The Macroeconomic Consequences of Regional Fiscal Decentralisation

Preliminary Draft: May 2019

Prepared for the IIOA Conference, Glasgow, July 2019

By

Katerina Lisenkova

Alastair Greig

Peter G. McGregor

Graeme Roy

and

J. Kim Swales

Fraser of Allander Institute

Department of Economics, University of Strathclyde.

Abstract

There has been a widespread move towards more decentralised regional fiscal systems, but there has been little research on how regional fiscal systems themselves may impact upon economic outcomes. This is the concern of the current paper. We explore the impact of a range of stylised regional fiscal systems on the response of regional economies to demand and supply side disturbances, and therefore on the incentive to the regional government to pursue growth promoting policies. These encompass a number of variants of block grant, partially devolved tax, full tax autonomy and equalisation fiscal regimes. We provide a partial and general equilibrium theoretical analysis that furnishes insights into the likely impact of these fiscal regimes on the responsiveness of regional economies to exogenous export and productivity changes, including that tax sharing regimes may not always improve growth incentives relative to more common block grants. However, the theoretical analysis proves incapable of delivering a comprehensive ranking across all the fiscal systems we consider. Furthermore, such analysis provides no indication of the likely scale of the responsiveness of regional economies to alternative fiscal systems, yet this is crucial information for the many economies contemplating a move to further regional fiscal decentralisation. Accordingly, we explore the issue further using an intertemporal, regional empirical general equilibrium model, extended to incorporate a range of alternative fiscal regimes. This allows us to isolate the impact of these regimes on the responsiveness of regional economies within a framework that uses a common database and model. The results offer confirmation of our theoretical analysis, but also generate some surprising results, including that variants of block grant systems could possibly provide even greater incentives to supply-driven growth than full tax autonomy.

Keywords: Regional fiscal autonomy; regional fiscal systems; applied general equilibrium.

JEL codes: D58, R13, R15, H24, H71

1. Introduction and background

Fiscal decentralisation and economic outcomes

In recent decades, there has been a growing trend toward greater levels of fiscal decentralisation, both in developed and developing countries. Much of this is motivated by the subsidiarity principle – the idea that public services should be provided by the most decentralised government level that is capable of delivering them effectively. Oates (2005) provides a summary of these arguments. Key is the social welfare advantages of being able to tailor fiscal policy more closely to the preferences of the local territory, and enhancing the accountability of local politicians for their policy decisions.

But fiscal decentralisation is not without economic costs - Rodgriguez-Pose and Gill (2005). Diseconomies of scale in administration can impact on public sector budgets. It can also create additional complexity and cost for taxpayers and employers. Whether the costs outweigh the risks has been a topic of debate and numerous empirical and theoretical studies. Overall, the results are inconclusive – see Martinez-Vazquez and McNab (2003) for a review of the channels through which fiscal decentralisation could influence economic outcomes.

Most of the studies focus upon the long-term links between public spending, tax and the key drivers of growth such as productivity, labour market participation and innovation. There has been less research into how the structure of the fiscal arrangement themselves could impact upon economic outcomes – see Rodden and Wibbels (2009). This is the focus of this paper.

Defining fiscal decentralisation

Cross country comparisons of fiscal decentralisation are not without their challenges. Common indicators of the degree of decentralisation, such as the share of devolved spending or revenue relative to total government spending, often provide an imprecise or even misleading picture of intergovernmental fiscal frameworks.

Ultimately, what matters is not just the level of spending or taxation earmarked for a particular region (or group of regions) but the level of discretion over the budget itself and the ability to manage fluctuations in the income of different tiers of government. See OECD (2013) for a summary.

The first key element is the source of funding. At one end of the spectrum is a model of full fiscal autonomy. In this framework, all devolved expenditures are funded by the taxes collected in that local community. Such a strong form of 'fiscal devolution' are few and far between. The Basque Country and Navarre in Spain are regularly cited as working under a model of full fiscal autonomy, but even then framework to ensure significant fiscal harmonisation with the rest of Spain.

At the other end of the spectrum, are frameworks where devolved governments are funded entirely by a grant from central government. Such grants can be 'needs based' – i.e. based upon an agreed metric to assess an equal level of spend across different parts of a nation state. Australia runs arguably the most overt programme of 'needs based' allocation of state grants and revenues, with a Commonwealth Grants Commission overseeing the distribution of funds and the methodology deployed.

Alternatively grants can be determined by other factors, often reflecting the bargaining power of individual legislatures. The Barnett Formula in the UK continues to survive as a means of transferring resources from the centre to the devolved nations based solely upon levels of population and comparable spending in England. The fact that it continues to fund higher public spending for

Scotland, despite most indicators of 'need' suggesting that Scotland is one of the better off parts of the UK, continues to attract much political debate.

Grants can also be 'earmarked' or come as a simple 'block grant'. Earmarked grants are monies which the devolved government is able to administer, but only on items pre-determined by the central government. A block grant gives the maximum amount of discretion.

Within the two book-ends of grants and full fiscal autonomy, is the role of tax-sharing. Most countries with substantial decentralised structures run some form of tax-sharing arrangement, with or without in-equalisation. Some systems of tax-sharing involve vertical integration, so that either the tax base or tax revenues (or both) are shared in some way between the central and devolved government. The new income tax powers of the Scottish Parliament are one such example[1].

Other systems also (or instead) incorporate a degree of horizontal integration. Here, tax revenues are shared across different states, landers or provinces – usually based upon 'need'. Whilst there might be some ability to retain at least a proportion of any improvement in revenues, equalisation helps to counterbalance differences in revenue raising capacity. Germany operates a sophisticated system of tax-sharing, with tax revenues first allocated on the basis of what each Lander (roughly) collected on its territory, but then a horizontal equalisation mechanism redistributes these revenues subject to pre-defined criteria (Vandernoot, 2014). Canada operates a similar system of "fiscal equalisation", where revenues are transferred horizontally from provinces with high fiscal capacities to provinces with low fiscal capacities to ensure a uniform standard of public services across the country (Ouimet, 2014).

The paper is structured as follows. In Section 2, we provide a partial equilibrium analysis of a number of stylised regional fiscal regimes. These include systems in which the devolved region's budget is driven (at least in part) by: a block grant from the central government; tax sharing; a needs based grant (with an equalisation scheme embedded); full fiscal autonomy.

Critical determinants of the region's budget, including population and tax revenues, are treated as exogenous in the partial equilibrium analysis. In Section 3 we relax these constraints and provide a general equilibrium analysis of the likely macroeconomic consequences of alternative stylised fiscal regimes, using a simple model of a small, open, regional economy. Our focus here is on permanent, asymmetric shocks that impact a small regional economy. This context motivates our use of a standalone theoretical model of a regional economy.

Section 4 outlines the structure of our (standalone) forward-looking, multi-sectoral CGE model of Scotland, augmented to incorporate the key elements of the stylised fiscal systems that we consider in our theoretical analysis. This use of a single database and modelling framework allows us to isolate the intrinsic properties of the alternative regional fiscal systems.

In Section 5 we explore the impact of the spectrum of fiscal regimes on the response of the regional economy to both demand- and supply-side macroeconomic disturbances. Since one of the main motivations for further fiscal autonomy, in addition to improving accountability, is to provide devolved governments with an incentive to pursue growth enhancing (and avoid growth reducing)

[1] Here, whilst the Scottish Government has full authority over the rates and bands of all taxes levied on earned income, the UK Government retains control over the definition of the tax base and is responsible for all unearned income.

policies, we consider the implications of our results for these incentives. However, we also consider their implications for equalisation.

Section 6 is a brief conclusion.

2. Stylised versions of regional fiscal systems

We consider four main alternative stylised fiscal systems (and some variants of each), reflecting various degrees of fiscal autonomy and concerns over equalisation. This allows us to identify key characteristics of regional fiscal systems while abstracting from very detailed aspects of fiscal arrangements.

At one end of the spectrum we consider a number of block grant systems, where the regional government's budget is determined entirely by a fiscal transfer from the central government. Second, we consider intermediate cases in which there is some devolution and sharing of taxes (operating in conjunction with a block grant adjustment). Third, we analyse the case of full tax autonomy, where the regional government's devolved expenditures are met entirely by the revenues that it raises through taxation. Finally, we consider the case where the regional fiscal system incorporates needs-based grants, and so embodies an equalisation mechanism. While we often use UK examples of fiscal structures to motivate particular stylised representations of fiscal systems, these are illustrative; the principles we elucidate have general applicability.

Block grant systems

The Barnett formula. To illustrate the operation of a simple block grant system we consider the example of the Barnett formula, which, in principle at least, characterised the operation of the Scottish (Welsh and Northern Irish) fiscal system until very recently. The Barnett formula links the change in the block grant to Scotland to the change in public spending in the rest of the UK (RUK) times the ratio of the population of Scotland to that in the rest of the UK.

$$(1) \qquad \Delta B_t = \left(\frac{P_t^S}{P_t^r}\right) \cdot \ (G_t^R - G_{t-1}^R)$$

Where: B_t is the nominal value of the Scottish block grant; the superscripts S and R refer to Scotland and RUK respectively; P_t is population at time t; and G_t^R is the nominal value of government expenditures in RUK; Δ is the first difference operator.

If there is no change in public spending in rUK, as we assume throughout our present single-region analysis, the Barnett formula implies that the Scottish block grant effectively becomes fixed in nominal terms ($B_t = B_0$).

"Real" Barnett. The Barnett formula used to be expressed in real, rather than nominal terms, and we explore this case because it provides a useful benchmark that is neutral in terms of its effect on the impact of both demand and supply disturbances. Here the *real* level of government budget is fixed, and given our assumption about rUK, this implies:

$$(2) \qquad \frac{B_t}{cpi_t^S} = \frac{B_0}{cpi_0^S}$$

Where cpi^S_t is the Scottish consumer price index in period t. So the nominal Scottish budget in any period is updated to reflect any changes in prices since the base period:

$$(3) \qquad B_t = B_0 \cdot \left(\frac{cpi_t^S}{cpi_0^S}\right) = cpi^S_t dot$$

Clearly, if prices do not change, then Real Barnett and Barnett are equivalent. However, if prices are rising (falling) Real Barnett generates a higher (lower) Scottish budget. This introduces an asymmetry in the relation between the response of Real Barnett and Barnett to macroeconomic disturbances, as we shall see in the Section 3 below.

