of Union in 1707. It is not Scotland alone that is affected by the excessive territorial centralisation of the UK, to which the Levelling Up agenda of the Johnson Government drew attention.

I appreciate that the Independent Report has a narrow remit and that it is a preliminary to the formal Fiscal Framework Review between the UK and Scottish Governments. However, it is vital to contextualise the technocratic exercise that is the calculation of BGAs and the working through of post-event reconciliations.

First, the 2016 Fiscal Framework was designed in a more stable era. Since then, the public finances have been disrupted by Brexit, COVID-19, and the cost-of-living/energy crisis. UK fiscal events, with implications for the Scottish Budget, have proliferated. The Scottish Spending Review 2022 has been compromised by subsequent events, including the change of UK Government and the 23 September 2022 fiscal event. The Scottish Parliament could justifiably claim that its financial procedures were more thorough and inclusive than those of Westminster, but the frequency of changes and their timing has severely compressed the Scottish Budget-setting process.

Second, power asymmetry between the UK and Scottish Governments has become more salient. There are legislative constraints on the Scottish Government: for example, it is not possible for the Scottish Government to introduce changes to Scottish Income Tax in-year. Before the start of each financial year, the Scottish Parliament must pass a Scottish Rates Resolution to set the rates for non-savings non-dividend income, which then applies for the whole of that financial year. Constraints on the UK Government derive not from law but from convention, in which experience has demonstrated that less reliance can now be placed, and from debt markets.

Third, the Scottish Government has a much narrower range of tax powers. When the Johnson Government increased UK-wide National Insurance Contributions mid-year and promised to reduce the rUK basic rate of income tax in 2024-25, this demonstrated the potential squeezing of devolved tax powers as pressure would build for a comparable Scottish reduction. It is important to distinguish two kinds of UK Government powers: those over the architecture and implementation of income tax across the UK; and the setting of rUK rates in parallel with the rate-setting decision of the Scottish Government. An example of the former is the setting of the Personal Allowance, which went up rapidly in the 2010s, and is believed to have differentially adversely affected Scottish income tax revenue because of the earnings profile. Another example with possibly adverse effects for Scotland is the 23 September 2022 announcement of relaxations of Off-Payroll working regulations (IR35) which are frequently used in the oil-related sector in North East Scotland which is a strong contributor to Scottish revenues.

Fourth, the acceptability of BGA arrangements to the Scottish Government will depend in part on the extent of its borrowing powers. Large reconciliation adjustments when final data become available are disruptive of budget planning, especially for the medium-term. The Scottish Government's borrowing limits are both too small and set in nominal terms, unexpectedly higher inflation drawing more attention to this constraint. Federal or central governments are sensitive to the debt exposure of subnational governments: Saskatchewan in the 1980s was a prominent case. Given the independence ambition of SNP Governments, it is unsurprising that the UK Government is wary of extending borrowing powers due to fears that borrowing would be irresponsible. Although that fear will constrain what the UK Government is likely to agree to, making the BGA system work requires a higher level of resource borrowing.

I do not comment on BGAs for social security, about which I have no knowledge.

b. To what extent do you think that the various approaches to calculating the Scottish block grant adjustments, outlined in the background note, are consistent with the Smith Commission's principles? How could the calculation of the BGAs be made more consistent with the Smith Commission principles?

The Smith Commission report did not come down from Mount Sinai as divine revelation. Instead, it was a rushed fix to resolve a pressing political problem after the close result of the 2014 Scottish Independence Referendum. Its contents acquired political significance because it was a Scottish multi-party commission to which the Cameron Government effectively delegated decision-making.

The first No Detriment Principle is that the act of devolving taxes should not damage the public finances of a devolved government taking on new taxation powers at the point at which this occurs. If this were not the case, it is unlikely that there would be much tax devolution.

The second No Detriment Principle is that subsequent policy decisions by either the UK or devolved government should not damage the other. I have always regarded this principle as incoherent, however reasonable it might sound at first sight. Unless it is assumed that both governments will always make the same tax decisions, then there are likely to be spillovers if one government decides, for example, to have lower/differently-structured income tax than the other. Quantifying such spillovers is likely to prove difficult, especially if the issue of compensation arises.

