Regional Economics and Constitutional Change in the UK

Submitted for the Degree of PhD in Economics

University of Stirling

July 2016

David Eiser

Registration Number: 2031086

Acknowledgments

I would not have been able to complete this PhD without the significant support of three people in particular.

Professor Rob Ball encouraged me to register for a PhD shortly after I started working at the University in 2011. He provided useful help and advice in support of my application, and during the first year of my PhD. He also secured from the Head of School a commitment to fund my registration fees during the first two years, utilising the overheads from Rob's ESRC grant.

Rob was thus the catalyst for this PhD. His sad passing in 2012 means he was not able to see the result.

Professor David King became my supervisor following Rob's death. David has been very supportive throughout my time at Stirling. His support and advice in preparing the first paper in this PhD – not to mention three other published papers that we worked on together – was invaluable. David also facilitated my connection with David Bell, and the transfer of my registration from Management to Economics.

Professor David Bell has been my supervisor since 2013. He has imparted a wealth of knowledge on a huge range of topics, from the technicalities of using Stata to analyse the Labour Force Survey, to the latest developments in Scottish fiscal policy. David has also done a great deal to further my academic career more generally. As well as jointly publishing a number of papers, reports, books and presentations, David has regularly opened up all manner of opportunities for me to contribute directly to policy and media debates, significantly raising my professional profile.

I am hugely grateful to Rob, David King and David Bell for all their help and support.

I would also like to thank Dr. David Comerford for his help and advice on some of the more mundane, but no less important, aspects of a PhD.

Abstract

The UK, traditionally one of the more fiscally centralised of OECD countries, is currently in the midst of an extensive programme of tax decentralisation. This is most evident in Scotland. Ten years ago the Scottish Government was almost wholly reliant on a block grant from the UK Government to fund its spending, and debate was focussed on how the determination of this grant should be reformed.

Today the Scottish Government has far greater fiscal autonomy. Income tax was almost fully devolved to the Scottish Parliament in April 2017, and around half of VAT revenues will be assigned to Scotland by 2020. As a result, the devolved Scottish budget will in future be linked much more closely to Scotland's economy, and Scottish politicians will be able to deviate from UK policy on the setting of income tax and various smaller taxes.

The objective of this PhD is to examine the economic and political motivations for and implications of greater fiscal decentralisation, with a particular focus on the Scottish case. Its key over-arching questions include:

- Which fiscal powers are more and less suitable for decentralisation, and what might constraints might a devolved government face in exercising devolved tax powers?
- To what extent are the objectives of fiscal decentralisation compatible with the goal of inter-regional equity in public good provision?
- To what extent is fiscal decentralisation likely to enhance the incentives faced by politicians in a devolved parliament to pursue particular types of policy? And to what extent does the answer to this question depend upon the way in which supporting fiscal institutions, notably including the design of block grant arrangements, influence this?
- What factors determine regional economic performance, and to what extent can devolved governments be held accountable for (or face the budgetary consequences of) those trends?
- To what extent might fiscal decentralisation assuage or accentuate demands for Scottish independence?

This PhD consists of four academic papers covering aspects of regional economics and constitutional change in the UK, with a particular focus on Scotland. Each of the four papers is preceded by an abstract. An introductory chapter provides theoretical and policy context within which the four papers are situated. A concluding section to the PhD is provided in Chapter 6.

The four papers cover the following topics:

- Paper 1 (Chapter 2) was published in the immediate aftermath of the Scottish independence referendum of 2014, and considers the issues and constraints involved in devolving further fiscal powers to the Scottish Parliament.

- Paper 2 (Chapter 3) considers the scope for replacing the Barnett Formula (used to allocate funding to the Scottish Government) with a form of spending-needs assessment, based on a comparative analysis of formulae used within England and Scotland to allocate health funding to territorial health boards.

- Paper 3 (Chapter 4) examines how regional labour markets in the UK responded to the 2008/9 recession and its aftermath, and considers which factors may have influenced regional resilience to the recession.

- Paper 4 (Chapter 5) examines the factors that determine differential growth in regional income tax revenues, and considers the extent to which it is reasonable to hold devolved governments wholly to account for differential economic performance.

- Chapter 6 concludes.

Contents

1. Introduction

i. Introduction

The UK, traditionally one of the more fiscally centralised of OECD countries, is currently in the midst of an extensive programme of tax decentralisation. This is most evident in Scotland – income tax, already partially devolved to the Scottish Parliament, will be almost fully devolved by 2017/18 and half of VAT revenues will be assigned to Scotland by 2020 (land and property taxes are already fully devolved to the Scottish Parliament). By 2021, the Scottish Government will be responsible for raising around half of its budget through devolved and assigned taxation, compared to less than 10% a few years ago.

Tax decentralisation to Scotland is partly the result of long-term objective to narrow the Scottish Parliament's 'vertical fiscal imbalance' – the gap between devolved expenditure and revenue raising responsibilities – but it has been accelerated by the Scottish independence referendum of 2014, during which time the Scottish electorate's desire for the Scottish Parliament to have greater fiscal autonomy became clear. But tax decentralisation is not unique to Scotland. Income tax is also being partially devolved to the Welsh Assembly Government, whilst Corporation Tax is being devolved to the Northern Ireland Assembly. Taxes are also being devolved to lower levels of government. In England, Business Rates, which have traditionally been pooled and shared at national level, will be fully devolved to local government by 2020.

Where taxes are devolved to a subnational government, so that the subnational government has some ability to vary tax rates or thresholds, the rationale is in part to provide the subnational government with the autonomy to vary its budget at the margin, and more generally to exercise some form of fiscal policy distinctiveness. But in all cases, i.e. whether a tax is devolved or simply assigned, the aim is to make local politicians more financially accountable by linking budgets to local economic performance. In the rhetoric, tax devolution or assignment enables devolved governments to benefit from policy decisions that increase revenues, and conversely to bear the costs of policy that reduces revenues.

This is particularly true in the Scottish case. The Smith Commission, established to agree which fiscal powers to devolve to the Scottish Parliament following the no-vote at the 2014 independence referendum, argued that tax decentralisation would make the Scottish Parliament 'more accountable and responsible for the effects of its policy decisions and their resulting benefits or costs' and that it would 'strengthen the Scottish Parliament's ability to pursue its own vision, goals and objectives' (Smith Commission, 2014).

Fiscal federalism theory does indeed recognise a number of potential advantages of fiscal devolution in terms of the fiscal incentives faced by politicians in devolved institutions and the accountability of those politicians for the budgetary decisions they make, and I review these arguments in this introduction. However, the literature recognises too that tax devolution itself is only part of story. The design of supporting fiscal institutions – how grants to devolved governments are allocated, and the nature of any borrowing and cash management powers – is critical in determining the set of fiscal incentives a devolved government faces, the fiscal risks it is exposed to, and the way in which these incentives, risks and choices are perceived and understood by the electorate.

In Scotland, whilst agreement on which taxes should be devolved happened rapidly following the 2014 independence referendum, the design of the supporting Fiscal Framework – which covers arrangements for the Scottish block grant and cash management powers – took much longer to agree. Disputes centred around the extent to which the devolved budget should be exposed to particular fiscal risks, and the tools SG should have access to in order to manage risks.

The objective of this PhD is to examine the economic and political motivations for and implications of greater fiscal decentralisation, with a particular focus on the Scottish case. Its key over-arching questions include:

- Which fiscal powers are more and less suitable for decentralisation, and what might constraints might a devolved government face in exercising devolved tax powers?
- To what extent are the objectives of fiscal decentralisation compatible with the goal of inter-regional equity in public good provision?
- To what extent is fiscal decentralisation likely to enhance the incentives faced by politicians in a devolved parliament to pursue particular types of policy? And to what extent does the answer to this question depend upon the way in which supporting fiscal institutions, notably including the design of block grant arrangements, influence this?
- What factors determine regional economic performance, and to what extent can devolved governments be held accountable for (or face the budgetary consequences of) those trends?
- To what extent might fiscal decentralisation assuage or accentuate demands for Scottish independence?

The papers in the PhD have to an extent responded to contemporaneous issues, and these issues have evolved rapidly and significantly since the commencement of the PhD in 2011. The first paper provides additional policy and economic context to recent changes to Scotland's devolved fiscal settlement. The second paper considers the scope for replacing the Barnett Formula – used to

allocate block grant to the Scottish Government – with a needs assessment based funding formula, using healthcare as an application.

The third and fourth papers both consider questions of differential regional economic performance, and the extent to which a devolved government should be held accountable, through its devolved budget, for this differential performance. Paper 3 looks at the question of differential regional performance from a UK-wide perspective, considering economic trends in UK regions since the recession of 2008/9. Paper 4 considers this question with a specific focus on the devolution of income tax to Scotland from April 2017. It asks what factors are likely to determine regional income tax revenue growth, using the period from 1999/00 – 2015/6 to shed empirical light on this question.

In the remainder of this introduction I sketch out some of the key theoretical and policy issues that have informed the papers in this PhD. The second section reviews the theoretical literature on the motivations for tax decentralisation, and the design of fiscal federal institutions. The third section reviews recent and ongoing changes to devolved taxation arrangements in Scotland, as a precursor to the issues that are discussed in more detail in Chapter 2. The fourth section turns attention to the political economy of Scotland's Barnett Formula determined block grant, asking why the calculation of Scotland's block grant has assumed totemic importance in any question of fiscal devolution, and providing context to the analysis in Chapter 3 which considers the scope for reform of the block grant arrangements. The fifth section reviews Scotland's new Fiscal Framework, focussing in particular on the arrangements for grant allocation, and describing the implications for the budgetary risks that the devolved Scottish budget is now exposed to. This discussion forms the motivating basis (indirectly) for Chapter 4 and (directly) for Chapter 5.

ii. Fiscal decentralisation and the design of fiscal incentives

Since Musgrave's seminal work (Musgrave, 1959), economists have devoted substantial effort to the study of fiscal federalism. This literature considers the rationale for fiscal decentralisation, which fiscal functions are most and least appropriate for decentralisation, and how the design of supporting fiscal institutions (including the allocation of grant to, and the design of fiscal rules for, subnational governments) can help maximise the benefits from decentralisation. This section considers what this literature says about the choices and trade-offs involved in issues of fiscal decentralisation; subsequent sections consider Scotland's new devolved fiscal settlement in that context.

The so-called first generation of fiscal federalism theory hypothesises that the benefits of decentralization are positively correlated with the (geographic) variance in preferences for publicly

provided goods (Musgrave 1959, Oates 1972). Decentralisation allows local governments to tailor outputs of local public goods in such a way as to meet the preferences of their respective jurisdictions – providing a higher level of welfare than if a central government provided a single uniform level of public output across all jurisdictions. The welfare gains from decentralisation are enhanced where households are mobile across local areas – allowing households to relocate to the jurisdiction which provides the bundle of local public goods most closely tailored to their preferences.

A potential disadvantage of fiscal decentralisation is that, where local public goods create spillovers (i.e. benefits to residents of other jurisdictions), it is likely that local governments will underprovide those public outputs. Oates (2005) thus summarises the decentralisation trade-off as: 'on the one hand, inefficiencies under centralised provision of public services stemming from uniform outputs that fail to reflect divergences in local tastes and conditions versus, on the other hand, inefficiencies in local provision resulting from the failure to internalise interjurisdictional externalities'.

The second generation fiscal federalism literature examines the workings of different political and fiscal institutions in a setting of imperfect information and control, with a focus on the incentives that these institutions embody and the resulting behaviour they induce (Oates, 2005). Whilst the second generation theory, like the first, identifies that the debate about the gains from decentralisation continues to hinge on the relationship between preference heterogeneity and spillovers (Besley and Coate, 2003), an additional key insight from the second generation theory is that decentralisation may be advantageous even where there is perfect homogeneity of preferences across jurisdictions. The case for decentralisation rests not only on differences in preferences, but also on the potential for enhanced local control and 'accountability' under decentralised provision, where accountability is taken to mean the extent to which political agents can be associated with outcomes (Tommassi and Weinschelbaum 2007; Seabright 1996). Local voters are more likely to trust political agents if they are clear that their loyalties are not divided between serving some external interests in addition to the local interest.

A further strand of literature considers the rationale for decentralisation where the underlying structure of national and sub-national government is not necessarily stable. This literature considers the factors that might cause one or more component parts of the nation to secede, and what this might imply for fiscal federal arrangements.

Alesina and Spolaore (1997) explore the tradeoff between the benefits of large jurisdictions and the costs of heterogeneity. The benefits of size derive from scale economies in taxation and public good provision, and internal free trade. But the cost of size is the difficulty of satisfying an increasingly

heterogenous population with heterogenous preferences. In this context, the granting of fiscal autonomy to a peripheral region by the centre may help assuage demands for independence.

This framework allows consideration of the effects of globalisation on the seccession threat (and hence indirectly on pressure for decentralisation). Alesina and Spolaore (1997) argue that, by lowering trade barriers, globalisation will reduce the costs of secession, thus leading to an increase in fiscal decentralisation in order to assuage demands for independence. Garrett and Rodden (2005) however argue that globalisation is likely to lead to pressure for centralisation of taxation functions (and to a lesser extent, spending too). This is on the basis that globalisation is associated with greater fiscal risk and volatility, and will thus increase the appeal of a strong risk pooling and sharing function of centralised government to insure against asymmetric economic shocks. They argue that the less closely correlated are regional business cycles, the stronger the case for inter-regional risk-sharing, a point we return to subsequently in thinking about the Scottish case.

So much for the motivations for fiscal decentralisation, what does the literature say about which fiscal functions should be decentralised?

The first generation theory of fiscal federalism relates very much to the provision of local public goods. Central government plays the key role in redistribution and macroeconomic stability, given the constraints faced by local governments in exercising these functions (namely the open nature of local economies and scope for factor mobility).

What about decentralisation of taxation? First generation fiscal federalism makes the point that distortions can result from the taxation of highly mobile tax bases, and is best made by the central government (or possibly higher level regional rather than local governments). Local governments should rely primarily on 'benefit taxes', such as property taxes, that can be linked to the provision of local services. However, in the same way that decentralisation of local public good provision can be justified where preferences vary geographically, a case can also be made for the decentralisation of redistributive tax policy where preferences for this vary across parts of a nation state (Bolton and Roland, 1997).

In general, it is easier to identify spending functions that are appropriate for local control than it is to identify taxation functions that are appropriate for local control, emphasising Bahl and Cyan's (2011) point that, if the theory of tax assignment was followed to the letter 'there would surely be no subnational government taxes'. But the second generation fiscal federalism theory in particular places a heavy emphasis on the idea that decentralised budgets should be determined as much as possible by own source revenues. Where decentralisation of expenditures is financed primarily by

transfers from above, this can effectively act as a 'soft budget constraint', undermining the accountability of subnational politicians, and providing incentives to exploit the fiscal commons.