Barnett plus influence. There is considerable controversy over the way that the Barnett formula was operated in practice. Historically, Scotland has enjoyed a significant advantage in terms of total expenditure per capita as compared to rUK. If equation (1) was rigorously applied we would expect to see a gradual movement towards equal expenditure per capita across the regions of the UK over the longer term: the "Barnett squeeze". However, in practice, this has not been observed: the suggestion is that "formula plus influence" has led to the approximate maintenance of public expenditure per capita in Scotland (Christie and Swales, 2010). This "Barnett plus", regime implies:

$$(4) \qquad B_t / P_t^S = B_0 / P_0^S$$

$$(5) \qquad B_t = B_0 \cdot \left(\frac{P_t^S}{P_0^S}\right) = P^S_t \text{ dot}$$

"Barnett plus" implies that nominal Scottish government expenditure per capita is held constant: the political influence of the Scots on Westminster has been such as to offset the "squeeze" on public spending that would otherwise be expected from rigorous application of the Barnett formula. It is clear from (5) that under Barnett Plus the block grant is greater than/ equal to/ less than that under "Barnett" if Scottish population is increasing/ constant/ contracting.

In terms of macroeconomic disturbances the Barnett plus regime reinforces the impact of any population changes: any stimulus to (contraction in) population is reinforced by expansions (reductions) in nominal government spending (which, of course, do not occur under Barnett). So the multiplier under the Barnett plus regime is always greater than under Barnett, irrespective of the source or the direction of the disturbance.

Tax sharing (with block grant adjustments)

Smith regime. Our first example of a tax sharing system is the Smith regime, which introduced the devolution of income tax (excepting thresholds) and the assignment of a portion of VAT revenues resulting from the Scotland Acts of 2012 and 2014. Scotland's budget under Smith (S_t) is determined by three elements: the operation of the Barnett formula, as given above (fixed in nominal terms in the present context); plus all devolved tax revenues R_t^S; less a block grant adjustment (BGA_t):

$$(6) \qquad S_t = B_t + R_t^S - BGA_t$$

We call the combined effect of the last two variables in the expression on the LHS of equation (4) the Net Smith Transfer (NST). This is the additional transfer to Scotland, which can be positive or negative, above the "pure" (simple) Barnett transfer. This is the change in the fiscal position

negotiated through the Smith agreement. It equals the Scottish revenues from devolved taxes minus the BGA. Therefore:

(7) $$NST_t = R_t^S - BGA_t$$

This means:

(8) $$S_t = B_t + NST_t$$

The implications of shifting to the Smith from the Barnett regime for the Scottish budget are clearly of considerable interest for the Scottish Government. The Smith regime (equation 6) results in a larger Scottish budget than Barnett only if the devolved tax revenues that accrue to the Scottish Government exceed the block grant adjustment; that is the Net Smith Transfer (in 7) is positive. We now explore the conditions under which this condition is satisfied.

The block grant adjustment under Smith[2], BGA_t is defined in terms of the initial level of the BGA times the ratio of Scottish to RUK population growth times the growth in RUK revenues from the devolved tax base:

(9) $$BGA_t = BGA_0 \frac{P_t^S}{P_0^S}\frac{P_0^R}{P_t^R}\frac{Y_t^R}{Y_0^R}\frac{t_t^R}{t_0^r} = BGA_0 \dot{P}_t^S \dot{y}_t^R \dot{t}_t^R$$

In (9): Y is the devolved tax base, y is the per capita tax base and t is the tax rate. We use the convention that uppercase are absolute values and lowercase are ratios. The dot notation indicates proportionate growth, so that $\dot{Z}_t = \frac{Z_t}{Z_0}$

We assume for simplicity that the tax rates don't vary over time or between places so that:

(10) $$\dot{t}_t^S = \dot{t}_t^R = 1$$

This means that (7) can be expressed as:

(11) $$BGA_t = P_t^S \left[\frac{BGA_0}{P_0^S} \dot{y}_t^R \right] = P_t^S \left[\frac{BGA_0}{P_0^S} \dot{\lambda}_t^R \dot{\varepsilon}_t^R \right]$$

In the expressions on the LHS of equation (11) the term in the square brackets is the per capita BGA in period zero times the proportionate increase in the rUK devolved tax base. In the expression at the extreme RHS we introduce the ratios λ and ε where E is the employment level. The variable λ is tax base per employee and ε is the employment level as a share of the population - the employment rate - so that:

(12) $$\lambda = \frac{Y}{E}, \quad \varepsilon = \frac{E}{P}$$

and therefore

[2] This is known as the indexed per capita (IPC) approach to adjusting the block grant adjustment over time.

(13) $y = \lambda\varepsilon$

We also adopt the terminology that bga is the per capita Block Grant Adjustment, so that:

(14) $bga = \frac{BGA}{P}$

Using (3) and (6) gives:

(15) $\dot{bga}_t = \dot{y}_t^R$

The proportionate change in Scottish per capita BGA equals the proportionate per capita change in the rUK revenue from the devolved taxes (here equal to the proportionate change in the rUK relevant tax base).

Equation (11) gives an expression for the BGA. R_t^S is straightforward to express in terms of variables already defined, so that:

(16) $R_t^S = P_t^S y_t^S t_t^S = P_t^S \lambda_t^S \varepsilon_t^S t_t^S$

Substituting equations (11) and (16) into (8) gives:

(17) $NST_t = P_t^S nst_t = P_t^S \left(y_t^S t_t^S - bga_0 \dot{y}_t^R \right)$

Equation (17) again uses the convention that the lower case represents the value per capita so that:

(18) $nst_t = \frac{NST_t}{P_t^S}$

We are interested in the circumstances where the Scottish Government gets a larger transfer under Smith than under the pure Barnett, that is where $S_t > B_t$. Using (8) and (17), it is straightforward to show that:

(19) $S_t > B_t \rightarrow NST_t > 0 \rightarrow nst_t > 0$

This means that the direction of the change in absolute Scottish revenues from introducing Smith depends on the sign of the per capita Net Smith Transfer. In the base year, NST_0, is calibrated to take the value zero. That is to say:

(20) $bga_0 = y_0 t_0^S$

In subsequent periods, for the Scottish Governments revenues to be greater under Smith than "pure" Barnett, using (10), (17), (18) and (19) provides the requirement that:

(21) $y_t^S t_t^S > bga_0 \dot{y}_t^R \rightarrow \dot{y}_t^S > \dot{y}_t^R$

Equation (21) is the fundamental equation for determining whether the Scottish Government receives more revenue under "pure" Barnett or under Smith. The condition is simply that the growth

of the devolved per capita tax base in Scotland should be greater than the per capita growth in the same tax base in rUK. Therefore when we consider whether a particular policy from the Scottish Government will be incentivised by the Smith changes, we only need consider whether that policy will affect the nominal per capita tax base in a positive manner. Note that the growth in the Scottish population affects the absolute size of the Net Smith Transfer (positive or negative) but not directly the sign (see equation 17). Therefore a policy which positively increases the Scottish population but reduces the per capita tax base would give a better outcome for the Scottish Government under "pure" Barnett.

Equation (13) can be used to decompose the expressions in (21). Therefore

(22) $\dot{\lambda}_t^S \dot{\varepsilon}_t^S > \dot{\lambda}_t^R \dot{\varepsilon}_t^R$

Given the exogeneity of rUK here, we set $\dot{\lambda}_t^R \dot{\varepsilon}_t^R = 1$, so that for Smith to incentivise the Scottish Government requires that:

(23) $\dot{\lambda}_t^S \dot{\varepsilon}_t^S > 1.$

Note that the dependence of the outcome on $\dot{\varepsilon}_t^S$ implies that our assumptions concerning the relationship between changes in employment and population (through migration) becomes important. If we assume that the long-run equality between the growth in the labour force and the growth in employment required to maintain the long-run constant unemployment rate also implies a similar proportionate increase in population, then for all long-run analysis with migration $\dot{\varepsilon}_t^S = 1$, all we require for "correct" incentives – that is incentives to grow that are greater under Smith than under Barnett - is:

(24) $\dot{\lambda}_t^S > 1.$

The IPC method currently being used to determine the size of the block grant adjustment (equation 9) has two properties that were considered desirable by the Scottish Government. First, if, as is projected to be the case for the foreseeable future, population growth is lower in Scotland than in rUK, the population growth term on the RHS of equation (9) is less than unity, and so reduces the scale of the BGA.

A second feature of the BGA under Smith that is considered beneficial to Scotland is the fact that it is governed by relative growth in per capita tax revenues: in particular, symmetric UK-wide shocks impact on R_t^R as much as on R_t^S so should not unduly adversely impact the Scottish Government's budget (although there will be some impact if effects are not proportional to relative populations).

It may be worth elaborating a little on the distinctive responses of the Smith and Barnett regimes to macroeconomic shocks. Any macroeconomic disturbance that tends to attract in-migration and so increase population in Scotland would also tend to increase the size of the block grant adjustment (equation 9). While such disturbances also stimulate devolved tax revenues, in-migration increases the probability that the impact on the BGA might exceed the stimulus to devolved tax revenues. This is the source of the result that the Scottish Government's budget could actually be lower under

Smith than it would be under Barnett and thus the incentive to grow the economy could be less under Smith than under Barnett.