My understanding is that there has not yet been a resolution between the UK and Scottish Governments on whether compensation is owed to the latter due to the UK Government increasing the UK-wide Personal Allowance, with an adversely differential effect on Scotland because of the different structure of its tax base. This case is distinct in that this is not the UK Government exercising powers (for example, over bands) for its own jurisdiction, but of it taking UK-wide decisions.

The second No Detriment Principle could open a can of worms of compensation claims. For a devolved tax system to work in the specific circumstances of the United Kingdom (e.g. asymmetric devolution, small geographic distances, fiscal history of excessive centralisation), this requires (a) conciliatory intergovernmental relationships, and (b) belief that political processes and potential taxpayer mobility will put constraints on divergence.

In summary, I do not believe that what is stated in the Smith Commission Report is decisive for the decisions to be taken in the Fiscal Framework Review, of which the BGA methods are an important element. The Smith Commission Report is now eight years old, and the proposal for partial VAT assignment has been forgotten about.

c. To what extent do you think the various approaches to calculating the Scottish block grant adjustments shares risks between the Scottish and UK governments appropriately? To what extent do you think it is important that the allocation of risks implied by the BGA mechanism aligns with the balance of risks held under the Barnett formula?

In relation to Scottish per capita spending on devolved services, the Barnett formula has not produced the degree of convergence on England = 100 that mathematical models of its operation predicted when limited information about the formula became public in the 1980s and 1990s. Data do not exist to make an authoritative assessment of the reasons why, but lower rates of population increase in Scotland are certainly a factor.

In terms of taxation revenues, population risk is the critical one for Scotland. Protection from this is justified by (a) the seemingly irresistible trend for UK population to become more concentrated around London and England's south-east corner, and (b) the fact that the UK Government controls most of the levers relevant to the distribution of economic activity and population. Whether the UK as whole cares about such territorial imbalance and, if so, what could be done, is outside the scope of this submission.

I do not think it is important that the allocation of risks implied by the BGA mechanism aligns with the balance of risks held under the Barnett formula. Although Scotland has done well in this respect on the spending side (only the increment works on up-to-date population, not the base), there is no direct parallel to the revenue side. If the argument is that there should be deep cuts in the block grant, then that argument should be spelled out clearly in public, not buried in technocratic matters which are inaccessible to all but a few specialists. Such a decision would open a range of fiscal and constitutional issues, beyond the scope of this submission. Making Scotland take the population risk seems likely to be disruptive of financial planning, so estimates of what the Comparable Method BGAs would have been over the transitional period will be informative.

I will be interested in analyses as to how the Indexed Per Capita method might be adjusted to deal with composition risk, including taxpayer profile and sectoral composition. It is worrying that some forecast income-tax BGAs have been larger than forecast SIT revenues, despite marginally higher tax rates for the upper part of the taxpayer distribution. On the basis of the studies of Taxable Income Elasticities published by HMRC and the Scottish Government, it is unlikely that behavioural responses, as opposed to compositional differences, are the reason for these forecasts. However, complicating BGAs further should only be done where the effect is substantial and the evidence is clear, as the credibility of the devolved tax system depends on BGAs not becoming regarded as a political fix.

One possibility for addressing this compositional problem would be banded BGAs, as in Wales. In practice, this would mean that Scotland would have to match at least some of the rUK bands, a step that would be interpreted as a partial reversal of the 2016 Act at a time when the Scottish Government calls for extended taxation powers. Moreover, the abolition of the rUK Additional Rate from 2023-24 makes a banded system for Scotland look less plausible.

At this juncture, before reading the Independent Report, my inclination is not to over-complicate an already complex system that defies accessible public explanation. The threshold for more complexity should be quite high. Trying to work out how much compensation the UK or Scottish Governments should receive for band and rate changes seems unproductive and conflictual, especially as I would expect differences in tax systems to remain relatively small.

Potentially In a separate category are the decisions the UK Government takes about the architecture of the income tax, for example Personal Allowances and regulatory matters such as Off-Payroll working. Given what is known about differences in taxpayer composition, it would be expected that large increases in Personal Allowances will have differentially adversely affected SIT revenues whereas the present freeze, notwithstanding high inflation, if sustained, would advantage SIT revenues. Estimates of the size of such effects would be welcome.

d. Do you have any other suggestions for how the block grant adjustments should be designed beyond the transitional period?