A trade-off thus emerges between, on the one hand, a desire to assign spending and taxation levers to appropriate levels of government, and on the other hand, the objective that sub-national governments should fund themselves through own sources of revenue as far as possible.

The design of intergovernmental grant mechanisms is a critical part of any decentralised fiscal settlement. Central government grant helps to close any vertical fiscal imbalance that subnational governments inevitably face as a result of the mismatch between advantages of spending over taxation decentralisation. But grants play an important role in relation to horizontal equity, redistributing between poorer and richer subnational governments, and in providing an insurance function to subnational governments against the risk of economic shocks, whether these shocks are common to the nation, or asymmetric across regions.

However the provision of these redistributive and insurance functions may create moral hazard problems, disincentivising subnational authorities from providing outputs or programmes which improve the economy's capacity to grow or respond to exogenous shocks. The challenge is thus to design fiscal institutions capable of realising the benefits of decentralisation while avoiding the potentially negative effects of soft-budget constraints.

In practice the literature says relatively little about how this challenge might be realised, recognising that it in large part reflects a subjective political choice. There is support however for the notion that, for the benefits of decentralisation to be realised, subnational governments should face incentives to adopt growth-enhancing policies. Weingast (2005) refers to this as 'market-preserving federalism', and argues that market-preserving federalism can be consistent with an objective to provide a risk-insurance or redistributive function from the centre. Institutions that harden local budget constraints can also include tight rules on subnational deficits, constraints on subnational borrowing, and clear specification as to how fiscal crises will be handled (Oates, 2005).

The interrelated literatures on fiscal federalism and secession thus provide a range of motivations for fiscal decentralisation. Fiscal decentralisation is recognised to have a number of potential merits, but only if the broader fiscal institutions have been appropriately designed. Rather than necessarily providing specific answers however, the literature serves largely to highlight the many trade-offs that exist in determining options for fiscal decentralisation.

iii. The evolution of Scotland's fiscal autonomy

When the Scottish Parliament was established it was characterised by a high level of vertical fiscal imbalance. On spending, the parliament has substantial responsibilities in relation to health, education, justice and policing, economic development, the environment, and culture and recreation. On tax however, only two relatively small property taxes were determined in Scotland – Council Tax (a tax on domestic property) and Non-Domestic Rates, a tax on business property.

Revenues from these two taxes amounted to around 10 per cent of the Scottish Parliament's spending budget (Scottish Government, 2016), with the remainder of the budget provided by a grant from the UK Government.

The influential Calman Commission report, published in 2009, argued that this vertical fiscal imbalance was problematic, arguing that 'Funding by block grant alone means that while the Scottish Parliament is completely accountable for the spending of its budget, it is not accountable for the total of that budget or how it is raised; it has no fiscal powers that can be used as policy instruments and it does not have a direct financial stake in the performance of the Scottish economy' (Commission on Scottish Devolution, 2009, para 3.87).

The Calman Commission recommended that this VFI be addressed through the partial devolution of income tax to the Scottish Parliament, alongside devolution of stamp duty land tax (a tax on property transactions) and landfill tax (a tax on waste to landfill). These recommendations were passed into legislation through the Scotland Act 2012 and came into operation in April 2014.

Meanwhile, the election to the Scottish Parliament of a majority Scottish National Party (SNP) administration in 2011 paved the way for a referendum on Scottish independence to take place in September 2014.

In the final days of the referendum campaign, the leaders of the three main UK unionist parties made a high profile 'vow' to devolve substantial further powers to the Scottish Parliament following the referendum, if Scotland voted to remain in the UK. (This 'vow' was made in response to a tightening in the polls and, indeed, one poll which suggested that a 'yes' vote was likely, 10 days before the referendum).

The first paper in this PhD (Chapter 2) asks which taxes are more or less suitable for devolution, based on fiscal federalism theory and experience in other countries. It considers the economic and constitutional constraints that may influence how effectively the Scottish Parliament will be able to exercise newly devolved tax powers, and it also asks whether welfare spending devolution is consistent with notions of Union.

Following the 2014 Scottish Referendum, the Smith Commission was established to agree which fiscal powers to devolve to the Scottish Parliament. The Smith Commission argued that tax decentralisation would make the Scottish Parliament 'more accountable and responsible for the effects of its policy decisions and their resulting benefits or costs' and that it would 'strengthen the Scottish Parliament's ability to pursue its own vision, goals and objectives'.

The Smith Commission recommended:

- That 'Non Savings, Non-Dividend' (NSND) income tax revenues should be transferred to the Scottish Parliament. NSND income tax revenues account for around 92% of all income tax revenues raised in Scotland (and include the tax paid on income from earnings, self-employment, pension income and property income). The Commission recommended that the Scottish Parliament be able to vary income tax rates and bands in Scotland without constraint. But the UK Government will retain authority to determine the income tax base. This means that the setting of the Personal Allowance, and the way in which the pensions tax relief is defined for example, are determined by the UK Government.
- That Air Passenger Duty (APD) should be devolved in full.
- It also recommended that a share of VAT collected in Scotland should be assigned to the Scottish Parliament. Specifically, the first ten pence of Standard Rate VAT and the first 2.5 pence of reduced rate VAT to be assigned to the Scottish Parliament. Given that the Standard and Reduced rates of VAT are currently 20% and 5% respectively, this means that under current policy, half the VAT revenues raised in Scotland will be assigned to the Scottish budget.
- It also recommended the devolution of Aggregates Levy, in full.

The Smith Commission's recommendations were enacted through the Scotland Act 2016. Table 1.1 summarises the extent of existing and planned tax devolution to the Scottish Parliament.

Table 1.1: Devolved, shared and assigned tax revenues in Scotland

Tax	Date of transfer/ devolution	Revenues raised 2015/16 (£m)	Degree of control by Scottish Parliament	Responsibility for collection
Council Tax	1999	£2,100	Fully devolved; complete autonomy.	Local government
Non-Domestic Rates	1999	£1,900	Fully devolved; complete autonomy.	Local government
Land and Business Transactions Tax (LBTT)	2015	£416	Fully devolved; complete autonomy.	Revenue Scotland
Landfill Tax	2015	£147	Fully devolved; complete autonomy.	Revenue Scotland
Income tax	2017	£11,214	The Scottish Government can set the rates and bands. But the UK Government defines the tax base and sets allowances.	HMRC
Air Passenger Duty	2018	£275	Fully devolved; complete autonomy	Revenue Scotland
VAT	2019	£5,000	Assigned revenues; no autonomy	HMRC
Aggregates Levy	tbc	£53	Fully devolved; complete autonomy	Revenue Scotland

Source: Government Expenditure and Revenue Scotland (GERS); author analysis

A new Scottish tax agency, Revenue Scotland, has been established to collect revenues for the fully devolved Scottish taxes (LBTT, Scottish Landfill Tax, Aggregates Levy, and Air Passenger Duty, which the Scottish Government has announced will be renamed 'Air Departure Tax'). Revenues from the partially devolved income tax and the assigned VAT in Scotland will continue to be collected by HMRC.

In addition to tax devolution, some devolution of social security benefits is also taking place. A number of UK-administered benefits, mainly related to ill-health, disability and caring will be devolved to the Scottish Parliament in around 2019/20. Spending on these benefits in Scotland by the UK Government in 2015/16 totalled around £3bn (Scottish Government, 2016).

By 2019/20, devolved revenues (over which the Scottish Government can vary rates and bands) will raise around 37% of Scottish Government spending. Devolved and assigned revenues (i.e. including assigned VAT) will account for 50% of Scottish Government spending.

How might these recent changes to the devolved Scottish fiscal settlement be seen through the lens of fiscal federalism theory reviewed earlier? The first round of tax decentralisation, operationalised through the Scotland Act 2012 following the Calman Commission report, was motivated very much

by the objective to narrow the Scottish Parliament's VFI and thus improve the accountability of Scottish politicians for budgetary decisions. The second round of tax decentralisation following the 2014 independence referendum was partly motivated by these same objectives, but arguably also by a desire on the part of the UK Government to assuage secessionary threats.

In neither case however has there been a particular sense that decentralisation can be motivated by the notion of Scots having different fiscal preferences. Indeed, Bell and Eiser (2015) argue that there is limited evidence for preference heterogeneity in public good provision in Scotland relative to the rest of the UK (rUK), but stronger evidence that demands for decentralisation are motivated by the Scottish electorate's higher trust in devolved relative to reserved legislatures, and the desire for greater 'local' accountability.

The decision to devolve income tax might be seen as surprising given its key redistributive role and the mobility of the income tax base across UK regions. But this decision must be seen in the context of the fact that few taxes meet all the criteria for devolution, and income tax does have the advantages of being visible to the electorate and a reasonably stable source of revenue. The social security levers being devolved are closely linked to existing devolved functions in health and social care, and do not include benefits that are either very cyclically sensitive or that are particularly effective for redistributive purposes.

But will these new fiscal levers enhance the fiscal incentives faced by Scottish politicians? Will the ability to vary the budget at the margin and influence redistributive policy enhance the accountability of Scottish politicians to the electorate? And will this result in different policy choices in Scotland than might otherwise have been made?

To answer these questions it is necessary to consider the design implications of Scotland's new Fiscal Framework (HM Treasury, 2016). The Fiscal Framework sets out the supporting fiscal institutions required to operationalise the new fiscal powers, with the key aspect of this being the way in which block grant is allocated to the Scottish Parliament following the recent rounds of tax decentralisation. Whilst the decision about which taxes to devolve to the Scottish Parliament was resolved within a matter of weeks following the 2014 independence referendum, it took around 15 months of negotiations before the UK and Scottish Governments were able to agree the design of the Fiscal Framework.

iv. The political economy of the block grant

The Scottish Parliament's high vertical fiscal imbalance means that it has been reliant on a block grant from Westminster to finance the majority of its spending. This block grant is determined by the Barnett Formula.

The Barnett Formula is ostensibly a population based formula with no element of equalisation for spending need or revenue capacity. The Formula calculates the change in the Scottish block grant from one year to the next. Specifically, the Barnett formula allocates to the Scottish Government its population share of changes in 'comparable' spending by the UK Government. The change is added to the Scottish block grant of the previous year.

For example, if the UK Government announces a £100m increase in health spending in England, then because the Scottish population is about 10% of the English population, the Scottish budget will be increased by £10m (although the Scottish Government is not obligated to spend this money on health).

Despite being a population-based formula with no explicit equalisation element, the Barnett Formula is recognised as providing a relatively generous budgetary allocation to Scotland. It enables the Scottish Government to support spending of around 17% higher per person than on equivalent public services in England (Deaner and Phillips, 2014). The Barnett Formula is also used to allocate grant to the Welsh Assembly, although it is perceived as being relatively less generous to Wales in the context of Wales' likely spending needs (McLean and McMillan, 2005; MacKay and Williams, 2005).

The Barnett Formula has long been criticised for what has been perceived as inequity in the way it distributes resources to the UK's devolved administrations, and there have been repeated calls to replace it with a system of spending needs assessment. Despite the criticism of the Barnett Formula, one of the objections to replacing it with a system of needs assessment is that, because what constitutes spending need is inherently subjective and contestable, the UK and devolved governments would never be able to agree on what a needs based grant formula should look like.

The second paper in this PhD (Chapter 3) examines the hypothesis that the UK and Scottish Governments would be unable to agree on how to assess spending need, with an application to health spending. Specifically, the paper exploits the fact that both the UK and Scottish Governments have developed detailed funding formulae for allocating resources to territorial health boards within their respective territories. The paper takes the formula used by the UK Government to allocate NHS resources to Primary Care Trusts in England, and applies it to Health Boards in Scotland. Effectively, the paper asks what level of resource those Health Boards would receive if the English formulae was

17

applied on a UK-wide basis. Then, the formula used by the Scottish Government to allocate NHS funding to Scottish Health Boards is applied to PCTs in England.

The rationale for this analysis is to compare how the English and Scottish formulae assess the relative spending needs of regions with varying socio-economic and health characteristics. Arguably, the more similar the allocations are, the greater is the likely scope for acceptable compromise on the development of a spending needs replacement for the Barnett formula. This is not to say that a Barnett replacement should use either the English or Scottish formulae for health, but rather to identify where disagreements around needs assessment are likely to arise, and to help inform the legitimate starting points for the respective governments around what might constitute a reasonable suite of needs indicators, and where the limits of a 'fair' block grant settlement may lie.

The Barnett Formula's relative generosity to Scotland can be explained by various mechanical features of the formula, but it is the influence of political economy factors that explain why the generosity of the formula to Scotland has been allowed to continue unchecked.

One of the mechanical explanations of the formula's generosity relates to the way that relative population change is treated by the Barnett formula. If population growth is the same in Scotland as in England, then the Barnett Formula should lead to convergence in spending per capita in Scotland with equivalent English spending per capita. This is because the Barnett Formula allocates the same cash terms increase in spending per person to Scotland as in England; over time therefore, any initial difference in spending per head should be eroded over time (and this convergence in spending per head will be faster the faster that comparable spending increases). However, the rate of convergence can slow or even go into reverse if there are large differences in relative population growth. To see this, imagine that the population of England increases, and the UK Government increases spending on health in England to maintain spending per person at a constant level. In this case, the Barnett formula allocates to the Scottish block grant a population share of the increase in English spending, regardless of whether or not the Scottish population is also increasing. For many years, relatively slower population growth in Scotland has limited the extent of Barnett convergence.

A second mechanical explanation why convergence in per capita spending has not occurred relates to a flaw in the treatment of income from Non-Domestic Rates in the Barnett Formula. Effectively, the Scottish budget was protected from budget cuts to local government spending in England after 2010 (Phillips, 2014). This flaw in the Barnett Formula was corrected in 2016, but the relative generosity it implied for the Scottish budget was retained within the grant baseline.

Why has the apparent generosity of the Barnett Formula been allowed to continue? The political economy literature has sought to explain the Barnett formula's longevity by characterising it as part of a more nuanced 'formula plus influence' funding system funding which has advantages for both the central and devolved governments (Christie and Swales, 2010), and which represents the 'codification of a bargaining outcome' (Smith, 2006). Simply put, this argument asserts that Scotland's relatively generous grant reflects its bargaining power, due to a combination of the threat to secede from the Union and its geographical and political distance from the centre (McLean and McMillan, 2005; Christie and Swales 2010; Hallwood and MacDonald, 2009).

Given the relative generosity of Scotland's block grant allocation, the Barnett Formula takes on totemic importance in any debate about Scotland's constitutional future. Indeed, one of the arguments made by supporters of independence in the lead up to Scotland's 2014 Independence Referendum was that, if Scotland voted to remain in the UK, the UK Government would 'punish' Scotland by abolishing the Barnett Formula and replacing it with a less generous mechanism for allocating grant to the Scottish Government. In response, the UK Government made a high profile 'vow' to maintain the Barnett Formula – alongside its promise for greater fiscal devolution – if Scotland voted to remain in the UK.