Treasury mechanism. The Treasury's preferred method of updating the block grant adjustment method did not have the term in relative population growth in equation (9), and while it was not actually implemented, it may be in future. This so-called Comparable Method (CM) approach to adjusting the block grant through time preferred by the Treasury implies:

$$BGAT_t = BGA_0 \cdot \left(\frac{Y_t^R}{Y_0^R}\frac{t_t^R}{t_0^R}\right) \quad (25)$$

Under the Treasury method of indexation total Scottish government tax revenues would have to grow more rapidly than those in rUK for the Scottish budget to be greater under Smith than under Barnett. This was thought to be an overly demanding requirement, especially given the expectation of lower population growth in Scotland than in rUK. Of course, while the current formula effectively insulates Scotland from lower population growth, it does nothing to offset the projected more rapidly ageing of the Scottish relative to the rUK populations, or any other source of adverse demographic impact (Eiser, 2017).

Note that given that rUK is exogenous in the present analysis the second term on the RHS of (25) is unity, so that the Treasury block grant adjustment is fixed at BGA_0 in these circumstances. Accordingly, the change in the Scottish budget in response to any macroeconomic disturbance will simply be equal to the induced change in devolved tax revenues, as inspection of equation (6) confirms. There is effectively no additional block grant adjustment arising in response to any permanent, asymmetric macroeconomic disturbance to the Scottish economy under the preferred Treasury formulation. We therefore expect that the Treasury version of the BGA will result in bigger reinforcement of macroeconomic shocks than is true under Smith. The incentive to grow the Scottish economy is greater under the Treasury than the Smith BGA, but the system also allows movement away from equalisation.

Full tax autonomy

Here we consider the case of a region that has full fiscal (specifically tax) autonomy and in which all tax revenues raised within the region remain there. As noted in the Introduction there are few examples of regional economies that have such autonomy. A stylised representation of a region with full fiscal autonomy provides us with a useful limiting case with which we can compare other regional fiscal systems.

In the Scottish context we are effectively assuming that North Sea Oil revenues would be sufficient to cover Scotland's fiscal deficit vis a vis the Westminster government.[3] (Accordingly, we abstract from the fact that, currently at least, Scotland's fiscal deficit would almost certainly have to be partially funded from tax revenues raised in Scotland.) In this case the Scottish budget in any period, which we take to be balanced, would simply equal the sum of all tax revenues generated in Scotland:

$$FTA_t = TR_t^S \quad (26)$$

[3] Alternatively we could consider a system in which tax is fully devolved, but welfare payments continue to be funded by the centre. This is effectively the case for Scotland, at least as far as benefits that are likely to be sensitive to economic activity. In this case there is effectively no fiscal deficit that requires financing.

Where FTA is the government budget under full tax autonomy and TR^S is the sum of *all* tax revenues raised in Scotland. Details on devolved and total taxes are provided in the discussion of the model in Section 4.

Clearly, all tax revenues are greater than currently devolved revenues:

(27) $TR_t^S > R_t^S$

and the response of revenues to changes in Scottish economic activity is enhanced since many of the currently non-devolved taxes are positively related to economic activity, including corporation tax. Accordingly, the impact of macroeconomic disturbances are reinforced more strongly in this case than under the Treasury BGA model, which in turn is associated with a higher multiplier than the Smith model.

In terms of incentivising growth the tax sharing cases can be ranked unambiguously and both are naturally inferior to the full tax autonomy case on this criterion. Notice here that the ranking of multipliers associated with the regime does not depend on the source or direction of the disturbance. For all disturbances it is the case that the multiplier is greatest under full tax autonomy, followed by the Treasury grant adjustment regime and then Smith. However, equalisation varies inversely with the scale of the multiplier.

This symmetry of responses across the direction and source of disturbances does not hold for variants of block grants. Furthermore, comparisons of tax sharing and block grant systems are more complex, as we have seen.

Equalisation: needs-based grants

As noted in the introduction, a number of modifications to grant systems are based upon concerns about inequalities developing across regions. We illustrate the operation of an equalisation regime by augmenting the block grant mechanism with an equalisation formula intended to reflect need in broad brush terms. Here we assume It takes the form:

$$(28) \qquad E_t = B_0 - EBGA_t$$

Where E is the budget under the equalisation regime and EBGA is the block grant adjustment under that regime:

$$(28) \qquad EBGA_t = \beta \cdot \frac{(Y_t/P_t)}{(Y_0/(P_0))} \cdot B_0$$

The EBGA is bigger the greater the growth in GDP per capita (relative to other regions). We parameterise this relationship using the equalisation regime in Australia (Poghosyan et al, 2016).

The equalisation term moderates the impact of any disturbance as compared to the pure Barnett formula since it reduces the budget in the face of asymmetric expansion and increases it in response to an asymmetric contraction. The multiplier is always smaller under this regime than under Barnett, as is the incentive to grow the regional economy. However, it does act, unsurprisingly, to equalise budgets across regions.

While we focus on adjustments to Barnett here, we could add the equalisation adjustment to any of the other stylised fiscal regimes. It would have a similar impact of reducing the scale of e.g. the tax autonomy or Smith multipliers, and promote equalisation although at the expense of reducing the incentive to grow the region.

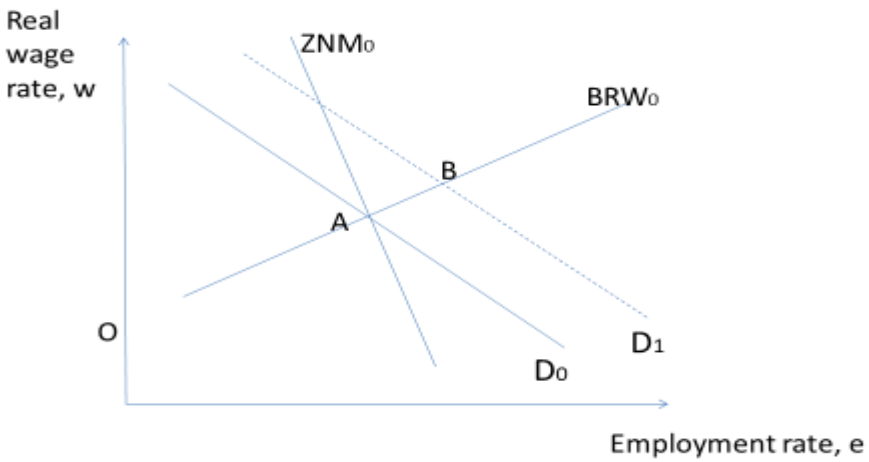
We turn next to a system-wide or macroeconomic analysis of these alternative fiscal systems.

3. Comparing the system-wide consequences of the alternative fiscal regimes

Our interest is in the way in which alternative regional fiscal systems impact the response of the regional macro-economy to supply- and demand-side disturbances. Clearly, this cannot be fully achieved through the partial equilibrium analysis since that is silent on how either Scottish tax revenues or Scottish population, for example, are determined yet, as we have seen, these are critical proximate determinants of the sub-central region's budget. In this section we provide a diagrammatic analysis that allows us to compare the macroeconomic impact of a number of stylised decentralised fiscal systems in a simple model of an open regional economy (Layard, Nickell and Jackman, 1991; Lecca et al, 2014) that are likely to alter responses to macroeconomic disturbances. We focus here exclusively on permanent, *asymmetric* disturbances to a *small* regional economy so that we can concentrate solely on the region of interest, treating other regions as exogenous to the analysis. This simplifies the analysis considerably.

In Figure 1 the initial long-run equilibrium occurs at point A where the regional (bargained real) wage curve (BRW_0) intersects the zero net migration condition (ZNM_0) and the labour demand curve (D_0) in employment rate (e) – real wage rate (w) space. The slope of the BRW curve reflects the fact that the greater is workers bargaining power, as reflected in e, the higher the bargained real wage will be. For zero net migration, however, a high real wage would have to be offset by a lower employment rate. Above and to the right of ZNM, w and/or e are "too high" to meet the zero net migration condition: there will be in-migration, which will tend to reduce w and e (while increasing the levels of population and employment – see below), pushing them back to a point on ZNM. The labour demand curve is a general equilibrium relationship, which incorporates all of the effects of a change in real wages, including those on consumption demand. While it is not inevitable, it is likely that this curve is negatively sloped for an open regional economy with competitiveness effects dominating the income effects of wage rises (as in our default empirical model).

Figure1. The impact of an export stimulus on the real wage and employment rates.



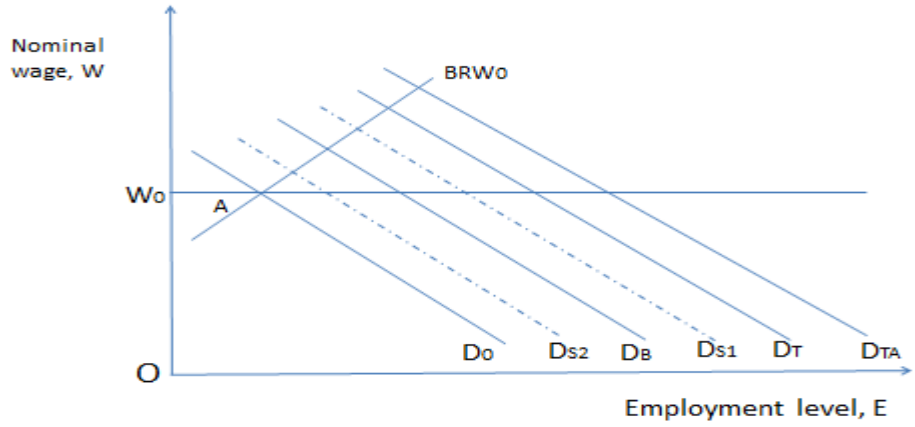
Consider now the impact of a stimulus to demand in the form of an increase in exports. In all of our stylised fiscal systems an increase in exports shifts the labour demand curve to the right (to D_1). Of course, the extent of this shift will vary with the extent to which the stylised fiscal system reinforces or offsets the stimulus to exports, an issue we explore shortly. However, for now it suffices to note that in all cases the stimulus to demand tends to increase employment and the employment rate, which results in an upward push on the real wage. In the short-run, before migration responds there is a shift from A to B, which is associated with a higher w and e, and so with net in-migration (since ZNM_0 is unaffected). This increase in population gradually generates a shift backwards in the demand curve for labour. Given the flow migration assumptions underlying ZNM this process continues until the system returns to point A at which real wage and employment rates return to their original levels. At this new long-run equilibrium nominal wages and prices have also returned to their original levels, although population and employment are permanently increased.