No.

Given the complexity of these matters, I have doubts whether it will be possible to move from the transitional period to a permanent setting. Instead, there might have to be periodic re-settings, each followed by a new Fiscal Framework renegotiation.

e. Do you have any suggestions for how understanding of block grant adjustments among stakeholders can be improved?

I commend the efforts of the Treasury, Scottish Government, Scottish Fiscal Commission, Institute for Fiscal Studies and Fraser of Allander Institute for their contributions to transparency about the funding of the Scottish Parliament.

What was a simple Barnett Formula mechanism in 1999, even though few then understood it, the reality with devolved taxes has now become bewilderingly complex. As can be seen from media reporting, readers' letters pages in newspapers, broadcasting and social media, public understanding is extremely low, and much commentary is infused with partisanship on party and constitutional-preference lines.

Such efforts at putting explanatory information into the public domain should undoubtedly continue, but the ground is likely to be stony.

I proposed a tax-varying power on income tax in 1976 and strongly believe that the Scottish Income Tax should play a significant role in funding the Scottish Parliament, which means that BGAs must be calculated. I now have four main concerns.

First, that the complexities of tax devolution will be regarded as gobbledygook by most people, so that – whatever the transparency of data and procedures – accountability and legitimacy will be diminished rather than increased in the eyes of the population.

Second, the accountability benefits of tax devolution are believed to occur through the greater scrutiny expenditure claims will receive when Parliaments and local authorities have to raise a non-negligible proportion of their revenues from their own citizen-taxpayers. However, UK Governments take a dismissive attitude to Westminster parliamentary financial scrutiny and the Scottish budget process has been curtailed in recent years by lateness in UK decisions, to some extent attributable to the COVID-19 pandemic. The frequency of UK Government fiscal events is destabilising budgetary planning by the devolved governments. The level of intergovernmental co-operation necessary for fiscal devolution to work is now lacking. The UK Government is spending in devolved areas (rather than the money going through Barnett) and taking rUK decisions on the limited number of taxes held by devolved governments, either without regard for the implications or even with the intention of creating political pressures to force the devolved governments to mimic UK Government decisions.

Third, there are profound asymmetries in UK governance. In territorial terms, the devolved governments only cover 16% of the UK population, with what UK governments decide for England being of massive financial consequence for them. In terms of risk-bearing capacity, the devolved governments have much less than UK governments which have, *inter alia*, the full panoply of tax powers, no balanced budget requirement, unconstrained borrowing powers, their own central bank, and control over the Westminster electoral cycle. This is a difficult system to make work, especially without intergovernmental goodwill. My fear is that devolved taxes will become seen as unusable, with the devolved governments pressurised to mimic UK decisions, after which it would be argued that the transaction and compliance costs of unused parallel systems cannot be justified and that abolishing the tax powers would remove the compositional risks. This would be a regrettable retreat from aspirations for greater fiscal accountability at the times of the Calman and Smith Commissions, and would have unpredictable effects on the constitutional debate.

Fourth, transparency (on which progress has slowly been made) is important, but so is trust in the political system. In this context, trust has two dimensions. First, there has to be confidence in the competence of the bureaucracy and of the political leadership. This has become more questioned in the UK, particularly since the global financial crisis. Second, there has to be confidence that political and bureaucratic leaderships will usually act in the public interest, for example, not opportunistically for private gain or partisan advantage. This has also been shaken, sometimes unfairly given the magnitude of crises that have required addressing. The UK is particularly vulnerable to losses of trust, because so much of its governance depends on convention rather than law. Unlike in a law-based symmetric federal system, the UK devolved governments depend on forbearance from the UK Government. Since 2019, associated with Brexit, UK Governments have intruded on devolved fiscal and policy space, with unknown long-term consequences.

The above goes beyond the remit of the Independent Report in order to contextualise decisions on the Fiscal Framework at a critical juncture in the history of the United Kingdom. These are matters which should be carefully considered by the UK and Scottish Governments.

Glasgow, 26 September 2022