Even following the devolution of substantial tax revenues as highlighted above, the Scottish Government will be reliant on a block grant from the UK Government to finance around half of its spending activity via a block grant. In line with the commitment made by the UK Government prior to the Independence Referendum, this block grant will continue to be determined by the Barnett Formula.

But the Barnett Formula will clearly need to be adjusted to account for the transfer of revenue responsibility from the UK to the Scottish Government. How this adjustment should take place was at the heart of the negotiation over Scotland's Fiscal Framework.

v. Adjusting Scotland's block grant

An explicit objective of tax devolution to Scotland was to enable the Scottish budget to capture a greater proportion of the fiscal benefits of Scottish Government policy. This is true both of the first round of tax devolution realised through the Scotland Act 2012 (recall the previous quote from the Calman Commission), and the subsequent round of tax devolution realised through the Scotland Act 2016.

The design of supporting institutional frameworks, and in particular the manner in which a block grant is allocated, influence the incentive structure facing a devolved government, and the degree of exposure to economic shock that the devolved government faces.

The Smith Commission made recommendations about which fiscal powers to devolve, but did not set out specific details about the design of the supporting institutional architecture. The Commission did however set out some broad principles that the design of the Scottish Fiscal Framework should adhere to.

As well as recommending that the Barnett Formula should remain the mechanism for allocating block grant to the Scottish Government, these principles included that the Scottish and UK Government budgets should be no larger or smaller simply as a result of the initial transfer of revenue powers (the so-called 'no detriment from the decision to devolve' principle). The principles stated that the devolved Scottish budget should 'benefit in full' from policy decisions by the Scottish Government that increase revenues (and conversely bear the costs in full of policy decisions that reduce revenues or increase expenditures); and that the UK Government should manage economic shocks and budgetary risks that effect the whole UK.

The detail behind the Scottish Fiscal Framework was negotiated by the UK and Scottish Governments during 2015, with the Smith Commission's principles forming a sometimes contentious backdrop.

The basics of the new Fiscal Framework were fairly straightforward to agree. The requirement to retain the Barnett Formula following the transfer of revenues from the UK to the Scottish Government implied that, for each tax being transferred to the Scottish Government, a deduction would have to be made from the Barnett Formula calculated block grant. This 'block grant adjustment', or BGA, would effectively represent a counterfactual estimate of the revenues that the UK Government had foregone as a result of transferring the tax in question to the Scottish Government.

The BGAs would consist of two elements: an 'initial deduction' and an 'indexation mechanism'. The initial deduction is the revenue raised in Scotland (by the UK Government) from the tax to be devolved in the year prior to devolution (year 0). The indexation mechanism is a measure of the rate at which 'comparable revenues' have grown in the rest of the UK (rUK) between year 0 and any subsequent year in question.

By basing the BGA calculation on the revenues raised in Scotland from the tax to be transferred (the 'Initial Deduction'), this was seen to meet the Smith Commission's 'no detriment from the decision to devolve' principle. Scotland's budget is not disadvantaged (at least initially) relative to its Barnett

allocation, even if its tax capacity is lower than that of rUK. In other words, the approach equalises away Scotland's lower tax capacity at the point the revenues are transferred.

Subsequent application of the Indexation Mechanism to this initial deduction effectively means that the BGA increases each year in line with the growth in 'comparable' rUK revenues. This approach has the merit that, if Scotland's revenues grow more quickly than the comparable revenues of rUK – whether because of a tax rate increase in Scotland or more rapid growth in the Scottish tax base – then Scottish revenues would be higher than the BGA, and thus the Scottish budget would be better off than without tax devolution (and vice versa). But the general approach also means that the fiscal effects of a UK-wide economic shock would be borne by the UK Government – a decline in Scottish revenues would (if reflecting an equivalent decline in comparable rUK revenue) be offset by an equivalent decline in the BGA. In this respect the approach to indexing the growth of the BGA on the growth in comparable rUK revenues was seen to meet the Smith Commission principles in respect of protection against the budgetary implications of UK-wide fiscal shocks.

However, despite broad consensus between the governments as to the general structure of the BGAs, finding consensus on how specifically to calculate the indexation mechanism proved difficult (Bell et al., 2016).

Ultimately, the two governments agreed to index the BGA using a method known as 'Indexed Per Capita' (IPC), for the period until 2021 (at which point the indexation method will be renegotiated). IPC indexes the growth of the BGA according to the growth rate of per capita tax revenues in rUK, and the growth rate of Scottish population. In other words:

$$BGA_t = BGA_{t-1} \cdot \left(\frac{Pop_t^{Sco}/Pop_{t-1}^{Sco}}{Pop_t^{rUK}/Pop_{t-1}^{rUK}} \right) \cdot (TR_t^{rUK}/TR_{t-1}^{rUK})$$

The BGA in time period t is equal to the BGA in the previous time period, multiplied by the growth in tax revenues (TR) in rUK and a factor that accounts for the growth of rUK population relative to the growth of Scottish population.

Recall that the BGAs for each tax are deducted from the Scottish Government's block grant. What is added back into the Scottish budget are the revenues that are actually raised from each tax in Scotland.

Block grant — Block grant adjustment + Tax revenues raised in Scotland = Scottish budget
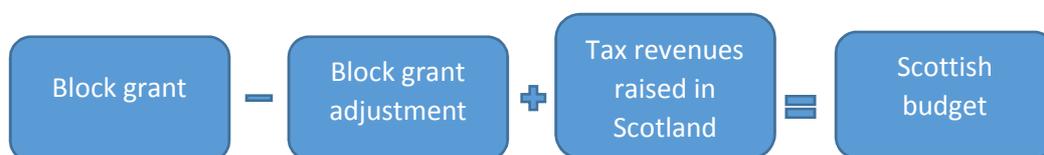
The key implication of the BGA arrangement is that, if the sum of the revenues raised from the devolved/ assigned taxes is greater than the sum of the BGAs, then the Scottish budget will be better off than it would have been without tax devolution.

This could happen under two circumstances: if the tax base grows relatively more quickly in Scotland than in rUK; or if tax rates in Scotland increase relative to those in rUK.

Of course the reverse could happen – Scottish revenues may grow relatively more slowly than those in rUK, in which case the Scottish budget will be worse off than it would have been without tax devolution.

The principle of the new fiscal arrangements is that the Scottish budget should '*benefit in full*' from policy decisions by the Scottish Government that increase revenues, and conversely bear the costs in full of policy decisions that reduce revenues. This is fine as a principle but the reality of the arrangements is that the Scottish budget bears in full the effects of any differential growth in Scottish revenues relative to rUK revenues, regardless of the causes of any differential growth. Whilst the Scottish Government would certainly hope to be able to implement policy to grow the Scottish economy and tax base, the link between policy and growth is often weak, and many of the factors determining tax revenue growth are only dependent on policy to a limited extent.

The implication of Scotland's new Fiscal Framework is that relative labour market performance becomes critically important in determining the budget of the Scottish Government, given the importance of labour market outcomes (wages and employment) in determining income tax revenues (with income tax being the largest source of decentralised revenue under Scotland's new fiscal framework). This observation forms the backdrop and motivation for papers 3 and 4 of this PhD.

Paper 3 (Chapter 4) examines the question of differential regional labour market performance by considering labour market trends across the UK regions since the Great Recession of 2008/9. It finds that there has been a strong regional dimension to the UK story. London has been at the vanguard of the recovery from recession in employment terms, but has experienced relatively greater real wage falls.

Paper 4 (Chapter 5) then builds on the largely descriptive analysis in Paper 3 in two ways: first, by examining the empirical link from regional labour market performance to regional income tax liabilities; and secondly by examining empirically the factors determining differential growth in regional wages, a major determinant of income tax revenues.

The paper raises important questions about the extent to which devolved or local governments can influence trends in local economic activity which drive trends in tax revenues in their areas. The Scottish Government does determine policy in many areas which might be material in influencing income tax revenues, including full control over education, health, business support, housing and infrastructure development. But it has little autonomy to determine policy in areas including labour market regulation, welfare, migration, and product market regulation. Local governments have substantially less policy autonomy.

And of course it is likely to be the case that geographical variation in tax revenues over time is partly dependent on factors beyond the direct influence of policy of either central or devolved government. A historic reliance on industries now facing decline, or conversely being intrinsically well suited to a growing industry, will influence revenue growth and thus devolved budgets.

The fourth paper examines the factors that influence differential growth in regional income tax revenues. This analysis concludes that, as well as economic indicators such as regional wages and employment rates, differential growth in regional income tax revenues is also influenced by the way in which the income tax base is defined at UK level. The decision of successive UK Governments to increase in real terms the Personal Allowance, whilst reducing in real terms the thresholds at which the Higher and Additional Rates of income tax become payable, higher income earners account for a growing share of income tax liabilities over time. This has implications for the relative growth of regional revenues per capita.

Notwithstanding this important issue, the paper goes on to consider the factors that might influence relative growth in regional wages. Wage growth is a major determinant of income tax revenues, and thus understanding the factors that influence this becomes an important question for devolved policy making. Paper 4 examines empirically the factors that have contributed to convergence of Scottish wages with rUK wages between 1999 and 2016.

vi. My role in the submitted papers

Scotland's Fiscal Future in the UK (Paper 1). David Bell suggested I write a paper looking at the options for and issues around further devolution to Scotland. The precise direction of the paper was determined by me. I met with David Bell regularly throughout the research and drafting of the paper, and David provided comments on drafts. David Bell also prepared Figure 4.1. A slightly different version of this paper was published as a Discussion Paper on the website of Scottish Fiscal and Economic Studies (ScotFES) https://esrcscotecon.files.wordpress.com/2014/09/scotlands-fiscal-future2.pdf

Assessing relative spending needs of devolved government: the case of healthcare spending in the UK (Paper 2). The intellectual idea to compare the English and Scottish healthcare allocation formulae by applying each in turn to Health Boards throughout the UK was David King's. My role was to implement the idea. I collated information about the structure of the two formulae, collated and applied the necessary data, analysed and interpreted the results, and wrote up the paper. I met with David King and Rob Ball regularly throughout the process to discuss methodological and data issues, and both David and Rob provided comments on paper drafts. The paper in this PhD draws heavily on the resulting paper published as: Ball R., Eiser D. & King D. (2015) Assessing relative spending needs of devolved government: the case of healthcare spending in the UK, *Regional Studies*, **49** (2), pp. 323-336. However, it also includes some additional material comparing approaches to measuring education spending needs, and additional discussion around the technical and political challenges inherent in developing needs assessment approaches to allocate resources to devolved government.

Migration and fiscal policy as factors explaining the labour-market resilience of UK regions to the Great Recession (Paper 3). In early 2015, I noticed that the Cambridge Journal of Regions, Economy and Society was inviting papers for a special issue on geographical aspects of the Great Recession. I drafted an extended abstract which was accepted for submission as a full paper. I collated the data and drafted the paper, meeting with David Bell regularly throughout to discuss general direction. The focus on migration came about as a result of a referee's sugesstion. The paper was published as: Bell D. & Eiser D. (2016) Migration and fiscal policy as factors explaining the labour market resilience of UK regions to the Great Recession*, Cambridge Journal of Regions, Economy and Society*, **9** (1), pp.197-215. However, the version in this PhD has undergone substantial revision, following comment from the External Examiner on the published version of the paper.

Tax devolution and accountability for revenue growth: lessons from Scotland (Paper 4). The idea for this paper, and the identification of the methodology used, was mine. David Bell provided substantive support in helping me to understand the mechanics of the methodology and the interpretation of the results, and well as providing helpful comments on paper drafts.

2. Scotland's fiscal future in the UK

David Bell and David Eiser

Abstract. This paper looks in detail at the proposals for further fiscal devolution to the Scottish Parliament following the No vote in September's independence referendum. The proposals vary significantly in terms of the taxes proposed for devolution and the balance between welfare spending devolution and maintenance of UK-wide social welfare union, but all envisage Scotland evolving into a semi-autonomous state within a more federal UK. Drawing on theories of fiscal federalism and empirical evidence from comparator countries, the paper considers the economic and constitutional constraints (including strategic tax competition, block grant allocation, and institutional factors) that may influence how effectively the Scottish Parliament will be able to exercise devolved tax powers. The paper considers the extent to which the models of fiscal autonomy proposed will mitigate future secession demands.

i. Introduction

The Scottish electorate voted against independence in the referendum held on September 18th 2014. During the referendum campaign, the leaders of the Conservative, Liberal Democrat and Labour parties made a commitment to introduce legislation at Westminster on a relatively rapid timetable that would bring further tax and spending powers to the Scottish Parliament.

This paper looks in detail at the proposals that these parties have made for further fiscal autonomy of the Scottish Parliament. It does so through the lens of fiscal federalism theory, which is the appropriate framework for the analysis of the allocation of powers in a multilevel government. It considers the economic and constitutional constraints that may influence how effectively the Scottish Parliament will be able to exercise devolved tax powers, and asks whether welfare spending devolution is consistent with notions of Union.

Theory suggests that fiscal decentralisation can bring economic dividends in terms of more efficient public sectors, higher levels of growth and greater public engagement in policy processes. However, realising these dividends is conditional on decentralisation conforming to a set of implicit criteria relating to the balance between levels of fiscal, political and administrative decentralisation, and the institutional arrangements underpinning decentralisation. In relation to this 'ideal type' of fiscal federalism, the main limitation of the Scottish Government's existing settlement is its very low reliance on own source tax revenues. Addressing this high vertical fiscal imbalance (the gap between expenditure budget and revenue raising responsibility) has been the focus of much of the recent debate in Scotland. Section 2 outlines the background to this, and the way in which it is being addressed through the recently enacted Scotland Act.

As discussed in the Introduction to this PhD, there is an argument that globalisation and international market integration strengthen the case for decentralisation by reducing the costs of smallness. The counter-argument is that the volatility and risk associated with globalisation might create a stronger role for federal governments in macro-economic stabilisation and inter-regional risk sharing. Both of these arguments are reflected in the various proposals for fiscal devolution in Scotland, which we detail in Section 3. These proposals differ markedly in terms of the taxes that are proposed for devolution, the nature of tax-sharing arrangements, and the extent to which welfare spending devolution is compatible with notions of a UK-wide social union and macro-economic risk-sharing. Section 4 compares the proposals with arrangements in other OECD countries.

Whatever fiscal devolution to Scotland is implemented, a number of factors will influence the extent to which the Scottish Government will be able to exercise its powers. In Section 5 we consider the significance of three of these factors in light of theoretical and empirical evidence. First, whether

strategic tax setting considerations will constrain the extent to which the Scottish Government will be able to achieve its preferred tax rate. Second, how pressure to reform Scotland's block grant will have to strike a balance between protecting Scotland's overall budget without eroding the incentives that tax decentralisation is intended to create. Third whether UK-wide symmetric institutional reform is required to realise the benefits of fiscal decentralisation.