The impacts on the *level* of employment are clearer from Figure 2. Here the relevant curves are drawn in employment level (E), *nominal* wage (W) space. As before the initial equilibrium position is at A. In the present case we know that all long-run equilibria will be established along the horizontal line going through A (at W_0); demand disturbances ultimately have no long-run impact on wages or prices in this model (McGregor et al, 1996).[4] For simplicity we assume that exports and current government spending have no supply side impacts; only aggregate demand is affected. Recall that, in all cases the employment rate is invariant in the long-run, so that population moves equi-proportionately with the level of employment throughout. In all cases the demand expansion increases GDP, employment and tax revenues, although the extent depends on the precise nature of the fiscal regime.

Under Barnett the (long-run) general equilibrium demand curve shifts out to D_B. Since prices do not change in the long-run, here Real Barnett ultimately generates an identical budget to Barnett. We know that in this case the only source of a stimulus to demand is the increase in exports; real (and nominal) government expenditure is fixed, so there is no induced fiscal change.

[4] Of course, in the short-run wages and prices do change. We consider this in our sensitivity analysis below.

Figure 2. The long-run impact of an export stimulus on the employment and the nominal wage.



The Barnett Plus regime with fixed nominal government expenditure per capita will clearly generate a demand curve that lies to the right of that associated with Barnett given the stimulus to population (not shown), but we cannot unambiguously rank its impact relative to the tax sharing/ autonomy regimes.

The relative location of the labour demand curve under Smith is ambiguous given our partial equilibrium analysis. If the impact of the expansion in export demand is to stimulate devolved tax revenues more than proportionately to Scottish population, the relevant curve would lie to the right of D_B (but to the left of D_T), like D_{S1} in Figure 2. However, if population is stimulated more than proportionately to tax revenues the curve would lie in a position like D_{S2}, which implies a smaller expansion than under Barnett.

We know that under the Treasury CM method of adjusting the block grant, the demand curve must lie at a position like D_T, which lies to the right of D_B and D_{S1}, since additional tax revenues and current government expenditures are generated in this regime. The labour demand curve under full tax autonomy lies even further to the right, at D_{TA}, for in this case all tax revenues are stimulated, not just those currently devolved (as in D_T).

The impact of the equalisation case is strictly ambiguous. However, If GDP per capita rises, as seems most likely given that GDP is more directly impacted by employment than by population and that it reflects increased capital as well as labour input, then the demand curve must lie to the left of D_B. (if GDP per capita falls, the equalisation mechanism adds to the Scottish budget, stimulating demand by more than under Barnett.) The scale of the differences naturally depends on the extent of equalisation implied by the adjustment formula.

The implications for the scale of the expansion of the economy under our stylised fiscal regimes can simply be inferred from Figure 2 and the preceding discussion. Note, however, that this is also an indicator of the strength of the incentive to the regional government to stimulate regional growth. Of course, there is an incentive to encourage exports under all regimes since output and employment are increased as a consequence. However, the scale of that incentive is clearly

enhanced by full tax autonomy, since here the additional activity generated is at its greatest. However, it is not the case that Smith always provides a greater growth incentive than Barnett. The growth incentive is greater under the fixed nominal expenditure per capita case than under Barnett, since the increase in population results in higher public spending. The growth incentive associated with the equalisation regime is diminished relative to the Barnett case, since the latter tends to moderate the impact of any disturbances. (This would be true of any of our cases augmented by an equalisation adjustment.)

However, this last observation serves to draw attention to the fact that fiscal systems are often also assessed on their capacity to promote equity across regions. In the present case, of a beneficial asymmetric demand stimulus, the target region's economy naturally improves relative to all other regions. Only in the presence of equalisation is this tendency consciously constrained by the regional fiscal system. However, here equity moves inversely to the region's economic activity and to an extent governed by the same factors that provide a positive incentive to grow.

It is worth noting the implications of our analysis for a contraction in demand. The analysis would be very similar, and establish that the shock-absorbing capacity of a region is ranked inversely to incentives to grow its economy: tax autonomy, for example, would provide the least insulation against the impact of external demand disturbances. The Smith and Treasury regimes would be associated with similar qualitative results on the relative scales of multipliers, with the Treasury case being associated with a greater multiplier. With a contraction there is downward pressure on prices (and so on the budget associated with Real Barnett), whereas in an upturn prices tend to rise (at least initially) and increase the budget associated with Real Barnett, so again it has a bigger multiplier than Barnett.

Different fiscal systems are associated with demand "multipliers" of varying scale; whether this is judged to be an inherent advantage or not is likely to depend on the state of the regional economy and the sign of any disturbances.

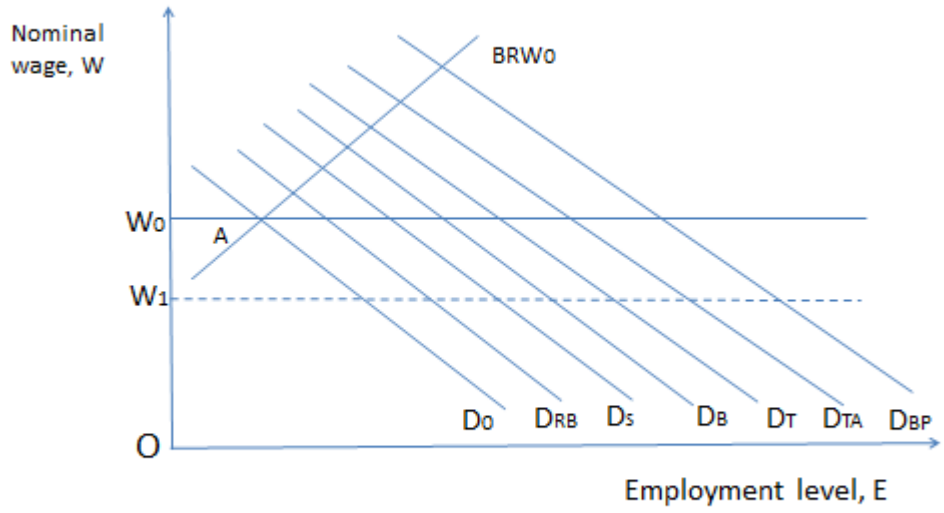
Next consider the impact of a supply-side stimulus, specifically an improvement in labour efficiency. This reduces the cost of an efficiency unit of labour and stimulates the demand for labour in efficiency units. However, the impact on the demand for labour in physical units (employment) is ambiguous. If the demand for labour is wage-elastic then the percentage increase in the demand for labour in efficiency units will exceed the percentage increase in labour efficiency and employment increases. In fact, in our default empirical model this is typically what happens in the long-run, after population and capital stocks are fully adjusted. In this case, the impact of the improvement in labour efficiency on real wage and employment rates can again be inferred from Figure 1: a stimulus to the real wage and employment rates will ultimately be offset by net in-migration, which ultimately restores the initial levels of the real wage and unemployment rates.

However, the analysis of the impacts on the levels of nominal wages and employment differs somewhat from the export case. In particular, although the real wage is ultimately unaffected, the nominal wage and prices fall (equi-proportionately) in response to increased production efficiency (for example to W_1 in Figure 3). The scale of this change is independent of the fiscal regime, as is therefore the stimulus to exports that is generated by the improvement in competitiveness (which is, of course, absent in the case of the pure demand stimulus).

The scale of the stimulus to demand does not follow the same pattern as for the export demand shock in part because most of the fiscal regimes are defined in terms of *nominal* values. This is a neglected, but potentially important, feature of regional fiscal systems. The long-run impact of Barnett for an export shock was identical to that for the case where government expenditures were held constant in real terms, since prices do not change in that case. Here prices fall, however, which leads to a fall in nominal government spending to preserve spending in real terms for the Real Barnett case; here the sole stimulus to demand comes through the productivity improvement. However, the fall in prices results in a substantial increase in real government spending in the fixed nominal expenditure (Barnett) case. This causes the demand curve corresponding to Barnett to shift substantially to the right of that associated with fixed real government spending (Real Barnett, so that, $D_B > D_{RB}$).

We now cannot be certain that the full fiscal autonomy case is associated with a greater stimulus to demand than Barnett (although it is in our simulations), but it must exceed the demand stimulus implied by Real Barnett (since activity is rising as are real tax revenues). (So we have $D_{TA} > D_B > D_{RB}$.) Of course, it must be the case that equalisation moderates the impacts relative to Barnett, so that the demand curve must lie to the left of D_B (but is not shown).

Figure 3. The long-run impact of a productivity stimulus on employment and the nominal wage.



In response to the productivity shock it is possible (and proves to be the case in our simulations) that nominal tax revenues and expenditure actually fall, while population is increasing, so that Smith is associated with a *smaller* stimulus than Barnett ($D_S < D_B$) . The maintenance of nominal government expenditure per capita under Barnett Plus must, of course, generate a stimulus greater than Barnett since population is increasing and so are nominal government expenditures in this case, implying

even greater rises in real government spending. In fact, this provides the biggest stimulus in our simulations (illustrated by D_{BP} in Figure 3).

It is quite clear that the scale of the growth incentive across fiscal regimes may vary significantly depending on the nature of the disturbance; the ranking of multipliers may vary with the source of disturbance. In particular, regimes with nominal rigidities can significantly augment the multiplier for supply-side disturbances and can, surprisingly, provide greater supply-induced growth incentives than even full tax autonomy. The nature of the regional price-indexation – or lack of it – in decentralised regional fiscal systems appears to be a neglected but potentially significant issue.