We conclude that, following the No vote in September 2014, even the more modest of current proposals envisage Scotland evolving into a semi-autonomous state within a fiscally federal UK. The more ambitious proposals envisage Scotland becoming one of the most fiscally autonomous regions in the developed world, although there are many practical obstacles involved in realising this vision. The extent to which realisation of these proposals will mitigate future threats of secession will depend on the interplay of factors including the extent to which globalisation strengthens the importance of the stabilisation role of federal governments, on the way in which the notion of a social union is reflected in UK government policy, on the reform to Scotland's block grant (Barnett) arrangement and, possibly, on the trend of North Sea revenues.

ii. Fiscal autonomy and the Scotland Act 2012

As outlined in the Introduction to this PhD, the Scottish Government has traditionally faced a large vertical fiscal imbalance – its spending substantially exceeds its ability to raise tax. Scotland's large vertical fiscal imbalance has been seen as a disadvantage because it reduces the accountability of the Scottish Government to its electorate. The Calman Commission, established by the Unionist parties to review the financial accountability of the Scottish Parliament argued that: "*Funding by block grant alone means that while the Scottish Parliament is completely accountable for the spending of its budget, it is not accountable for the total of that budget or how it is raised; it has no fiscal powers that can be used as policy instruments and it does not have a direct financial stake in the performance of the Scottish economy*" (Commission on Scottish Devolution, 2009, para 3.87).

The Calman Commission's recommendations were largely enacted through the Scotland Act 2012. The main proposal in the Scotland Act 2012 was the establishment of a 'Scottish Rate of Income Tax' (SRIT), a flat rate tax set at the same rate for each (UK government-determined) income tax band (Figure 2.1). From April 2016, the basic, higher and top rates of income tax levied on earned income by the UK Government in Scotland were to be reduced by 10p. It would then be up to the Scottish Government to determine the SRIT, the income from which will form part of the Scottish budget. If the Scottish Government chose to set the SRIT at 10p, the basic, higher and top rates of income tax in Scotland would remain at the same levels as in the rest of the UK. If the SRIT is set at 9p, then the tax rates paid in Scotland at the basic, upper and additional levels would be 19p, 39p and 44p. These

SRIT arrangements were subsequently superceded by the Scotland Act 2016, as discussed in the Introduction, which legislates for the devolution of NSND income tax in full to the Scottish Parliament. This paper, which was written in 2014, prior to the Independence Referendum, considers the proposals for further fiscal devolution to the Scottish Parliament that were being made at that time.

Together with the already devolved council tax and business rates and some smaller taxes that will also be devolved to Scotland through the Scotland Act 2012 (landfill tax and stamp duty), the Scottish Government will be responsible for taxes equivalent to around 27% of its spending, (assuming it sets a SRIT of 10%). Perhaps more importantly, it will gain the ability to vary its budget at the margin, although the level of gearing is relatively high – increasing the SRIT from 10p to 11p (so that the tax rates paid by basic, upper and additional rate payers in Scotland was 21%, 41% and 46%) would increase the Scottish Government's annual budget by just 1.25%[1].

The remainder of the Scottish Government's budget will continue to be funded through a (reduced) block grant. The approach to adjusting the block grant was discussed in the introduction.

Figure 2.1: Income tax arrangements under the Scotland Act
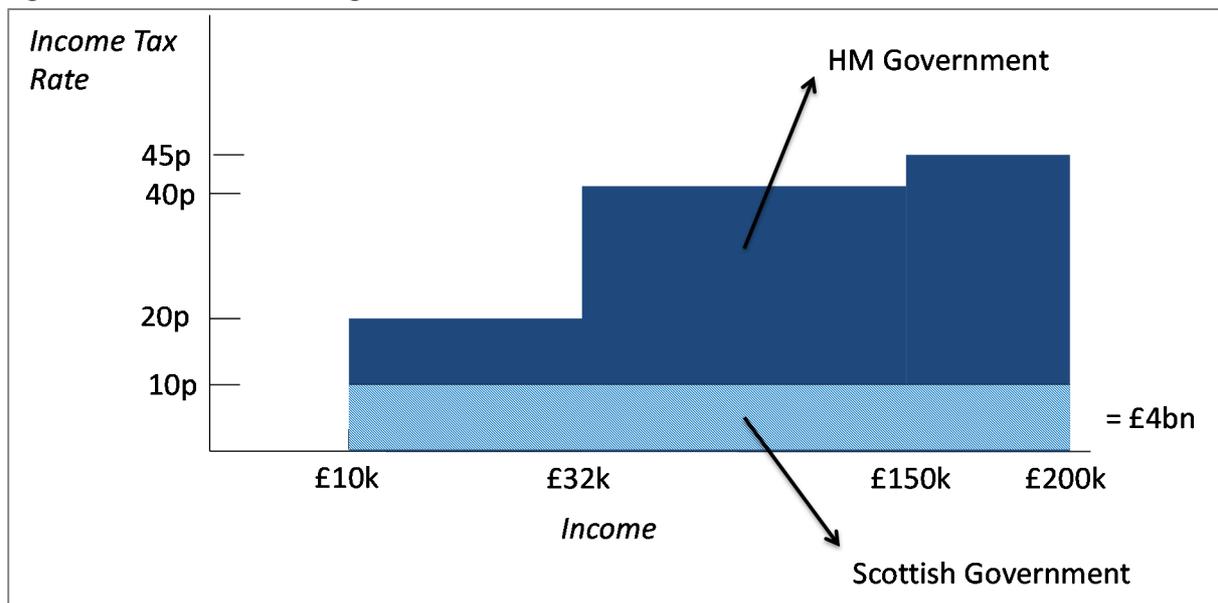
iii. Options for further fiscal devolution

Much of the debate on extended powers for Scotland has focussed on the scope for revenue devolution. The context was set by two high-profile think-tanks, the Institute for Public Policy Research (IPPR) and Reform Scotland, which published proposals for tax devolution under the

---

[1] In the fiscal federalism literature, this is known as high gearing because, to achieve a 1% increase in expenditure, the revenues from devolved taxation would have to increase by more than 1%.

banners 'DevoMore' and 'DevoPlus' respectively; following this, the Liberal Democrats, Labour and Conservative parties have each published their own proposals.

These proposals are summarised in Table 2.1. If the proposals have anything in common, it is a view that income tax is the most appropriate candidate for devolution. This reflects the view that it consists of a broad and stable base, is visible to the electorate, its burden falls largely on those who benefit from devolved services, and it is relatively easy to collect. DevoMore, DevoPlus, the Liberal Democrats and Conservatives all propose essentially 'full' devolution of income tax with the Scottish Government being given powers to vary tax rates and thresholds (although in most cases income tax on unearned income - investments, dividends and savings – remains reserved at Westminster given the scope for cross-border avoidance, and the Conservatives recommend that Westminster should continue to set the personal allowance). The Labour Party's proposals are somewhat more modest, amounting to an extension to the Scotland Act, so that the devolved part of income tax increases from 10p at each rate to 15p. The proposals also allow the Scottish Government to vary the progressivity of income tax, albeit it in a fairly restrictive way (the Scottish Government would be able to increase the progressivity of income tax rates relative to those in rUK, but not to reduce the progressivity). Note that giving Scotland full control over income tax would avoid the difficulties of vertical tax competition over a shared tax base, which we discuss subsequently.

The next most significant tax in revenue terms is the sales tax, VAT. A number of the proposals (including DevoMore and the Conservatives) argue that this is theoretically a suitable tax for devolution, but suggest that EU Law on state aid precludes the devolution of this tax *within* the UK. As a consequence, DevoMore recommends that half of the VAT revenues raised in Scotland should be assigned to the Scottish Parliament, a proposal which the Conservatives support in principle. The debate is whether such assignment provides the Scottish Parliament with sufficient revenue control to justify the exposure to revenue risk.

National Insurance Contributions (NICs) act like a tax on earnings, but their payment entitles individuals to certain 'contributory' social security benefits that are paid at UK level. It is this link between NIC contributions and benefit entitlement that has resulted in most proposals for tax devolution arguing that NICs should remain reserved. However, the connection between NICs and entitlement has weakened over time, and there is some disagreement over whether this contributory argument is a strong one mitigating against devolution of NICs, or whether NICs should be aligned with income tax and treated as such.

Beyond the big three taxes, (and given that property taxation, the most obvious candidate for decentralisation to a sub-national government is already devolved), there is not a very close correspondence between the proposals.

- Corporation Tax has been proposed for devolution by DevoPlus on the basis that economic and business development are devolved policy areas, although most proposals recognise the risks inherent in devolving corporation tax, namely the fact that the high mobility of the tax base may trigger tax competition.

- Proposals to devolve alcohol and tobacco duties are based on the fact that these 'sin' taxes have a clear link to devolved health policy, but there are serious practical difficulties in devolving these taxes given that they currently operate as a tax on production rather than consumption (Trench 2013), and some fears that differential rates might result in 'illicit trafficking (Scottish Conservatives, 2014).

- In terms of smaller taxes, there is some consensus that Air Passenger Duty should be devolved as a place-based tax, but less consensus on the suitability of devolving wealth taxes such as Capital Gains and Inheritance tax, which are low-visibility and low-yielding.

What emerges from this discussion is that, with the possible exception of income tax on earned income, there are no obvious candidates for further tax devolution to the Scottish Parliament, emphasising Bahl and Cyan's (2011) point that, if the theory of tax assignment was followed to the letter 'there would surely be no subnational government taxes'. The various proposals vary substantially in the proportion of the Scottish Parliament spending that would be covered by taxes raised in Scotland, from just under one third under Labour's proposals to 55% under the Lib Dem proposals, and around two-thirds under the DevoMore and DevoPlus proposals. What also emerges from Table 2.1 is that the tax base for several of the taxes proposed for devolution is lower in Scotland than in rUK. We return to this point subsequently.

Table 2.1: Proposals for tax devolution

	£m (2012/13)	Index of revenues per capita relative to UK	Devo-More	Devo-Plus	Scottish Conservatives	Scottish Liberal Democrats
Income tax	10,865	-12%	✓	✓	✓ (personal allowance reserved)	✓
VAT	9,347	0%	Shared		Potentially shared	
National insurance contributions	8,521	-2%	Devolved longer term			
North Sea revenue (geographic share	5,581	909%		Devolved longer term		
Corporation tax (excl North Sea)	2,872	0%		✓ (thresholds reserved)		Shared
Fuel duties	2,258	2%		Devolved longer term		
Council tax	2,006	-5%	Already devolved			
Non-domestic rates	1,981	-8%	Already devolved			
Tobacco duties	1,128	41%	Devolved longer term			
Other taxes, royalties and adjustments	1,082					
Alcohol duties	980	16%	Devolved longer term			
Vehicle excise duty	481	-4%		Devolved longer term		
Stamp duties	472	-38%	Devolved through Scotland Act			
Capital gains tax	292	-11%	Devolved longer term	✓		✓
Other taxes on income and wealth	271	4%				
Inheritance tax	243	-8%		✓		✓
Air passenger duty	234	0%	✓	✓	✓	✓
Insurance premium tax	207	-18%		✓		
Betting and gaming and duties	120	17%		Partially devolved		
Landfill tax	100	7%	Devolved through Scotland Act			
Climate change levy	62	14%				
Aggregates levy	45	107%	✓	✓		

Notes: Scottish Labour proposals are not shown as they relate mainly to income tax.

The main area of territorially identifiable public spending in Scotland that remains 'reserved' to the UK Government is welfare[2]. This includes virtually all cash transfer benefit payments associated with low income, unemployment and disability, as well as the State Pension and other old age benefits. Spending on these benefits amounted to £16 billion in 2011/12, representing half the value of the Scottish Government's existing devolved budget.

As well as debate around revenue devolution, there is also debate around which aspects of welfare benefit spending might be devolved to the Scottish Parliament. The areas of welfare spending most often cited as candidates for devolution include:

- Housing Benefit (a benefit to support those on low incomes with housing costs) has been proposed for devolution by Labour and DevoMore. It is a significant benefit in cash terms, accounting for £1.7 billion. The argument for devolving it is that it is linked to various areas of devolved policy, including social housing and planning. Furthermore, as a place-related benefit, Housing Benefit rates vary according to local conditions, and it is also reasonably stable over the business cycle.

- Attendance Allowance (AA) has been proposed for devolution by Labour, DevoMore and DevoPlus. AA is a benefit that is intended to help with personal care for those aged 65 or over who are physically or mentally disabled and policies. The rationale for devolving it is that Scotland already has a distinct policy with regards to elderly care, with the result that the UK and Scottish systems overlap. The same arguments could be made in respect of Disability Living Allowance paid to pensioners, though this was not part of any of the proposals.

- DevoMore and DevoPlus have also proposed devolving the Work Programme. The Work Programme is the UK Government's flagship programme for supporting the unemployed into work. The rationale for devolving it is that the Scottish Government has responsibility for skills and training policy.

- The Scottish Conservatives have also called for greater flexibility in the way that welfare benefits are provided in Scotland's devolution settlement, arguing for the Scottish Parliament to have the power to supplement existing welfare benefits legislated for at the UK level. This notion of being able to 'top-up' reserved benefits with own resources was similarly proposed by DevoMore.

---

[2] The other main area of reserved territorially identifiable spending in Scotland is associated with science and technology. Reserved spending that is not territorially identifiable largely relates to defence and foreign affairs.

However, the case for devolving aspects of welfare spending is arguably more difficult to make than the case for devolving tax responsibilities, for both reasons of principle and practicality.

In terms of principles, most proponents of further fiscal devolution have also been keen to stress the importance of maintaining the UK's social welfare union. The UK welfare state is seen as the key element in the risk sharing and resource pooling mechanisms that are seen as a defining characteristic of the Union. Thus the Labour proposals talk of 'rights enshrined at UK level that should be paid for from UK taxes', the Liberal Democrats talk of maintaining the UK's social welfare union. The Conservatives describe this social union as 'hugely important to what glues us [the UK] together'.

In this context, it is unclear to what extent welfare devolution is compatible with these principles. The rationale for devolving AA is that Scotland already has a distinct policy with regards to elderly care, with the result that there are both areas where Scottish and UK policy overlap, but also where there exists a different entitlement in Scotland relative to rUK. Scotland has taken a different policy decision on the issue of elderly care, and it is not clear that in these circumstances the universal insurance principle that underlies the UK benefit system continues to apply. Such a principle has to be built on consensus around the set of adverse events that the state should insure. Scotland has taken the view that personal care should be one of these: the rest of the UK has not.

In terms of practicalities, one issue is that the UK Government is currently in the process of combining Housing Benefit into a new benefit, Universal Credit, which brings together six existing means tested benefits for those on low incomes. The rationale for Universal Credit is to simplify the benefit system and avoid the current situation whereby benefit claimants sometimes face particularly high work disincentives as different benefits are withdrawn simultaneously. Devolving HB will almost certainly negate the proposed advantage of UC, namely that of receiving one benefit from one agency, rather than overlapping benefits from different agencies that tends to result in high benefit withdrawal rates. The practical issue that arises in relation to the Work Programme is that, given that success in placing people into work would generate positive externalities for the UK Exchequer, some kind of inter-governmental financial compensatory mechanism would perhaps have to be built into the arrangements to provide the Scottish Government with appropriate incentives for delivering the Work Programme effectively.

In summary, it is difficult to envisage meaningful devolution of welfare spending in a way that does not threaten the notion of there being some social welfare union that enshrines rights to fundamental welfare services across the UK. Because of this, there is arguably more contentiousness around welfare spending devolution than there is around revenue devolution. Furthermore,

devolution of welfare spending risks amplifying the issues around the vertical fiscal imbalance that the Scottish Government already faces.

iv. The Scottish proposals in an international context

There is often a perception in the UK debate that Scotland has far less fiscal autonomy than decentralised regions in federal countries. It is useful to consider how the proposals described previously would, if implemented, alter these comparative statistics.

Figure 2.2 compares fiscal decentralisation ratios in selected OECD countries with the proposals for further fiscal devolution to the Scottish Parliament. The horizontal axis plots the share of sub-central government (SCG) expenditure in total government spending, and the vertical axis plots the share of SCG tax revenue in total government revenues. For the UK as a whole, SCG (i.e. local authorities in England and the three devolved governments) account for 27% of total expenditure but only 5% of total revenues. However, the asymmetric nature of devolution in the UK means that this statistic is not particularly meaningful when considering devolution in Scotland specifically. Instead, it is more relevant to consider the SCGs (i.e. Scottish Government and local authorities) share of all revenues and spending in Scotland. On this basis[3], SCG accounts for 50% of all public spending but only 8% of total tax revenues in Scotland, shown by the point 'UK (Scotland)'.

The Scotland Act proposals result in the Scottish Government's revenue share increasing to 17%, whilst the Scottish Labour proposals would increase both the revenue and expenditure decentralisation ratios slightly further (the expenditure share increases because of the proposal to devolve expenditure on housing benefit to Scotland).

The Devo-More and Devo-Plus proposals result in radical increases to the Scottish Government's revenue share. When fully implemented, the Devo-More proposals bring the Scottish Government's revenues into line with spending (although this is achieved in part through the assignment of some VAT revenues; the extent to which tax assignment can be thought of as tax devolution is discussed below). The Devo-Plus proposals, when fully implemented, would see the Scottish Government responsible for 65% of all public spending in Scotland and 56% of all revenues raised. The Devo-More and Devo-Max proposals, fully implemented, would effectively imply that the Scottish Government is one of the most fiscally autonomous sub-central governments in the world[4].

---

[3] We allocate to Scotland per capita shares of UK spending on debt interest, non-identifiable public services such as defence and foreign affairs, and a geographical share of tax revenues from North Sea production.
[4] Some other countries also have asymmetric decentralisation settlements which are not reflected in this chart. In Spain for example, the Basque and Navarre regional governments operate under the devo-max model, and thus have a higher level of fiscal autonomy than the Scottish Government would under either Devo-More or Devo-Plus. Similarly, Quebec has a somewhat asymmetric fiscal settlement in Canada.

Figure 2.2: Decentralisation ratios in OECD countries (2010)
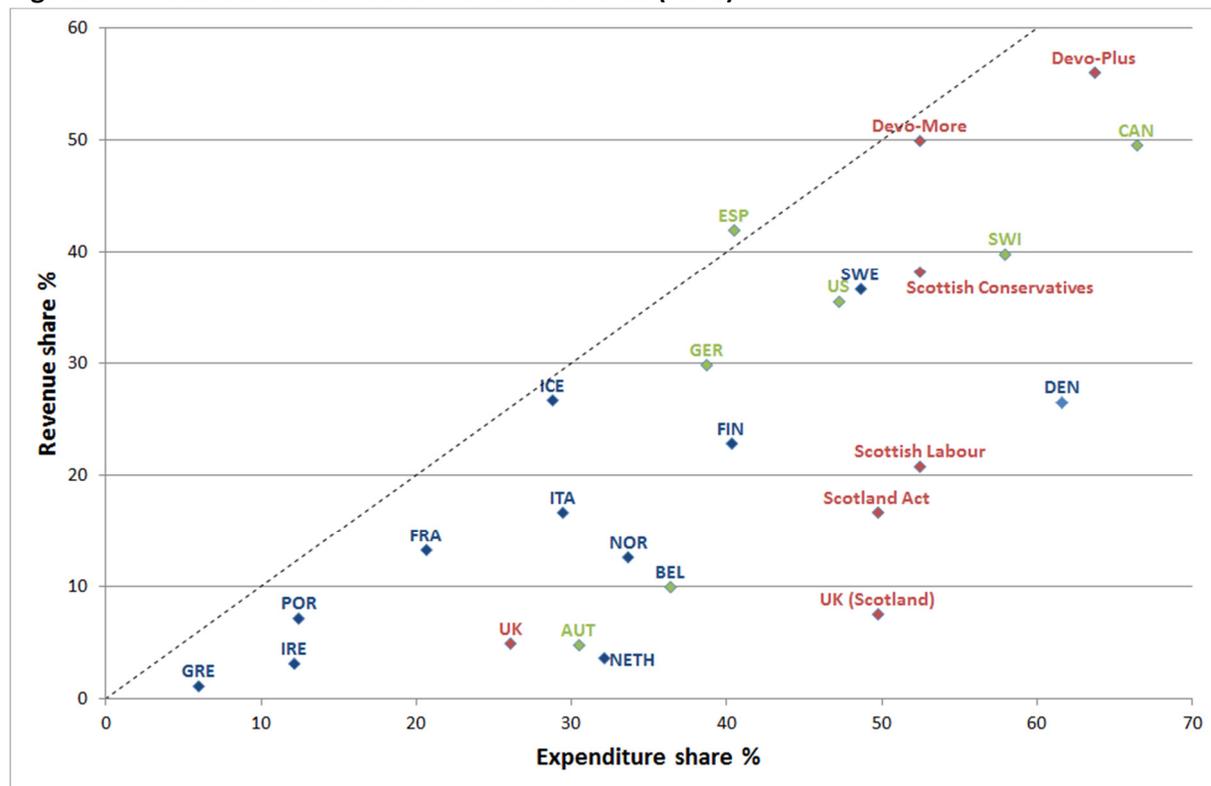
Source: OECD and author analysis. Federal countries are plotted in green, unitary countries in blue.

However, looking at simple measures of revenue shares does not tell the whole story. SCGs have varying levels of control over 'devolved' taxes. Broadly speaking, three types of tax decentralisation are most often implemented. First, SCG has full control over a tax if it can vary both the tax rate and the tax base (i.e. thresholds). The second category of tax decentralisation is where SCG can only vary the tax rate. Third, tax assignment arrangements occur when a tax is shared between different levels of government, with the tax rate and base either agreed bilaterally or, more often, unilaterally by the central government.

Figure 2.3 shows the tax power of SCG in selected OECD countries. Some countries, notably Germany, which initially appear to have high levels of tax decentralisation, actually use tax-sharing arrangements to a large extent, giving SCG little direct autonomy over tax revenues. Others such as Finland and Denmark constrain SCG to have control over rates only. The Scotland Act proposals will give the Scottish Government control over tax rates but not the tax base. The Devo-More and Devo-Max proposals will bring significantly more autonomy to vary tax rates and bases, although Devo-More also makes use of tax-sharing arrangements in the case of VAT. (As in the previous graph, the data points for Spain and Canada consider the average arrangements for Autonomous Communities

and Provinces respectively, and do not reflect the arrangements of 'special' regions – Basque and Navarre in Spain, and Quebec in Canada).

Figure 2.3: Tax power of sub-central governments in OECD countries

Source: OECD and author analysis.
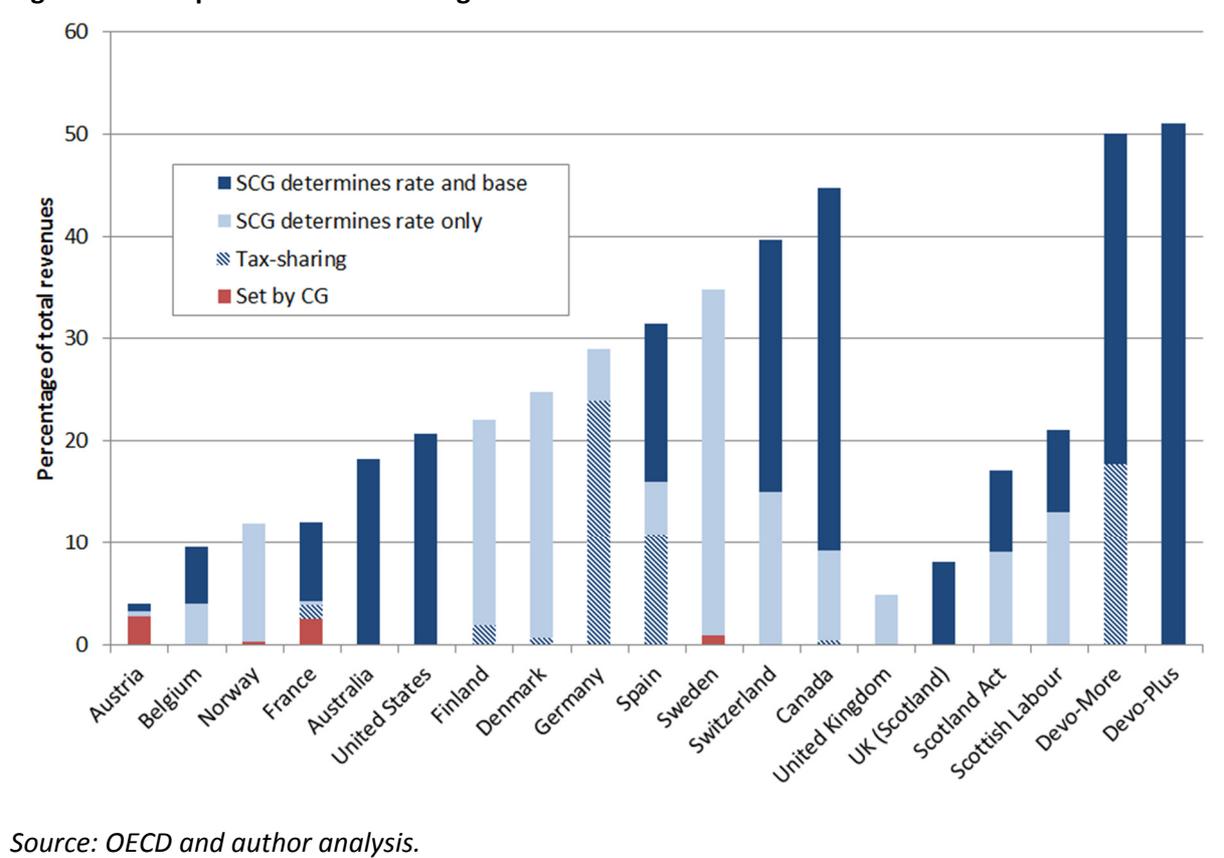
v. Issues and constraints

The debate around further tax and spending devolution to the Scottish Parliament takes place alongside several related fiscal debates, and we review three of these here. First, to what extent would the Scottish Parliament be constrained in how it uses new tax powers given fiscal interactions with the UK Government; second, how might fiscal devolution interact with block grant reform and what implications will this have for the Scottish Government's grant; third, whether institutional reform needs to occur in tandem with fiscal devolution.

Fiscal competition

Theoretical and empirical studies in the economics of fiscal federalism indicate that there are likely to be significant inter-dependencies between national and devolved governments in how the tax rate is set. The literature identifies two types of taxation externalities: a horizontal externality, working among governments of the same level; and a vertical externality, working between different levels of government.

Theory suggests that the presence of vertical externalities is likely to lead to tax rates that are too high compared to the social optimum, because different layers of government do not consider the effect of their tax rate on erosion of the tax base. Dahlby and Wilson (2003) describe this as 'the public sector version of a common property resource problem', leading to over-exploitation of shared tax bases (in effect, the regional government takes its marginal cost of public funds, MCPF, function as being lower than the social MCPF). However, the few empirical studies that have looked at this issue have found evidence that the tax rate set by regional governments can be both positively or negatively related to the national level tax, depending on the tax considered and various institutional factors (Esteller-Moré and Solé-Ollé, 2001).

On the other hand, theory suggests that horizontal externalities, which have received much more attention in the literature, are likely to lead to tax rates that are too low relative to the social optimum. Horizontal externalities are most often discussed in relation to fiscal competition (where, given tax base mobility, the tax rate on a mobile factor can affect the budget constraints of neighbouring jurisdictions). But they can arise through other channels, including expenditure spillovers (where the public services provided by one jurisdiction enter the welfare function of a neighbouring jurisdiction because their citizens might benefit from such expenditures) and yardstick competition (where the views of an imperfectly informed electorate as to the efficiency of their jurisdiction in the provision of public services are inferred in part by using neighbouring jurisdictions as a yardstick, which leads to mimicking) (Revelli, 2005).

A large number of empirical studies find evidence for horizontal tax interdependencies. Evidence suggests that horizontal tax interactions are stronger for more mobile tax bases (personal and business taxes) than property or consumption taxes, stronger among smaller local governments than larger regional governments, and are reduced by equalisation grants. There is evidence of spatial autocorrelation in tax rates across jurisdictions in Switzerland and Sweden. There is also evidence that much of this competition is driven by factor mobility. Milligan and Smart (2013) for example show that the shifting of taxable income between provinces accounts for about two thirds of the total tax avoidance in response to unilateral provincial tax changes in Canada. Day and Winer (2001) find evidence of fiscally-induced migration across Canadian provinces. Feld and Kirchgasser (2002) show that income tax-rate differences across Swiss Cantons influence firm location through the influence of firms' ability to attract skilled labour.

However, there is also evidence that tax competition between sub-central governments is not as strong as is sometimes supposed. One reason for this is that higher tax rates are generally associated with better public services, thus competition should be thought of in terms of fiscal competition

rather than purely tax competition. Moreover, although tax base mobility is increasing over time, particularly for the highest earners (OECD, 2013), the extent to which tax or fiscal competition triggers migration depends in part on the extent to which people respond to a tax rate increase by working less (rather than migrating), and the extent to which tax differences are capitalised into house prices. Hilber et al (2011) find that windfall central government grants to English local authorities are almost fully capitalised into house prices, (although whether such an effect will mitigate the migratory incentives of such grants or itself is the result of a migratory response is less clear). Brulhart and Parchet (2014) show that, although a 'domino-like' reduction in Swiss Cantonal inheritance tax rates over a long period was blamed on concerns about tax base competition, mobility of the relevant tax base (wealthy retirees) is in fact very inelastic with respect to the tax rate. They conclude that 'alleged pressures of tax competition did not seem in reality to exist'.