Our theoretical analysis is instructive and suggestive of the likely impact of alternative stylised fiscal systems on the responsiveness of the regional macro-economy to asymmetric, permanent supply and demand disturbances. There is a similar impact on the incentives to pursue growth policies that devolved governments face under these regimes. While theoretical analysis aids understanding it is limited in a number of respects. First, it cannot provide a clear ranking of the multiplier effects of all fiscal regimes (for example, Barnett and Smith), even for demand disturbances that ultimately have no impact on prices, since these depend on the *relative* strengths of the impact of disturbances on Scottish population and devolved tax revenues. Secondly, it is even more difficult to provide a complete ranking of the impact of alternative fiscal regimes on supply-side disturbances, because here prices do change, and many fiscal regimes are defined in nominal terms, significantly complicating the comparisons. However, in this case the probability of Barnett actually providing greater growth incentives than Smith is even more likely than for the export disturbance. And some block grant regimes may provide even stronger incentives for supply-induced growth than full tax autonomy. Finally, the analysis gives no indication of the *scale* of the likely effects of alternative regional fiscal regimes on the incentive to grow the region, yet this would appear to be crucial information for those contemplating a move to greater fiscal autonomy. To make further progress, we need to embed the key features of the alternative fiscal regimes within a numerical, general equilibrium model of the Scottish economy.

4. The CGE model

Outline of the model structure

The core AMOS CGE model used here is a development of one that has seen numerous applications (for example, Emonts-Holley et al, 2019; Lecca et al, 2013, 2014), so we provide only a very brief outline here. (A full model listing is available in Emonts-Holley et al, 2019.) This version of the single-region model forms the basis for integrating intergovernmental transfers and the devolution of taxes under a variety of stylised regional fiscal regimes. The model allows us to explore the properties of these stylised regimes in the context of a common modelling framework and data base. It is calibrated on a Scottish social accounting matrix for 2014.

This AMOS model has twenty-five sectors, and three domestic institutions; households, firms and government. External institutions are represented by Rest of the UK (RUK) and Rest of World (ROW).

In the myopic variant households are modelled as a single representative agent with myopic expectations and an exogenous savings rate. In the forward-looking model households have perfect foresight and consumption decisions reflect intertemporal optimisation. Aggregate household

consumption is allocated through a CES function. All international trade in the model adopts Armington assumptions, whereby imports, domestic production, RUK and ROW imports are considered imperfect substitutes.

In the default version of the model wages are determined in accordance with a regional bargaining function or real wage curve. This implies a negative relation between the bargained net of tax real wage and the regional unemployment rate. (Layard et al, 1991; Blanchflower and Oswald, 1994)

The model used here generally incorporates interregional migration based on the specification in Layard et al. (1991) and Treyz et al. (1993), where the net migration flow is positively related to the gap between the log of regional and national real wages, and negatively related to the difference between the log of regional and national unemployment rates.

In each sector gross output is produced by combining intermediate inputs, and value added. Labour and capital are combined in order to produce value added. Typically we use CES production functions which allow for a degree of substitution among factors of production. The demand for capital and labour is obtained from the first order condition for cost minimisation. This implies that the demand for both factors of production is positively related to the volume of value added and is a decreasing function of their prices.

Each industry in the region produces goods and services that can be exported or sold in the regional market. Intermediate goods that are produced locally (in Scotland) or imported are considered as imperfect substitutes. Regional and imported intermediate goods are mixed under the Armington assumptions through a CES function.

Investment goods are, likewise, produced domestically and can be imported (from RUK and ROW). Within each period the distribution between domestically produced and imported investment goods through the Armington CES function.

Capital stock adjustment is assumed to be subject to adjustment costs so that actual capital stocks adjust to their desired levels only gradually. In the myopic version of AMOS desired capital stock is derived from cost minimisation and is a function of value added and the price of value added relative to the user cost of capital. Investment closes a proportion of the gap between actual and desired capital stocks in each period. In the forward-looking case investment is the outcome of an intertemporal optimisation process. (See Hayashy (1982), Go (1994) and Devarajan and Go (1998)).

Specification of the stylised regional fiscal regimes

This section outlines how we operationalise the stylised regimes analysed in Sections 2 and 3 above. Government is split into two levels: the central Government (CG) in the autonomous region and sub-central Government (SCG). The key assumption in the model is that the SCG is sufficiently small to ensure that any overall regional deficit or surplus is sustainable.

We do not impose strict constraints on the accumulation of foreign and public debt, because interregional transfers can sustain perpetual capital flows. However, we can track implied foreign and public deficits and debts and impose sustainability constraints if required.

[Note to reader: the notation here has yet to be reconciled with that in Section 2 above - this will be corrected in the next draft of the paper. The reader can avoid any possible confusion by moving directly to Section 5 below.]

The CG is assumed to be exogenous, and so its behaviour does not respond to changes in regional prices, government revenues, or the behaviour of the SCG. Its transfers (inter-governmental transfer payments and welfare payments) are indexed to RUK prices. Since these are exogenous, all CG transfers are fixed in nominal terms. It is assumed the CG consumption of each good *i*, $QG_{i,t}$, is fixed in real-terms to its initial expenditure profile, $\delta_i^g \cdot GE_{t=0}$ where δ_i is a share parameter.

The SCG has a balanced budget, and its expenditure, SGE_t, is will be constrained by its income, SGY_t.

$$SGE_t = SGY_t \quad \textbf{(A.1)}$$

$$SQG_{i,t} = \delta_i^{sg} \cdot SGE_t \quad \textbf{(A.2)}$$

$$QG_{i,t} = \delta_i^{g} \cdot GE_{t=0} \quad \textbf{(A.3)}$$

Simple Blockgrant Adjustment

A central block grant adjusted periodically to reflect population shares, characterises the traditional UK system of regional funding. Intergovernmental grants, known as the "Barnett formula" (Keep, 2018), in effect, provided the only source of revenue for a SCG. Simply put, the Barnett system of intergovernmental grant to regions is determined as a population share of relevant expenditures.

The classic "Barnett case" is, therefore, a situation whereby the SCG revenue, is fixed by the intergovernmental transfer between the central government ($IG_{t=0}$),

$$SGY_t = IG_{t=0} \quad \textbf{(A.4)}$$

This conforms to a great deal of systems in place around the world.

Devolving Taxes and block grant adjustments

In Scotland, the devolution of tax revenues (HM Government, 2015) was complicated by the issue of how to maintain a block grant system and devolve taxes. A huge amount of debate between the SCG and CG took place over how a block grant adjustment indexation mechanism should operate to changes in income tax revenue. This is due to the idea that without an indexation mechanism Scotland would not have an incentive to increase income tax revenues (Bell et al., 2016).

The final, agreed upon Scottish block grant adjustment indexed a block grant adjustment to income tax per capita growth. In short, if Scottish income tax per capita grows at the same rate of income

tax per capita as the rest of the UK, the Scottish budget would be the same as without tax devolution. This "Smith scenario" can be defined in the following system,

$$SGY_t = IG_{t=0} - BGA_t + (IT_t^{SG} + \sum_i VAT_{i,t}^{SG}) \quad \textbf{(A.5)}$$

$$BGA_t = \alpha \cdot \frac{\Delta LS}{LS_{t=0}} \cdot \left(IT_{t=0}^{SG} + \sum_i VAT_{i,t=0}^{SG} \right) \quad \textbf{(A.6)}$$

Where the block grant adjustment, BGA_t, is deducted from the nominally fixed Barnett grant and revenues from the devolved components of income tax and VAT ($IT_t^{SG} + \sum_i VAT_{i,t}^{SG}$) is added to the Scottish budget constraint.

Our model specification, while containing some interregional movement of workers, has a considerable amount of the population "exogenous". The parameter, α, attempts to capture endogenous population change (e.g., the proportion the population captured by endogenous changes in labour supply, LS) captured by the model. It is assumed that $0 < \alpha \leq 1$, which indicates that a proportion of the population is captured by changes in labour supply, and that $\alpha = 0.65$ (the proportion of the working-age population in Scotland) in the default case. It is theoretically possible $\alpha > 1$. This would indicate that the population increases at a faster rate than labour supply in periods of net in-migration, and is considered unlikely, however.

For simplicity, the block grant adjustment nomenclature is used, and the adjustment can be considered static or dynamic. A Smith case, with no block grant adjustment, would therefore replace A.6 with A.7, and the initial devolved tax revenues would be deducted from the blockgrant.

$$BGA_t = BGA_{t=0} = IT_{t=0}^{SG} + \sum_i VAT_{i,t=0}^{SG} \quad \textbf{(A.7)}$$

A similar approach could be deployed to "devolve" any tax revenues from other taxes on labour like national insurance, NI_t, to indirect business taxes, IBT_i. All taxes could even be devolved, and the SCG budget constraints in A.5 and A.6 are replaced with;

$$SGY_t = IG_t - BGA_t + (IT_t^{SG} + \sum_i VAT_{i,t}^{SG} + NI_t + \sum_i IBT_i + \cdots) \quad \textbf{(A.8)}$$

$$BGA_t = BGA_{t=0} = IT_{t=0}^{SG} + \sum_i VAT_{i,t=0}^{SG} + NI_{t=0} + \sum_i IBT_i + \cdots \quad \textbf{(A.9)}$$

Other systems (Barnett Plus and Equalisation)

In practice, there are a number of ways in which an intergovernmental transfer might be adjusted. In many cases, however, the mechanism by which an intergovernmental grant is adjusted is exogenous to SCG policy making. These cases in situations in funding systems which depend wholly on economic or demographic dynamics that occurs outside the autonomous region (i.e., in situations where the intergovernmental grant comprises a "pot" which is then shared between regions).

This is not always the case, however. In the UK, for example, the Barnett system is often considered "flawed" but primarily because it has resulted in per head government expenditure differences across the UK, and so UK central government macro policy does not affect every region equally (Institute of Fiscal Studies, 2014).