The OECD (2013) finds that, in most countries, SCG tax rates have converged rather than diverged over time, but that they have converged upwards (perhaps counter to the expectation of those who predict a 'race to the bottom' in tax rates). This might reflect the fact that, for most taxes, the elasticity of revenues with respect to the tax rate is less than one, so that SCGs have to raise rates to raise revenues in response to demand for higher spending; taxes for which the elasticity of revenues with respect to the rate is greater than one are generally not devolved (e.g. corporation taxes), precisely because of the race to the bottom argument.

A priori, in cases of a shared tax (such as income tax under the Scotland Act) it is not clear whether the horizontal or vertical externality is likely to dominate, (Brulhart and Jametti, 2006). This is particularly the case when considering the Scotland Act, as it will involve the Scottish Government facing vertical and horizontal fiscal externalities from the same Westminster Government (which is in effect setting a UK-level income tax and the 'devolved' element of income tax in rUK).

More generally, the extent to which the Scottish Government will feel constrained in setting a different tax rate from that in rUK for any devolved tax will clearly depend on the mobility of the tax base. In relation to income tax, although there is some evidence that households are generally not very mobile between Scotland and rUK, mobility increases rapidly with income, and with the highest earning 10% of Scottish income taxpayers accounting for almost half of Scottish income tax revenues (Bell and Eiser 2013), it may be the case that the Scottish Government faces a relatively high marginal cost of public funds where these are funded through income tax. The Scottish Government has always had powers to vary income tax rates slightly, but these have never been exercised. On the other hand, a tax rise could have a relatively elastic effect on revenues if preferences are conducive to this (Comerford and Eiser 2014). The timing of tax decisions is also important, and it

may be that the Scottish Government is more inclined to 'follow' the UK Government if the UK Government 'goes first'.

A related question is the extent to which such tax interaction effects would be any less if Scotland was fully independent. One the one hand, vertical fiscal externalities would no longer be relevant. But the risk of mobility induced horizontal tax competition would still be present, particularly in relation to income tax rates for high earners (Landais et al 2013). In the long-run, horizontal externalities in tax rates between countries may be just as high as those between a semi-autonomous Scotland and rUK, given global integration.

Block grant

The Scottish Government relies on a block grant from the UK Government to finance the large majority of its expenditure. Even under some of the more radical proposals for revenue decentralisation to Scotland, the Scottish Government will continue to face some level of vertical fiscal imbalance. The way in which this grant is allocated, and the extent to which it attempts to equalise differences in spending need or revenue capacity will have profound implications for the Scottish Government's budget.

The Scottish Government's existing block grant is determined by the Barnett Formula, which is also used to allocate funding to the Welsh Government and Northern Ireland Assembly (HM Treasury, 2010). The Barnett Formula has been extensively criticised on several grounds:

- First, it takes no account of the relative spending needs of the UK territories and is therefore seen as inequitable. It is generally perceived to provide a particularly generous grant allocation to Scotland. This has long created a source of considerable tension in UK territorial politics which has intensified in recent years, with the UK media often quoting an estimate by Gerald Holtham that the Scottish Government is 'overfunded by £4 billion'. (However, the generous allocation to Scotland is largely the result of a generous baseline allocation rather than with the formula itself).
- Second, the formula bases changes to the Scottish Government's budget on policy changes in England. The notion of basing grant to a devolved government on the policy decisions of another parliament is argued to be inconsistent with the aims of devolution, effectively forcing the Scottish Government to become a 'policy taker' (Trench, 2012).
- Third, the formula – although simple in theory – lacks transparency in how it is operated in practice by the UK Treasury.

Despite these criticisms, the Barnett Formula has remained the preferred mechanism for allocating grant to the devolved governments since 1979, suggesting that it also has some advantages. One is that it provides stable and predictable funding (Select Committee on the Barnett Formula, 2009, paragraphs 49-50). The stability arises because the formula applies only to change in grants, ensuring that the Treasury cannot attack the budgets of the DGs collectively or sequentially.

The formula's longevity is usually explained by characterising it as part of a more nuanced 'formula plus influence' system which has advantages for both the central and devolved governments (Christie and Swales, 2010), and which represents the 'codification of a bargaining outcome' (Smith, 2006). Simply put, this argument asserts that Scotland's relatively generous grant reflects its bargaining power arising from the threat to secede from the Union (McLean and McMillan, 2005; Christie and Swales, 2010; Hallwood and MacDonald, 2009). In this respect, it is noted that the extent to which Scotland is 'over-funded' through the Barnett formula is almost exactly offset by the value of North Sea oil and gas taxation revenues derived from Scottish waters (Ashcroft 2013; Armstrong and McLaren 2014).

Following the implementation of the Scotland Act 2012, the Scottish Government will continue to receive a Barnett Formula determined block grant from the UK Government, but this grant will be adjusted downwards to reflect the revenues raised through the SRIT. The arrangements for this deduction were discussed in the introduction.

Dissatisfaction with the Barnett Formula has led to growing calls for it to be replaced with a formula which would allocate grant based on assessment of the devolved territories relative spending needs (Select Committee on the Barnett Formula, 2009; Independent Commission on Funding and Finance for Wales, 2010; All Party Parliamentary Group on Taxation 2013; Local Government Association in England, 2013). Calls to replace Barnett with a needs-based formula are particularly strong in Wales, indeed desire for devolution of income tax powers to Wales is predicated on reform of the Barnett Formula (The Silk Commission). In contrast, most of the Unionist party's proposals for devolution in Scotland envisage a continuation of Barnett, with only the Scottish Liberal Democrats explicitly in favour of adopting a needs-based formula.

Scotland's spending on devolved services is some 20% per capita higher than England's. Most analysis suggests that Scotland faces relatively higher costs (i.e. spending needs) for delivering public services, but that the costs gap is not as large as 20 per cent. The Holtham Commission estimated that Scotland's per capita spending need for devolved services is 6% higher than England's. Scotland's relative per capita spending need for health services has been estimated as 7-11% higher than England's (Ball et al. 2015), whilst per capita spending need for school-age education services is

estimated as 1-2% lower in Scotland (King and Eiser, 2016); weighting these two service areas according to their share of the total Scottish budget also suggests an aggregate relative need 6% higher than England's.

However, although needs assessment is used extensively within each of the UK territories to distribute resources to health boards, local authorities, schools, and colleges (Smith, 2006; National Audit Office, 2011), allocating grant to a semi-autonomous devolved government according to spending need is likely to be particularly challenging. The normative aspect of determining needs raises difficult questions about which needs are 'legitimate' rather than reflecting policy choice, and which indicators should be used to measure needs; consensus between the UK and Scottish Governments may be unlikely (Midwinter, 2006; McLean and McMillan, 2005). Implicitly, needs assessment requires a broad consensus between governments about the size and scope of the public sector, so that needs assessment might be problematic if this consensus does not exist (because of divergent policies in the funding of major public services for example). The dominance of England in UK population terms is likely to undermine any attempt to derive some measure of 'average' policy across the UK (King and Eiser, 2016).

Issues around the allocation of block grant on the basis of needs assessment are discussed in the subsequent paper in this PhD.

It is also not clear how needs assessment arrangements could be given the statutory backing that would provide any guarantee of their longevity. Laws (2008) argues that "entrenchment" – insulating legislation from future repeal – is inherently difficult in the absence of a written constitution. However, he argues that statutes are now regarded as implicitly either 'ordinary' statutes or 'constitutional' statutes. He cites as examples the Act of Union, the Scotland Act (1998) and the Government of Wales Act (1998). In his definition, a constitutional statute is one which enlarges or diminishes the scope of what we would now regard as fundamental constitutional rights. Whether legislation to support the use of needs assessment as a mechanism for allocating resources to the constituent parts of the United Kingdom could be thought of as affecting fundamental constitutional rights seems doubtful. Embedding such an arrangement as part of a wider agreement between the home nations that clarifies their inter-relationships would perhaps be a necessary condition to ensure the longevity of the allocation mechanism.

The issue of whether a needs-based grant system can be consistent with the notion of semi-autonomous devolved government is a matter of contention. Canada has repeatedly rejected the idea that grant should be allocated to provinces based on spending need, as this is seen to threaten provincial autonomy (Lecours and Belland, 2010). Germany does make limited grant allocations to

Lander on the basis of spending needs, but this must be seen within the context of Länder which have little policy autonomy and largely deliver federal policy at a decentralised level. Spain does allocate grant to its Autonomous Communities (ACs) on the basis of spending need, but the formula is very simple, reflecting the parameters that the ACs could agree on, and largely consists of a small number of basic demographic variables (Bosch, 2009). It is only in Australia where extensive needs assessment is used to allocate grant to States. Although the system relies on a complex set of calculations to determine average policy levels, it is not without controversy.

Given that Scotland's spending needs for devolved services are somewhat higher than those in rUK, a block grant allocated to Scotland without any allowance for differential spending needs, but instead allocated on a per capita basis, is likely to result in an inferior overall level of public services in Scotland compared to rUK. Although the assessment of spending needs is difficult in the case of devolved governments, it may be that there is scope for acceptable compromise around a relatively simple formula. Securing a grant which allocated it 106% of equivalent English spending would make a marked difference to the Scottish budget, whilst costing the UK Government less than the existing Barnett arrangement.

In addition to having somewhat higher spending needs, Scotland also has a lower tax capacity (i.e. ability to raise revenues on a given base at a given rate) than rUK for the taxes that are most likely to be devolved (Table 2.1). This suggests that the block grant to Scotland should also make some adjustment to equalise revenue capacity, if equality of public services across the Union is a political objective.

Unlike spending needs, equalisation of revenue capacity is undertaken in all federal countries. Such revenue equalisation grants aim to strike a fine balance between securing some notion of comparable public services at a given level of taxation on the one hand, and providing adequate tax incentives to devolved governments on the other (OECD, 2013).

In practice, revenue equalisation varies significantly across federations (Table 2.2), reflected in differences in marginal equalisation rates – the rate at which a jurisdiction's own revenue is taxed away or cancelled out in the form of lower grant. In Germany, any Länder with a fiscal capacity less than 99.5% of the average has over three quarters of its shortfall corrected. The extent of equalisation implies that every additional euro collected by a state on its own leads to a reduction of receipts from transfers by an almost equal amount, providing little incentive for Länder to generate additional revenue from economic growth or tackling tax fraud (Feld, 2011; Buettner 2008). In Canada the extent of equalisation is relatively less, but this in turn creates tensions as fiscal

disparities widen, in turn placing strain on tax harmonisation arrangements (Lecours and Belland, 2010).

The mechanism by which the Scottish block grant will be adjusted to account for tax devolution were outlined in the Introduction. In essence the 'block grant adjustments' mean that Scotland's lower initial tax capacity is fully equalised at the point of tax transfer. However, the Scottish budget is fully exposed to the risk of differential tax base growth in future. The implications of this arrangement for the Scottish budget are discussed further in the final paper in the PhD.

Table 2.2: Marginal equalisation rates of sub-central government in federal countries

Country	Marginal equalisation rates of SCG in federal countries
Austria	88% for Lander below average fiscal capacity, 0% for those above
Canada	70-100% for provinces below average fiscal capacity, 0% for those above
Germany	100% for Lander below 99.5% of average fiscal capacity, 15-85% for those above
Spain	75%
Switzerland	40-60%

Source: OECD (2013)

Allocation of block grant to the Scottish Government through the Barnett Formula has been a source of tension for many years. The demise of Barnett has been forecast repeatedly since 1999, but successive UK Governments have been unwilling to address the issue, even though the formula has no statutory basis. There is a strong case for arguing that the Barnett mechanism is not suited to a more fiscally autonomous style of devolved government in Scotland. Furthermore, if some income tax devolution to Wales is implemented, this will increase pressure for Barnett reform. However, the political economy literature (outlined in the Introduction to this PhD) would suggest that it may continue to be in the interests of the UK Government to fund Scotland relatively generously if it wants to mitigate future secession demands. And given the difficulties around the design of a spending needs formula, and issues around the incentive effects of revenue equalisation grants, it remains unclear as to how high a priority replacement of the Barnett formula will be.

Institutions

A view is sometimes expressed that significant tax devolution to Scotland needs to coincide with institutional reform, and a move to a more symmetric, federal structure of government for the UK as a whole. The Liberal Democrats Home Rule proposals for example envisage a federal UK government – with competency in areas such as foreign and defence affairs, currency, etc. – combined with various home rule governments with equivalent devolved powers, including some form of government for England.

However, the UK is probably closer to being a fiscally federal state than is sometimes recognised, when this is narrowly defined. One such definition of a federal system is where 'two levels of government rule the same land of people, with each level having at least one area in which it is autonomous, and where there is some constitutional guarantee of the autonomy of each government in its own sphere' (Riker 1964). Building on this, Sorens (2014) identifies the 'ideal-type' of fiscally federal system as being one where sub-central governments: have pragmatic autonomy (i.e. exclusive authority to determine a subset of government policy); face a hard budget constraint and fund expenditure largely from own-source revenues; where there is a common market for goods, labour and capital across borders which the SCG cannot disrupt; and where the system is institutionalised. Federalism so defined does not require the participation of sub-central governments in central government decision-making.

Thus a key difference between devolved government in a unitary system and a federal system is that, in a federation, the jurisdiction of each level of government is constitutionally recognised. The Scottish Government is clearly more than an administrative creature of UK Government, with complete political autonomy (i.e. the UK Government cannot appoint politicians to the Scottish Parliament) but it is not constitutionally entrenched, and the UK Government could theoretically abolish, change or over-rule it, or legislate within devolved spheres. However, all pro-Union parties now support the constitutional entrenchment of the Scottish Parliament. Once entrenched constitutionally, and with a substantially reduced vertical fiscal imbalance, the Scottish Parliament would, arguably, accord with a minimalist definition of federalism.

More pertinently, it could be argued that the Scottish interest is well served by a more specific form of federalism which Lluch (2012) refers to as 'autonomism'. Autonomism, according to Lluch, is based on the general federalist principle that multiple levels of government can lead to better governance in multinational states and is by definition assymetric. But whereas federalism flourishes where there is trust and reciprocity between federal and subcentral governments, autonomists tend to be wary of federalism because they believe it has homogenizing and uniforming tendencies. Instead, autonomism is 'the search for gradually expanding spheres of self-government within existing state structures'.

Thus symmetry of fiscal decentralisation is not a prerequisite for federalism, and nor does federalism require the participation of regional governments in central government decision-making. But federalism does require that the devolved government is not subordinate to the centre, that it is constitutionally entrenched, that it has pragmatic policy autonomy, and that is has reasonable fiscal autonomy.

This is not to say that, beyond the entrenchment of the Scottish Parliament, no further constitutional reform is necessary. In particular, there seems likely to be a need for reconsideration of the institutional machinery within which constructive intergovernmental relations can be managed (particularly given the fractious nature of the debate leading to the independence referendum). But there is no reason to believe that fiscal devolution to the Scottish Parliament requires major reform to the UK's institutional structures. Scotland is on its way to becoming an autonomous territory within a federal UK. Like Spain, Canada and Belgium, the UK is likely to evolve towards a 'holding together federalism', fulfilling the main requisites of federalism through unique means and structures, rather than seeing a 'federal moment' (Sala, 2014).

vi. Conclusions

The No vote in the referendum signals the start of negotiations around which taxes might reasonably be devolved to the Scottish Parliament. There is a general consensus that many of the theoretical benefits of decentralisation (as discussed in the Introduction to this PhD) are conditional on the institutional form of fiscal federalism conforming much more closely to the 'ideal type', implying much greater reliance on own-source taxation and the constitutional entrenchment of the Scottish Parliament.

The proposals for tax devolution currently on the table from Unionist parties vary significantly in terms of which taxes should be devolved, and the extent to which they address the Scottish Parliament's vertical fiscal deficit. There is perhaps a danger that Unionist parties are raising expectations beyond what is feasible for political reasons. There is however consensus that income tax is the most appropriate for devolution initially, and even the most modest proposals envisage the Scottish Parliament taking responsibility for over half of income tax revenue. The most ambitious of the tax proposals would see Scotland become one of the most fiscally autonomous regions in OECD countries, on a par with the Basque Country and Navarre in Spain, or Quebec in Canada.

The post-referendum negotiations will also consider the scope for devolution of aspects of welfare spending. This is potentially a more difficult and contentious area, given the practical challenges in disentangling the UK welfare system and differences in views as to how the desire for autonomy or self rule should be balanced by considerations of UK-wide social union and solidarity.

What is clear is that all Unionist parties support the constitutional entrenchment of the Scottish Parliament. On becoming constitutionally entrenched, and with a greater degree of fiscal autonomy, Scotland will thus evolve into a semi-autonomous state within a 'holding together' type of federal UK.

It has often been hypothesised that the Scottish Government's relatively generous grant from Westminster is implicitly intended to assuage demands for independence. It remains to be seen whether greater tax autonomy might mitigate future independence demands. As discussed in the introduction to this PhD, Bolton and Roland (1997) frame the secession decision as a trade-off between the efficiency benefits of Union (risk-sharing, no efficiency losses from tax competition) against the benefits of having a government 'closer to the people' (i.e. one that better reflects local preferences). The evidence from the campaign and, in particular, the closeness of the vote suggests that there is a strong appetite for the latter within Scotland.

The implicit assumption of the Unionist proposals is that fiscal devolution can realise both of these benefits, minimising the trade-off. Studies of strategic tax setting between governments in a federal setting suggest that the relatively high mobility of capital and labour between Scotland and rUK may mitigate the extent to which the Scottish Government feels able to exercise its powers; it is likely to face a relatively high marginal cost of public funds. Whether tax devolution mitigates the independence threat will thus depend on the extent to which the fiscal constraints for Scotland within the Union are perceived to be greater than those if it were an independent country.

Moreover, none of the Unionist party proposals envisage devolution of North Sea oil and gas revenues; given the symbolic nature of these revenues within the independence debate, this is likely to remain a source of friction in the future. Nor do any of the proposals envisage full devolution of welfare spending functions, an area in which the Scottish Government is keen to stress dissatisfaction with Westminster policy. Furthermore, if tax devolution is accompanied by reform of the mechanism for determining Scotland's block grant – either as part of a quid pro quo for more powers, or in order to operationalise tax devolution in Wales – the Scottish Government's budget may face a decline in its spending power relative to rUK. Issues around the potential reform of the Scottish block grant are discussed in the second paper in this PhD (which looks at the scope for replacement of the Barnett Formula determined block grant with a system of spending needs assessment) and in the fourth paper in this PhD (which looks at the implications of the adjustment to Scotland's Barnett determined block grant to account for devolved tax powers on the budgetary risks that the Scottish budget is exposed to).

On the other hand, globalisation arguably strengthens the case for fiscal autonomy within the Union as opposed to independence, given that globalisation both increases the importance of macro-economic stability functions of a federal government, and increases fiscal harmonisation pressures between countries (Garrett and Rodden, 2003). Thus whether tax devolution strengthens or mitigates the independence threat is likely to depend on the interaction of factors including the

long-term trend in North Sea revenues, the extent to which the Scottish Government feels able to exercise its tax powers in line with its preferences, and the extent to which Scotland feels that the notion of a 'social union' is reflected in future funding arrangements.

3. Assessing relative spending needs of devolved government: the case of healthcare spending in the UK

Rob Ball, David Eiser and David King

Abstract. The block grants allocated to the UK's devolved administrations (DAs) are not determined by any estimate of their spending needs. There are increasing calls to replace the current grant allocation mechanism with one which explicitly considers the DAs' spending needs. This paper compares two existing formulae for estimating healthcare spending needs (used by the NHS to allocate resources within England and Scotland) by applying both formulae to the DAs. It is found that these formulae provide very similar estimates of the DAs' healthcare spending needs. A similar analysis of patterns of education funding in England and Scotland reveals that resources for education are allocated to local authorities in broadly similar patterns within the two countries. Nonetheless, despite the similarity of these formulae, it is argued that significant political opposition to the use of needs assessment formulae for allocating grant to the devolved governments is likely to remain..

i. Introduction

Globally there is a trend towards decentralisation of public sector activity (Lago-Peñas et al. 2011). As the significance of devolved, or sub-central, tiers of government in public service delivery increases, the question of how to fund devolved levels of government becomes more important. A study of 23 OECD countries found that on average around a quarter of central government tax revenue was spent on inter-governmental grants (Blöchliger and King, 2006).

Among OECD countries, a variety of allocation mechanisms have been established for allocating grant to sub-central government, often involving some assessment of the spending needs of those governments. At one end of the spectrum, Australia has one of the most sophisticated spending needs assessments (Kay et al, 2005; Select Committee on the Barnett Formula, 2009), with States' expenditure needs estimated separately across 14 expenditure areas (Commonwealth Grants Commission, 2010). At the other end of the spectrum, Spain uses a comparatively simple formula to inform the allocation of resources to Autonomous Communities (Bösch, 2009).

For over 30 years, the devolved administrations (DAs) in the UK's three devolved territories (Scotland, Wales and Northern Ireland) have relied on a block grant from the UK government to finance the majority of their spending. This block grant continues as the main source of funding for the DAs even following the substantial devolution of policy and spending autonomy that began in 1997.

Recent years have seen increasing debate around how the DAs should be funded, and in particular the extent to which they should have greater tax raising powers. The Scotland Act 2012 will allow the Scottish Parliament to vary income tax rates and retain a proportion of all income tax raised in Scotland, and devolution of further fiscal autonomy seems likely in the longer term. The Commission on Devolution in Wales is exploring the case for devolution of similar tax-raising powers to the Welsh Assembly (Commission on Devolution in Wales 2012), whilst a Ministerial Working Group is exploring the possibility of devolving Corporation Tax to Northern Ireland (NI) (HM Treasury 2011).

Yet with the debate around fiscal autonomy it is easy to forget that some element of block grant will remain an important– perhaps the most important – part of the DAs funding package for the foreseeable future (this is true for Wales and NI even if Scotland were to vote for independence in 2014). There has for a long time however been substantial dissatisfaction with the way in which the DAs block grant is determined. This dissatisfaction largely centres on the fact that the determination of the block grant is not based on any estimate of the DAs' spending needs. Instead, changes in grant from one year to the next are calculated as a function of the change in spending in England, and the population share of each DA.

The question of how the block grant to the UK's DAs should be determined has recently been reignited by two high profile reports. The House of Lords Select Committee on the Barnett Formula (Select Committee on the Barnett Formula 2009) and the Holtham Report (Independent Commission on Funding and Finance in Wales, 2010) recommended that the current approach to determining the block grant should be replaced by some form of expenditure needs assessment, which would explicitly allocate grant to the DAs based on estimates of their relative spending needs. Other authors however have highlighted the difficulties in estimating spending needs, given that the question of what constitutes a need is inherently normative (Midwinter, 2002).

However, it is also true that spending needs assessments are used regularly to estimate the spending needs of local authorities, schools and health trusts in the UK (Smith, 2006), with grant allocations often being determined on the basis of these estimates. This raises a question about how some of these existing methods of spending needs assessment might estimate the spending needs of the UK's devolved territories: would different formulae produce widely different estimates of the UK territories' spending needs (thereby justifying the views of the needs assessments sceptics), or would different formulae produce similar estimates of the UK territories' spending needs?

This paper compares two formulae for assessing the healthcare spending needs of territories. One formula is used by the National Health Service (NHS) in England to allocate resources to Primary Care Trusts (PCTs); the other formula is used by the NHS in Scotland to allocate resources to Health Boards in Scotland (throughout the remainder of this paper, the term PCT is used to refer to Primary Care Trusts in England, Health Boards in Scotland and Wales, and Health and Social Care Trusts in NI). The paper compares the two formulae by applying each in turn to the UK territories.

The two formulae have been developed separately over several decades, overseen by different governments with varying political priorities. Thus it might be expected that the two formulae would provide quite different estimates of the spending needs of the UK territories, in which case there would seem to be limited scope for using them as part of a UK-wide system of estimating spending need. On the other hand, if the two formulae provide similar estimates of the UK DAs' relative spending needs, the scope for developing an approach to estimating the spending needs of the DAs looks more possible. This is not to say that a Barnett replacement should use either the English or Scottish formulae for health or education, but rather to identify where disagreements around needs assessment are likely to arise, and to help inform the legitimate starting points for the DAs in any negotiations with the UK Government around what might constitute a reasonable suite of needs indicators, and where the limits of a 'fair' block grant settlement may lie.

The analysis focuses on health services for two reasons. First, the National Health Services in both England and Scotland have developed sophisticated mechanisms for measuring the healthcare spending needs of territorially based health trusts, and these formulae form the basis of our analysis. Second, healthcare is an important element of devolved policy, accounting for around one third of the budget of each DA. Healthcare is of course just one element of DA spending, and thus it could be argued that developing an estimate of the DA's healthcare spending needs alone would not necessarily help in developing an alternative to the Barnett Formula (given that the Barnett Formula is used to determine all elements of the DA's block grant). The objective of this paper however is to consider in the first instance whether there are similarities in the English and Scottish approaches to estimating healthcare spending needs, and following this to consider what the implications are for the development of alternative funding mechanisms for the UK's DAs in the future.

The remainder of the paper is structured as follows. Section 2 describes the political economy of the current funding arrangements for the UK DAs, and sets out issues relating to the territorial differences in healthcare spending. Section 3 provides an overview of the healthcare needs assessment formulae that are used to allocate healthcare resources within England and Scotland, while Section 4 describes in more detail how these formulae were applied to the UK territories. Section 5 presents the results, focusing on how the two formulae assess the health spending needs of the UK territories. Section 6 extends the analysis to school age education services, and assesses the similarity of the English and Scottish health and education needs formula empirically. Section 7 concludes.

To preview our results, we find that the existing English and Scottish NHS healthcare allocation formulae do both provide similar estimates of the devolved territories' healthcare spending needs, despite having been developed from different data in different socio-political contexts. We argue that these findings justify the current interest in replacing the Barnett Formula with a formula for estimating spending needs, but that there remain a number of political challenges to overcome if this is to be achieved.

ii. The political economy of funding the devolved territories

The spending of the DAs is largely financed through a block grant from the UK government (the DAs have some ability to raise funding indirectly through council tax and business rates, although this position is evolving). This block grant is determined as a function of two things: the baseline and an increment. The baseline is simply the grant in the previous year, whilst the increment is determined by the Barnett Formula. The Barnett Formula determines the change to the DAs' budget based on

changes in spending on comparable[5] services in England, and the population share of each DA (HM Treasury, 2010a). For example, if the UK government announces a £100m increase in health spending, 99% of all health spending is devolved, and Scotland's population is 10% of England's, then the Scottish Government would see an increase in its budget of £9.9 million.

The Barnett formula has been criticised because it makes no attempt to estimate the spending needs (other than population shares) of the devolved territories (McLean et al, 2008). This has led to the formula being claimed variously to be overly generous to some territories (in particular Scotland but also NI) and less generous to others (Wales), (Morgan, 2001; McLean and McMillan, 2003; MacKay and Williams, 2005).

Given this criticism, a case has been made to replace the Barnett formula with some form of expenditure needs assessment (Independent Commission on Funding and Finance for Wales, 2010), and specifically one which ensures '*those parts of the UK which have a greater need receive more public funds to help them pay for the additional levels of public services they require as a result*' (Select Committee on the Barnett Formula, 2009, Paragraph 81).

Criticism of the Barnett formula's insensitivity to spending needs is not new. A legitimate question to ask is why Barnett has remained the preferred mechanism for allocating grant to the UK territories, especially given the apparent disadvantage to citizens of England and Wales made by the current pattern of Barnett-based allocations.

The continued reliance on Barnett is usually explained in relation to issues of political economy. On the one hand, some commentators have characterised the Barnett formula as part of a more nuanced 'formula plus influence' system of funding which has advantages for both the central and devolved governments (Christie and Swales, 2010). According to these authors, the grant allocations to the devolved territories reflect 'the greater cultural and geographic distance of these regions from the centre and the greater political coherence and the more powerful bargaining power that the peripheral [i.e. devolved] regions held' (Christie and Swales, 2010). Modelling this as a sequential game, Hallwood and MacDonald (2009) make a similar point about the role of devolved territories' threat to secede from the Union in maintaining a territories grant at a level above that which the territory might receive should grant be based on 'need' alone, an argument that was similarly made

---

[5] Comparability factors are set out by HM Treasury for English UK Government department. The comparability factors measure the extent to which the services provided by each department are comparable to the services provided by the Scottish Government in Scotland. In effect the comparability factor measures the extent to which the UK department's services have been devolved. The comparability factor for education is 100% (education is a wholly devolved matter) whereas for the Department for Work and Pensions the comparability factor is 1.4% (social security is essentially a matter for the UK Government).

by McLean and McMillan (2003) in commenting on the formula's apparent generosity to Scotland and NI relative to Wales.

Other commentators highlight the normative dimension of spending needs assessment, and the potential difficulties in politicians reaching consensus on what a needs-based funding formula for the UK territories would look like in respect of which indicators might be used, and how they might be weighted (Midwinter, 2002).