It has been argued (Christie and Swales, 2010) that the Barnett system of intergovernmental grants might be considered one which fixes intergovernmental grants per head (relative to UK per head government spending). Under such a scenario it would, theoretically, mean that economic activity in the autonomous region would affect the size of the block grant. In short, a "Barnett 2 case" would constitute an intergovernmental grant indexed to incremental changes in the population, and in the case of the AMOS model would yield,

$$GY_t^{SCG} = IG_t - BGA_t \quad \textbf{(A.10)}$$

$$BGA_t = -\left(\alpha \cdot \frac{LS_t}{LS_{t=0}}\right) \cdot IG_{t=0} \quad \textbf{(A.11)}$$

These block grant adjustments can operate under various forms. In Australia, they perform a "needs-based" function. Here, all forms of government transfers operate on a needs assessment. Needs-based block grants are – to some extent – responsive to endogenous pressures. Poghosyan et al (2016) found that block grant allocations (and even household transfers) were inversely related to GDP per capita. This would yield,

$$BGA_t = \beta \cdot \frac{(GRP_t/(\Delta LS + LS_{t=0} \cdot \alpha^{-1}))}{(GRP_{t=0}/(LS_{t=0} \cdot \alpha^{-1}))} \cdot IG_{t=0} \quad \textbf{(A.12)}$$

where the β value is derived empirically and has a value of 0.07 for intergovernmental grants and 0.16 for household transfers (Poghosyan et al, 2016).

5. Simulation results

We have provided both ex ante partial and general equilibrium analyses of the response of a range of stylised regional fiscal systems to demand- and supply-side disturbances. Our empirical model allows us, at least in principle, to identify the likely balance of the countervailing forces identified in this analysis and provides estimates of the scale of impacts. For policy choices it is likely to matter how much a shift in the degree of fiscal decentralisation changes the incentive to grow the host regional economy.

Here we use the default CGE model outlined in Section 4 to provide an empirical analysis of the impact of alternative stylised regional fiscal systems on the effects of demand and supply side disturbances. Specifically, we simulate 5% increases (and decreases) in exports and in labour productivity and compare the macroeconomic outcomes across the four main types of regional fiscal system: block grant, tax sharing, full tax autonomy and equalisation regimes. While the positive shocks could, for example, be interpreted as the outcome of successful, Scottish-specific, industrial policies we simply take them here to be illustrative of exogenous demand- and supply-side disturbances. Recall that our focus here is on permanent, asymmetric disturbances that impact Scotland, but not rUK; the results should be interpreted accordingly.

The impact of alternative fiscal regimes on the response to a rise in the demand for exports

The long-run results for the impact of a 5% stimulus to Scottish exports using our default model are reported in Table 1.The broad-brush qualitative results of the export demand stimulus are similar across fiscal regimes and align with our theoretical analysis. The stimulus to exports increases the demand for Scottish produced goods, increasing regional GDP and employment. However, the scale of the impact does typically vary across fiscal systems. For simplicity we focus exclusively here on the long-run equilibria of our default model, which is characterised by flow migration and regional wage bargaining that focuses on the real post-tax wage. The importance of these assumptions is explored in our sensitivity analysis.

As implied by our theoretical analysis of system-wide effects some key characteristics are shared across all fiscal regimes. In particular, the equilibrium levels of real wage and unemployment rates are ultimately unaffected by the export stimulus, as are prices and nominal wages. In the long-run such a system operates like an augmented input-output (IO) model (McGregor et al, 1996). Accordingly, exports increase by the full 5% stimulus in the long-run across all regimes. (This is not true in the short-run since it is characterised by the presence of capacity constraints, and wage and price adjustments, as we discuss in our sensitivity analysis.)

Table 1. The long-run impact of a 5% stimulus to exports across stylised regional fiscal systems (% changes from base)

	Real Barnett	Barnett	Barnett Plus	Smith	Treasury	Full Tax Autonomy	Equalisation
GDP (£m)	2.50%	2.50%	2.99%	2.58%	2.79%	3.47%	2.47%
GDP per capita	0.99%	0.99%	1.09%	1.00%	1.05%	1.19%	0.97%
Household Consumption	1.36%	1.36%	1.68%	1.42%	1.55%	2.00%	1.31%
Investment	2.72%	2.72%	3.09%	2.78%	2.94%	3.46%	2.69%
Total Exports	5.00%	5.00%	5.00%	5.00%	5.00%	5.00%	5.00%
Export RUK	5.00%	5.00%	5.00%	5.00%	5.00%	5.00%	5.00%
Export ROW	5.00%	5.00%	5.00%	5.00%	5.00%	5.00%	5.00%
Total Imports	2.32%	2.32%	2.68%	2.38%	2.53%	3.04%	2.29%
WAGES:							
Labour cost WF	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
Gross HH Wage WHG	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
Net HH Wage WHN	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
Real Labour cost WF/CPI	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
Real Gross HH Wage WHG/CPI	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
Real Net HH Wage WHN/CPI	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
PRICES:							
CPI	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
Govnt price index	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
Value Added Price index	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
Export price index	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
Repl cost of capital (PINV)	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
User cost of capital (PINV)	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
LABOUR MARKET:							
Unemployment level	6.00%	6.00%	6.00%	6.00%	6.00%	6.00%	6.00%
Unemployment Rate (pp difference)	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
Employment	2.31%	2.31%	2.89%	2.41%	2.65%	3.47%	2.27%
Labour supply	2.31%	2.31%	2.89%	2.41%	2.65%	3.47%	2.27%
Population	1.50%	1.50%	1.88%	1.56%	1.72%	2.26%	1.48%
NOMINAL TAX REVENUE:							
Indirect tax IBT	2.74%	2.74%	3.05%	2.79%	2.92%	3.36%	2.71%
Total HH Tax	1.84%	1.84%	2.30%	1.91%	2.10%	2.75%	1.78%
Employer NI	2.31%	2.31%	2.89%	2.41%	2.65%	3.47%	2.27%
Employee NI	2.31%	2.31%	2.89%	2.41%	2.65%	3.47%	2.27%
Income Tax	2.31%	2.31%	2.89%	2.41%	2.65%	3.47%	2.27%
Corporation Tax	2.72%	2.72%	3.09%	2.78%	2.94%	3.46%	2.69%
VAT Revenues	1.02%	1.02%	1.60%	1.11%	1.36%	2.18%	0.97%
Council Tax	1.36%	1.36%	1.68%	1.42%	1.55%	2.00%	1.31%
Smith Taxes / Head	0.41%	0.41%	0.61%	0.44%	0.52%	0.80%	0.39%
SCOT GOV SPENDING:							
Transfers to HH from Gov	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
Transfers to HH from RUK	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	-0.16%
Scottish Government Expenditure	0.00%	0.00%	1.88%	0.31%	1.10%	3.75%	-0.07%
Real Scottish Government Consumption	0.00%	0.00%	1.88%	0.31%	1.10%	3.75%	-0.07%

The first column shows the long-run impact of the export shock under the Real Barnett regime, within which real government expenditure is fixed. Here this yields identical results to those for the Barnett regime, characterised by fixed nominal government expenditures, which are reported in column 2, since prices do not change.

As expected, the export stimulus raises aggregate demand, which tends, in the short-run (not shown), to push up the real wage and reduce the unemployment rate as the labour market tightens and induces in-migration. This ultimately restores the real wage and the unemployment rate to their original values. However, there are substantial increases in GDP and employment (of 2.48% and 2.29% respectively), but ultimately no change in prices or wages. Of course, in this case the sole stimulus to demand is through exports; in particular, there is no induced fiscal impact: nominal government expenditure is fixed by Barnett, and since prices do not change in the long-run real government expenditure is also fixed.

Again, as our theoretical analysis implies, Barnett Plus provides an additional stimulus to demand, as is apparent from column 3 of Table 1. It is clear that population is stimulated by the increase in exports under Barnett. However, under Barnett Plus this leads to an increase in government expenditure of 1.88%, and so a greater stimulus to GDP and employment of 2.99% and employment of 2.89% than under Barnett.

Under Smith, qualitative long-run results are identical, but the quantitative results differ (column 4). Here devolved tax revenue per capita does increase – tax revenues increase by more than population – so that government expenditure is higher than under Barnett by 0.31%. GDP and employment effects are therefore greater too, at 2.58% and 2.41% respectively. Given our base model configuration, the net Smith transfer is positive and so provides a greater incentive to grow the host regional economy than does Barnett, at least in the face of a stimulus to exports.

We know that the Treasury adjustment of the block grant under partial tax autonomy is associated with a greater multiplier than is Smith, and this is confirmed by the results reported in column 5 of Table 1. Government expenditure in this case increases by 1.10% compared to 0.31% under Smith, and GDP now increases by 2.79% and employment by 2.65%.

Under full tax autonomy the stimulus to demand is maximised, as inspection of column 6 confirms. Here government expenditure rises by 3.25%, GDP by 3.47% and employment by 3.47%. Since in this case all of the increases in tax revenues are recycled by the Scottish Government, the export multiplier is largest in this case.

The equalisation regime adjusts the Barnett settlement downwards (upwards) in response to any increase (decrease) in Scottish per capita GDP. Under Barnett GDP per capita in Scotland increases by 0.99% in response to the increase in exports and the Scottish budget increases. Accordingly, under the equalisation regime, the results of which are reported in column 7 of Table 1, government expenditure actually falls, albeit by a modest 0.07%, resulting in slightly smaller increases in GDP (of 2.47%) and employment (2.27%) than under Barnett.

Our empirical model provides a clear ranking of the fiscal regimes in terms of the long-run impact of an export stimulus: FTA>BP>T>S>B=RB>E, and an idea of the differences in scale implied by the alternatives. Not surprisingly, the biggest multiplier for an export stimulus is associated with full tax autonomy and the smallest with the equalisation regime. However, we know that only a subset of these rankings are general; others depend on the net effects of countervailing forces.

The impact of stylised fiscal regimes on the response to increases in productivity

The 5% increase in Scottish labour productivity represents a major stimulus to the supply-side of the regional economy, which simultaneously tends to reduce prices and increases output and employment, through the impact of increased competitiveness on exports and, typically also, through an induced significant rise in real public expenditure. As our theoretical analysis implies, across all fiscal regimes: real wages and unemployment rates are unaffected; nominal wages and CPI both fall by 2.24%, and exports rise by 6.47% (Table 2). However, the scale of the response of the real economy to the productivity change varies with the fiscal regime in accordance with our theoretical considerations. The source of these differences is the differences in real government expenditures associated with each regime.

Table 2. The long-run impact of a 5% stimulus to productivity across stylised regional fiscal systems
(% changes from base)

	Real Barnett	Barnett	Barnett Plus	Smith	Treasury	Full Tax Autonomy	Equalisation
GDP (£m)	4.72%	5.80%	6.15%	5.64%	5.76%	5.96%	5.63%
GDP per capita	4.50%	4.74%	4.81%	4.70%	4.73%	4.77%	4.68%
Household Consumption	1.00%	1.68%	1.90%	1.58%	1.66%	1.78%	1.42%
Investment	4.33%	5.14%	5.41%	5.02%	5.11%	5.26%	4.99%
Total Exports	6.47%	6.47%	6.47%	6.47%	6.47%	6.47%	6.47%
Export RUK	6.42%	6.42%	6.42%	6.42%	6.42%	6.42%	6.42%
Export ROW	6.56%	6.57%	6.57%	6.57%	6.57%	6.57%	6.57%
Total Imports	0.06%	0.83%	1.07%	0.71%	0.80%	0.93%	0.65%
WAGES:							
Labour cost WF	-2.24%	-2.24%	-2.24%	-2.24%	-2.24%	-2.24%	-2.24%
Gross HH Wage WHG	-2.24%	-2.24%	-2.24%	-2.24%	-2.24%	-2.24%	-2.24%
Net HH Wage WHN	-2.24%	-2.24%	-2.24%	-2.24%	-2.24%	-2.24%	-2.24%
Real Labour cost WF/CPI	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
Real Gross HH Wage WHG/CPI	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
Real Net HH Wage WHN/CPI	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
PRICES:							
CPI	-2.24%	-2.24%	-2.24%	-2.24%	-2.24%	-2.24%	-2.24%
Govnt price index	-3.99%	-4.00%	-4.00%	-4.00%	-4.00%	-4.00%	-4.00%
Value Added Price index	-4.73%	-4.73%	-4.73%	-4.73%	-4.73%	-4.73%	-4.73%
Export price index	-3.09%	-3.09%	-3.09%	-3.09%	-3.09%	-3.09%	-3.09%
Repl cost of capital(PINV)	-1.78%	-1.78%	-1.78%	-1.78%	-1.78%	-1.78%	-1.78%
User cost of capital(PINV)	-1.78%	-1.78%	-1.78%	-1.78%	-1.78%	-1.78%	-1.78%
LABOUR MARKET:							
Unemployment level	6.00%	6.00%	6.00%	6.00%	6.00%	6.00%	6.00%
Unemployment Rate (pp difference)	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
Employment	0.33%	1.57%	1.97%	1.38%	1.52%	1.74%	1.40%
Labour supply	0.33%	1.57%	1.97%	1.38%	1.52%	1.74%	1.40%
Population	0.21%	1.02%	1.28%	0.90%	0.99%	1.13%	0.91%
NOMINAL TAX REVENUE:							
Indirect tax IBT	1.47%	2.14%	2.35%	2.04%	2.11%	2.23%	1.99%
Total HH Tax	-1.24%	-0.27%	0.03%	-0.42%	-0.31%	-0.14%	-0.52%
Employer NI	-1.92%	-0.71%	-0.32%	-0.89%	-0.75%	-0.54%	-0.88%
Employee NI	-1.92%	-0.71%	-0.32%	-0.89%	-0.75%	-0.54%	-0.88%
Income Tax	-1.92%	-0.71%	-0.32%	-0.89%	-0.75%	-0.54%	-0.88%
Corporation Tax	2.47%	3.27%	3.53%	3.15%	3.24%	3.38%	3.12%
VAT Revenues	-1.67%	-0.42%	-0.02%	-0.61%	-0.47%	-0.25%	-0.67%
Council Tax	1.00%	1.69%	1.90%	1.58%	1.66%	1.78%	1.42%
Smith Taxes / Head	-2.05%	-1.62%	-1.49%	-1.69%	-1.64%	-1.56%	-1.71%
SCOT GOV SPENDING:							
Transfers to HH from Gov	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
Transfers to HH from RUK	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	-0.75%
Scottish Government Expenditure	-3.99%	0.00%	1.28%	-0.61%	-0.15%	0.56%	-0.33%
Real Scottish Government Consumption	0.00%	4.16%	5.50%	3.53%	4.00%	4.75%	3.82%

Under Real Barnett GDP rises by 4.72% and employment by 0.33% in response to the 5% rise in labour productivity. (See column 1 in Table 2.0 In fact this is the smallest stimulus to GDP and employment of any of the fiscal regimes. The reason is clear: in this case real government expenditure is fixed. The price index for Scottish Government expenditure falls by 3.99%, substantially in excess of the fall in the CPI because the former has comparatively little import content (the prices of which are unchanged), and nominal government spending falls by the same amount. In all other cases real government expenditure increases significantly, reinforcing the expansionary impact of the increase in productivity. This case isolates the impact of the export stimulus alone. The key transmission mechanism is the major stimulus to exports, which rise by 6.47% as a consequence of the improvement in competitiveness. Notice that, in our default model, the substitution and competitiveness effect of the labour productivity stimulus exceeds its direct

dampening effect in the long-run and employment rises even in the absence of a reinforcing fiscal expansion.

Given the fall in prices (especially in the deflator for government spending), we know that Barnett results in a larger budget for the Scottish Government than Real Barnett (Table 2, column 2). Here the maintenance of a fixed nominal expenditure causes real government expenditure to increase by 4.16%. This impact is additional to that associated with improved competitiveness (which was isolated under Real Barnett) and so the stimulus to GDP (5.80%) and employment (1.57%) is much greater here. The impact of the stimulus to government expenditure is itself maximised by what is, across long-run equilibria, effectively a passive supply side due to the accumulation of capital and migration flows. There is no crowding out in the long-run under any of the fiscal regimes. While the impact of the productivity stimulus may seem disproportionate in this case, it simply reflects the fact that this fiscal regime acts to reinforce the long-run competitiveness effect with a major fiscal expansion.

Under Barnett Plus (column 3), government expenditure per capita is maintained. Since under Real Barnett population increases (by 0.21%), Barnett Plus is associated with an increase in nominal government expenditure. The real fiscal stimulus is therefore even greater than under Barnett, which further stimulates in-migration, so that population ultimately increases by 1.28%. Real government expenditure ultimately rises by 5.5% - the biggest fiscal stimulus across all the regimes. Of course, any fiscal expansion in these circumstances is very effective since there is no induced crowding out and output and population are further increased: GDP ultimately rises by 6.15% and employment by 1.97%.

It is clear that, under Smith, (Table 2, column 4), both GDP and employment are actually stimulated less than under Barnett (5.64% and 1.38% as against 5.80% and 1.57% respectively). Here the block grant adjustment exceeds the increment to devolved tax revenues: the net Smith Transfer is negative so that the Scottish budget is smaller than it would have been under Barnett. Nominal Government expenditure actually falls by 0.61% and the rise in real government expenditure is 0.63 percentage points less than under Barnett. Here devolved nominal revenues per capita fall and we know that this ensures the multiplier under Barnett is greater than that under Smith. This contrasts with our finding for the export stimulus: the ranking of regional fiscal regime multipliers depends on the source of the disturbance.

Under the Treasury variant of the block grant adjustment (column 5), the results confirm our theoretical analysis, with impacts being greater than under Smith with GDP increasing by 5.76% and employment by 1.52%. As for Smith, there is actually a fall in nominal revenues in this case, although on a lesser scale than for Smith, so that the increase in real government expenditure is greater than Smith. However, the fiscal stimulus is, in this case, less than under Barnett.

The results of the full tax autonomy case are reported as column 6 of Table 2. As our earlier analysis implied as compared to the Treasury model there is a greater stimulus to real government expenditure of 4.75%, and a correspondingly bigger impact on GDP (5.96%) and employment (1.74%). Perhaps the most surprising feature here is that this case does not generate the greatest supply-side multiplier: here the Barnett Plus regime does that because of the effect of falling prices on the fixed nominal budget generates the greatest fiscal expansion.

As our earlier analysis suggests, the equalisation case moderates the scale of the region's response to the productivity stimulus, generating a rise in real government expenditure of 3.82%, and a stimulus to GDP and employment of 5.63% and 1.40% respectively (as compared to 5.80% and 1.57% respectively). The region's multiplier is, as expected, diminished by the operation of an equalisation mechanism.

Figures 4 and 5 summarise the changes in the Scottish Government budget in response to the export and productivity shock respectively, for the stylised fiscal regimes considered here. Only the regime that fixes government expenditure in real terms identifies the "pure" effect of the supply and demand disturbances. All of the other regimes, except the equalisation case, generate a real fiscal expansion that reinforces the impact of the disturbance, although this varies across regimes in a manner consistent with our theoretical analysis. The equalisation case moderates the regional multiplier associated with whatever regime it is attached to (here the Barnett case).

Figure 4. Changes in nominal government expenditures in response to the export stimulus.

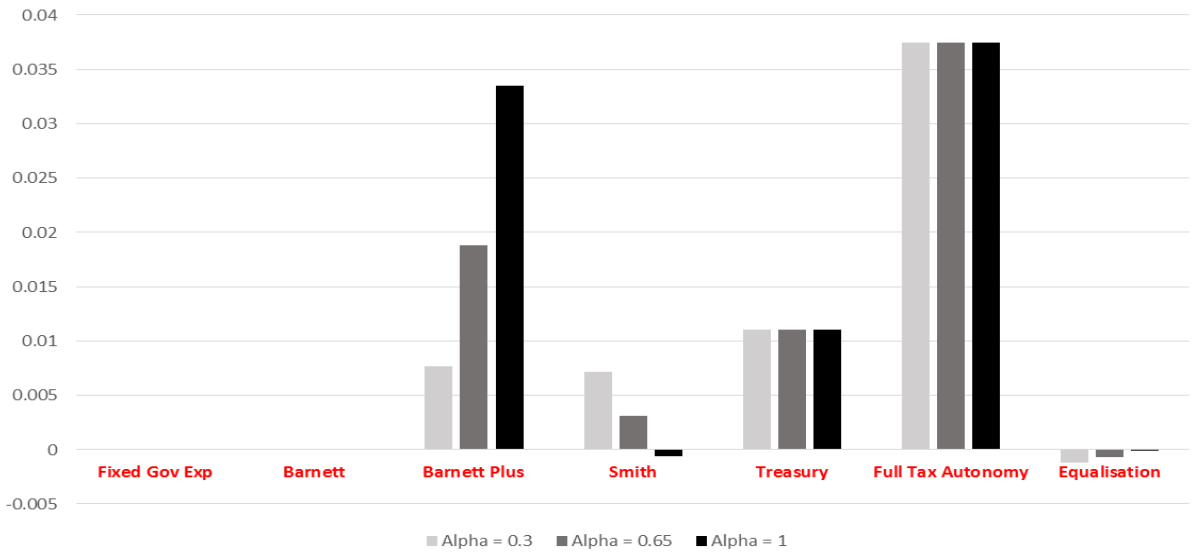
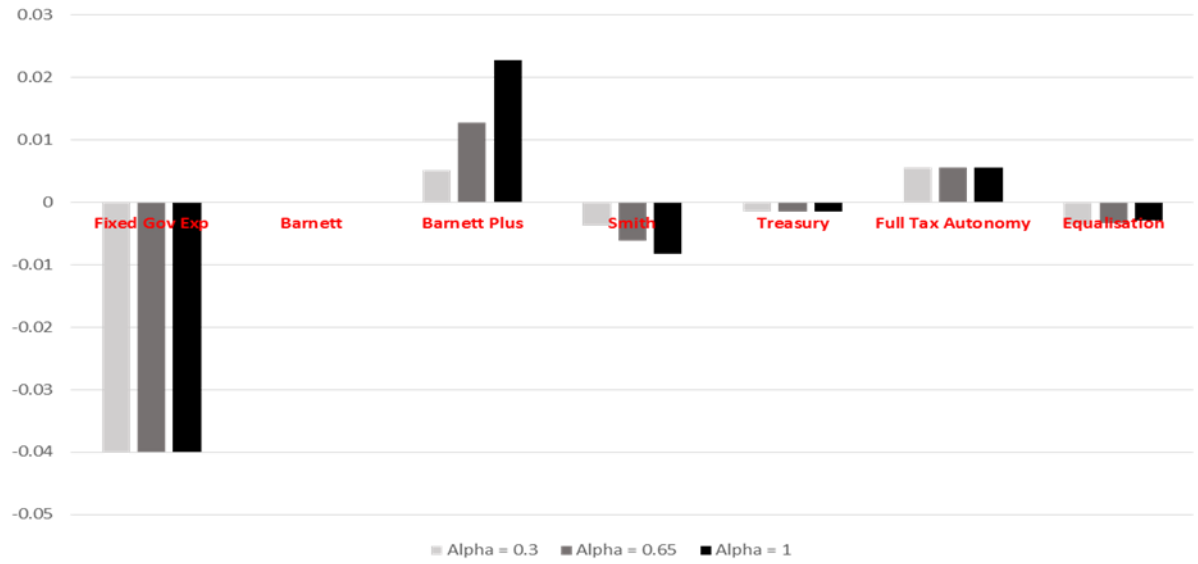


Figure 5. Changes in nominal government expenditures in response to the 5% stimulus to productivity.



6. Conclusions

There has been a widespread move towards more decentralised regional fiscal systems, but there has been little research on how regional fiscal systems themselves may impact upon economic outcomes. This is the concern of the current paper. We explore the impact of a range of stylised regional fiscal systems on the response of regional economies to demand and supply side disturbances, and therefore on the incentive to the regional government to pursue growth promoting policies. These encompass a number of variants of block grant, tax-sharing systems, full tax autonomy and fiscal regimes that incorporate an equalisation mechanism.

Our partial and general equilibrium theoretical analyses provide insights into the likely impact of these stylised fiscal regimes on the responsiveness of regional economies to exogenous export and productivity changes, including that tax-sharing regimes may not always improve growth incentives relative to block grants. However, useful though this analysis is it proves incapable of delivering a comprehensive ranking across all the fiscal systems we consider. Furthermore, the theoretical analysis provides no indication of the likely scale of the responsiveness of regional economies to alternative fiscal systems, yet this information is likely to be crucial where further regional fiscal decentralisation is being considered.

Accordingly, we explore the issue further using an intertemporal, regional empirical general equilibrium model, extended to incorporate a range of alternative fiscal regimes. This allows us to isolate the impact of these regimes on the responsiveness of regional economies within a framework that uses a common database and model. The results offer confirmation of our theoretical analysis, but also generate some surprising results, including that variants of block grant systems could possibly provide greater incentives to supply-driven growth than full tax autonomy. This in part reflects a rather neglected aspect of decentralised fiscal systems, namely that they are often framed in nominal terms, without reference to regional prices.

In future research we plan to extend our analysis in a number of directions. First, the impact of the link between population and labour supply could be further explored. One dimension of this is the behaviour of migration, and alternative perspectives on this process are likely to prove significant,

potentially for our theoretical as well as empirical analysis. Second, since the shorter-run is often of policy interest it would be useful to compare a selection of the adjustment paths to the long-run equilibria which are the focus of this paper. Third, while there is compelling evidence supporting the adoption of a regional wage curve over the longer term, experience since the Great Recession suggests that this might have broken down, and it would be interesting to analyse the consequences of this. Fourth, the analysis could be extended to consider impacts on the distribution of income within the target region. Finally, the single region focus considerably simplifies the analysis, but an extension to a multi-regional context would undoubtedly prove valuable.

References

Blanchflower, G.D., Oswald, A.J., 1994. Estimating a Wage Curve for Britain. *The Economic Journal*, 104, 1025-43.

Bell, D., Eiser, D., and Phillips, D. (2016) Scotland's Fiscal Framework: Assessing the Agreement, Institute of Fiscal Studies, IFS Working Paper, W16/05

Eiser, David (2017) A primer on the Scottish Parliament's new fiscal powers : what are they, how will they work, and what are the challenges? Fraser of Allander Economic Commentary, 41 (2). pp. 26-41.

Christie, A. & Swales, J.K. (2010) The Barnett Allocation Mechanism: Formula Plus Influence?, Regional Studies, 44:6, 761-775, DOI: 10.1080/00343400903107710

Devarajan, S. & Go, D., 1998. The Simplest Dynamic General Equilibrium model of an Open Economy. *Journal of Policy Modeling*, 20, 677–714.

Emonts-Holley T., Greig A., Lecca P., Lisenkova K., McGregor P.G. & Swales J.K. (2019) A Scandinavian 'high-tax, high-spend' model for regions? The impact of enhanced regional fiscal autonomy, Spatial Economic Analysis, DOI: 10.1080/17421772.2019.1568536

Hayashi, F., 1982. Tobin's Marginal q and Average q: A neoclassical Interpretation. *Econometrica*, 50, 213-224.

Go, D., 1994. External Shocks, adjustment policies and Investment in a developing economy: Illustration from a forward-looking CGE model of the Philippines. Journal of Development Economics, 44, 229-261.

HM Government (2015) Scotland and the United Kingdom: an enduring settlement, *Presented to Parliament by the Secretary of State for Scotland by Command of Her Majesty*, January 2015

Institute of Fiscal Studies (2014) "Flaw" in Barnett Formula protects Scotland and Northern Ireland from hundreds of millions of cuts, Press Release

Layard, R., Nickell, S. Jackman, R., 1991. *Unemployment: Macroeconomic Performance and the Labour Market*. Oxford University Press, Oxford.

Lecca, P, McGregor, P & Swales, J 2013, Forward looking versus myopic regional computable general equilibrium models: how significant is the distinction?, *Economic modelling*, 31, 160-176.

Lecca, O., McGregor, P.G. & Swales, J.K. & Yin, Y.P 2014, Balanced budget multipliers for small open regions within a federal system: Evidence from the Scottish variable rate of income tax. *Journal of Regional Science,* 54(3), 402-421.

Keep, M. (2018) The Barnett Formula, *House of Commons Library*, Briefing Paper

Martinez-Vazquez, J., McNab, R.M. (2003): Fiscal Decentralization and Economic Growth, *World Development* 31(9), pp. 1597–1616.

McGregor, P. G., Swales, J. K. and Y P Yin "A Long-Run Interpretation of Regional Input-Output Analyses", *Journal of Regional Science*, 1996, Vol. 36, No 3, 479-501.

Oates, W.E. (2005) Toward A Second-Generation Theory of Fiscal Federalism, *International Tax and Public Finance* 12(4), pp. 349-373.

OECD (2013), *Fiscal Federalism 2014: Making Decentralisation Work*, OECD Publishing, Paris, https://doi.org/10.1787/9789264204577-en.

Ouimet, H.R. (2014) Quebec and Fiscal Federalism: From Tremblay to Séguin and Beyond, *Canadian Journal of Political Science* 47(1), pp. 47-69

Poghosyan, T., Senhadji, A. and Cottareli, C. (2016) "The role of fiscal transfers in smoothing regional shocks: evidence from existing federations", IMF Working Paper, WP/16/141.

Rodden, J. And Wibbels, E. (2010), Fiscal decentralization and the business cycle: an empirical study of seven federations. Economics & Politics, 22: 37-67. doi:10.1111/j.1468-0343.2009.00350.x

Rodríguez-Pose, A., & Gill, N. (2004). Is There a Global Link between Regional Disparities and Devolution? *Environment and Planning A: Economy and Space*, *36*(12), 2097–2117. https://doi.org/10.1068/a362

Vandernoot, J. (2014) Funding of German Lander, Mechanisms, and Solidarity, *Research in Applied Economics* 6(2), pp. 1-17.

Treyz, G.I., Rickman, D.S., Hunt, G.L., Greenwood, M.J., 1993. The dynamics of US internal migration. *Review of Economics and Statistics*, 75, 209-14.