Thus although there is widespread recognition of the major limitation of Barnett in relation to its insensitivity to spending need, the formula's longevity seems due to its more nuanced advantages as the 'codification of a bargaining outcome' (Smith, 2003), as well as the difficulties of agreeing how a needs formula might be developed. In relation to the latter issue, both the House of Lords and Holtham Commission reports made suggestions for the type of indicators that might be included in a simple but transparent needs assessment formula, arguing that the implementation of such a formula could be overseen by an independent commission, similar to the operation of the funding of Australian states. In relation to the former issue, there is an argument that interest in replacing Barnett may become stronger following devolution of some level of tax autonomy to Scotland and the other devolved nations, combined with increasing divergence in political priorities across the UK territories following devolution (Jeffery 2012); and the emergence of a stronger English political identity that questions the capacity of the current UK-level political institutions to pursue and defend English interests (Lodge et al. 2012).

Turning to healthcare specifically, per capita spending on health services varies substantially across the UK territories (Table 3.1). As an index relative to English per capita spending on health services, per capita spending on health services in Scotland is almost 14% higher, while it is 8% higher in Wales and 9% higher in NI.

Table 3.1: Spending on healthcare in UK territories

	England	Wales	Scotland	NI
Healthcare spending per capita	£1,643	£1,769	£1,871	£1,783
Index of per capita health spending (England=1)	1.000	1.077	1.139	1.085
Spending on health as % all devolved spending	-	34.3%	33.0%	34.2%

Source: Public Expenditure Statistical Analyses (HM Treasury 2012)
Note: Figures on spend per capita are averages over 2005/6 – 2010/11 in 2012 prices. Devolved spending for Wales and Scotland is defined as spending on health, education, local government, transport, environment, housing, culture and tourism, and policing, but excluding defence, foreign affairs and welfare. Devolved spending for NI is defined the same way with the exception of policing, which is considered non-devolved. Calculation of an equivalent figure for English spending as a percentage of devolved spending is problematic given that the UK Government has never published a figure for 'English equivalent expenditure' (Heald, 2014).

The differences in per capita spend on health services are largely a function of differences in Barnett formula-derived grant allocations to the DAs. Indeed, spending on health makes up a similar proportion of total DA spending in each territory, accounting for around one third of devolved spending (Table 3.1).

As discussed in the preceding section, these Barnett-Formula derived grant allocations have nothing to do with any estimate of spending needs for health (or any other service). Thus although there is some evidence that the UK's devolved territories (and Scotland in particular) may have higher health spending needs than England on some measures (Hanlon et al., 1999; McLaren et al., 2010; NAO, 2012), the spending patterns across the UK territories do not reflect any type of estimate of healthcare spending need. If the Barnett allocations do coincide with any estimate of health spending need, this would be by accident, rather than design.

If they so chose, the DAs could increase spending on healthcare, but this would necessarily come at the expense of reductions in spending elsewhere within the DAs portfolio of devolved services. Given the significance of healthcare spending in DA budgets, the scope for varying healthcare spend without negatively impacting on other areas of the budget is extremely limited.

iii. Needs-based funding models in the NHS

The way in which grant is allocated to the UK's devolved territories, which makes no attempt to estimate those territories' spending need, contrasts markedly from the way in which resources are allocated within the territories to Primary Care Trusts, where sophisticated needs estimates models are used to estimate spending needs. The NHS in England has used an expenditure needs assessment formula to calculate revenue allocations to PCTs since 1980. The NHS in Scotland has used its own needs assessment formula for determining revenue allocations to Health Boards since

1979. Wales and NI also use a formula funding approach to allocate resources to Health Boards and Health and Social Care Trusts respectively, although the detail of these formulae is not considered in this paper.

The English and Scottish formulae use a weighted capitation approach to determine the level of resources allocated to PCTs (Department of Health, 2009a; NHS Scotland Resource Allocation Committee, 2007). The principle of the weighted capitation approach is to calculate shares of the overall budget that should be allocated to each PCT, in order that each PCT can provide similar services, given the characteristics of the population in its area. This is done by adjusting the population in each PCT to take account of both the health needs of its population and any additional costs associated with delivering services in that area.

The health needs of an individual or a given population cannot be observed directly. Instead the capitation approaches to funding allocation used in both England and Scotland rely on statistical modelling to examine the utilisation of health services at a small area level as a function of population characteristics of each area. The statistical models also include 'supply' variables to take account of the fact that utilisation will tend to be higher where the availability of healthcare services is greater.

'Needs' formulae are typically estimated using regression equations of the form:

$$E_i = \beta_1 X_i + \beta_2 S_i + \epsilon_i$$

Where E_i is the expenditure on a particular element of healthcare (e.g. GP prescribing costs) in small area i as a function of a vector of socio-economic characteristics X_i and a vector of characteristics S_i designed to control for supply.

The basic assumption underlying these resource allocation approaches is that healthcare 'need' can be empirically determined by modelling the relationship between utilisation and socio-economic variables whilst correcting for supply factors. Based on the resulting empirically estimated models, data on level of such socio-economic variables in small areas can then be used to allocate resources in a way that reflects needs.

The healthcare allocation formulae have been criticised on a variety of grounds. Despite the sophistication of some of the modelling involved, the use of local utilisation data as a proxy for need clearly runs the risk of reflecting and reinforcing existing inequalities in the relationship between underlying healthcare needs and the resources allocated to address them. A reliance on utilisation data presupposes that historical patterns of service uptake by different care groups are appropriate, but it ignores the possibility that some groups may underutilise healthcare services. There are also

difficulties in reliably controlling for supply in influencing current patterns of utilisation. In short, there are marked difficulties in modelling need where need, utilisation and supply are so inextricably linked (Asthana and Gibson, 2008).

The models have also been criticised on the basis of the data and statistical methods employed leading to assertions that the formulae produce results which are 'arbitrary and unfair' (Galbraith and Stone, 2011).

A further criticism relates to the question of what type of healthcare need the models should be allocate on the basis of, and the implications for equity criteria. The Scottish and English allocation models have traditionally been motivated by the objective to promote 'equal opportunity of access for equal needs'. But if the objective were instead to promote an 'equal opportunity to be healthy', funding would need to be targeted so as to reduce the health gap between the most advantaged and least advantaged groups. As we will see later, recent iterations of the English formula have attempted to build-in elements designed to capture an objective to reduce healthcare inequalities.

In summary, the formulae on which the allocation models are based capture the existing average responses of NHS Trusts to 'need' in their areas, including any equity criteria that those Trusts apply. Thus the capitation formulae technically offer comparable public sector organisations the opportunity to deliver some average level of service, assuming average responses to social and economic circumstances, and an average level of efficiency. Whether the resulting allocations reliably assess spending 'need', including any specific equity criteria, remains a moot point.

Despite these criticisms, utilisation-based capitation formulae remain the preferred mechanism of the NHS in England, Scotland, Wales and NI to allocate healthcare resources.

The English and Scottish allocation formulae each consist of a number of blocks, or elements, representing different aspects of healthcare delivery (for example, Acute Care, Maternity Care, Care of the Elderly, Mental Health Care, and so on). For each of these blocks, both the English and Scottish capitation formulae adjust each PCT's population to take account of three broad factors that are seen to influence spending need:

- The *age* structure of the population, recognising that need for healthcare varies according to age, and in particular, the young and elderly, whose populations are not evenly distributed across the country.
- *Additional needs*, i.e. the fact that populations of the same age distribution display different levels of need for healthcare. Higher levels of underlying morbidity (illness) and socio-economic deprivation are likely to increase demand for healthcare.

- *Cost*, i.e. the fact that the cost of providing healthcare may be higher in remote rural areas because of diseconomies of scale, or in large urban areas where market forces result in higher factor costs.

To illustrate, consider the derivation of the index for Acute Care in the Scottish formula (Acute accounts for 52% of resources allocated through the Scottish formula in total). PCTs are allocated a sum for every individual within each of 20 age categories. For example, a PCT is allocated £989 for each male aged 0-1, and £1,749 for each male aged 75-79. Each PCT's *additional need* for resources related to morbidity and life circumstances is then calculated based on standardised rates of mortality and long-term limiting illness in the area. Finally, the *excess costs* are derived based on the proportion of the PCT population living in one of ten categories of rurality.

The English formula's approach to assessing spending needs for Acute Care (which accounts for 51% of resources allocated through the formula) is as follows. Age and additional needs are assessed simultaneously using a 'one-stage' approach. This allows the relationship between age and additional need to vary across 18 different age bands, with 12 separate needs indicators entering the equation. Unavoidable excess costs are calculated by considering differences in wage, land and building costs relative to the national average.

The capitation formulae are used to inform the allocation of significant resources. In the 2009/10 financial year, the weighted capitation formula informed the allocation of some £80 billion to NHS Trusts in England (Department of Health, 2009b), representing over 80% of health revenue expenditure in the country (HM Treasury, 2010b). Scotland's capitation formula informs the allocation of some £7 billion, around 85% of all funding to Health Boards and 67% of all NHS Scotland expenditure (NHS Scotland Resource Allocation Committee, 2007). Both the Scottish and English formulae are updated on a regular basis, overseen by independent committees in each country.

iv. Approach and data

This section describes in further detail the structure of the English and Scottish healthcare allocation formulae, and how the two formulae were applied to the territories of the UK. It is worth re-iterating that these formulae have not been developed by the authors; the formulae have been developed over many years by the respective health services in the two countries, and the aim of this paper is to apply each formulae to all PCTs in the UK.

Although the general structure of the English and Scottish formulae is the same, there are some differences between the formulae at a more detailed level. Firstly, there are some differences between the formulae in terms of the components of health covered. The English formula is used to

allocate resources to three broad components of healthcare: Hospital and Community Health Services (HCHS); Prescribing (the drugs bill); and Primary Medical Services (PMS, i.e. GP surgeries and out-of-hours services). The weights attached to each of these three elements are 76%, 12% and 11% respectively. The Scottish formula also allocates resources for HCHS (weighted 85%) and Prescribing (weighted 15%), but not PMS (in Scotland, resources for PMS are allocated outwith the formula).

Secondly, there are differences between the two formulae in terms of the specific indicators used and the weights attached (Table 3.2 shows the indicators used). Further detail on how these indicators are weighted and combined to estimate the relative spending needs of each PCT are set out in guidance documents (NHS Scotland Resource Allocation Committee, 2007 for detail of the Scottish formula; and Department of Health, 2009a for details of the English formula). The precise structure of the two formulae is too complex to outline here, but readers who wish to understand the mechanics of the formulae in more detail can consult the respective guidance documents referred to above, or Ball et al., (2012) which sets out details of how the Scottish formula was applied, and Ball et al., (2011) which sets out details of how the English formula was applied.

An interesting difference between the two formulae is in relation to cost. In assessing the costs associated with delivering healthcare in different geographical areas, the English formula relies largely on indicators of factor costs, specifically wages, land values and premises costs. The Scottish formula in contrast focuses more on the effects of sparsity on healthcare delivery costs. The implications of this difference for the pattern of resources made by the two formulae is discussed in the results section.

A further difference between the formulae is that the English formula includes an adjustment for health inequalities (HI), whereas the Scottish formula does not. Before the introduction of HI to the English formula, the objective of the English formula was to allocate resources so as to enable PCTs to commission similar levels of healthcare for populations with similar healthcare needs. The HI element introduces an additional objective to the formula, specifically to allocate resources so as to 'contribute to the reduction of avoidable inequalities in health'. The effect of the health inequalities element on the allocation of resources can be significant as it implies that resources should be distributed not only to reflect the existing burden of sickness but also targeted to reduce the health gap between the most and least advantaged groups. In Scotland, health inequalities are addressed outwith the formula, via specific grants. The implication of these different approaches to addressing health inequalities are discussed in the results section.

In order to apply the English and Scottish healthcare capitation formulae across all UK territories, the data described in Table 3.2 was collated for each PCT in England, Wales, Scotland and NI. This

resulted in data being collated for some 174 PCTs in total[6]. The data was applied to the two formulae to determine a relative expenditure needs score for each PCT, and by aggregation, each DA.

---

[6] Data was collated for 148 PCTs in England, 14 Health Boards in Scotland; seven Health Boards in Wales and five Health and Social Care Trusts in Northern Ireland. There are in fact 151 PCTs in England, but due to boundary issues the three Birmingham PCTs were amalgamated into one and the two Cheshire PCTs were also amalgamated into one, for the purpose of this analysis.

Table 3.2: Summary of indicators used by the English and Scottish healthcare allocation formulae

	English formula	Scottish formula
HCHS	**Age indicators**	
	• GP-registered population in 18 age groups, 2009	• GP-registered population in 20 age groups, by male and female, 2009
	Additional need indicators	
	• Age-specific death rates within each of 18 age-bands, 2001-5 • The standardised proportion of the population with no qualifications, 2001 • The proportion of young people not staying in education, 2006 • Standardised % population with a limiting long-term illness, 2001 • Proportion of Pension Credit claimants, 2005 • Proportion of births that are low birthweight, 2001-5 • An indicator to reflect income deprivation affecting children • Proportion of under 16 population claiming Disability Living Allowance, 2005 • Proportion of young people claiming New Deal, 2004 • Proportion of the total population claiming Disability Living Allowance, 2005 • % of the population claiming Incapacity Benefit or Severe Disablement Allowance, 2005 • Proportion of the over 60 population claiming DLA, 2005 • Number of births (2006) • Proportion of births that are low birthweight (2009) • Mean house price (2003-9) • Under 65 comparative mortality factor, 2009 • Proportion aged 60+ claiming Income Support, 2009 • Psycho-social morbidity index (a function of ethnic minority population, a standardised illness ratio, and an index of income deprivation) • Prevalence of HIV/ AIDS diagnoses (2009) • Disability free life expectancy (a function of illness and death rates), 2001/9	• All cause standardized mortality rate, age 75 and under (2001-5) • Limiting long-term illness rate, age-sex standardized (2001) • % of 16+ population claiming Severe Disablement Allowance (2004) • Proportion of households that are one person (2001) • Proportion of households that are social rented (2001) • Mean house price (2003-9) • Birth rate per female population aged 15-44 (2009) • Urban rural supplementary variable (2004)
	Cost indicators	
	• Standardised spatial wage differentials, 2007-9 • Land and buildings cost index, 2004	• Proportion of population living in each of ten geographical categories (from Primary city to very remote island), 2001 •
Prescribing	**Age indicators**	
	• GP-registered population in 9 age groups, males and females, 2009	• GP-registered population in 9 age groups, males and females, 2009
	Additional need indicators	
	• Percentage population with long-term limiting illness • Proportion Disability Living Allowance Claimants • Low Income Scheme index • Proportion births that are low birthweight	• All cause standardized mortality rate, age 75 and under (2001-5) • Limiting long-term illness rate, age-sex standardized (2001)

English formula	Scottish formula
• Disability free life expectancy, 2001/9	

			Scottish formula
PMS	Age indicators	n/a	
	• Population in 7 age groups, male and female, 2009		
	Additional need indicators		
	• Age-sex standardised proportion of population with long-term limiting illness (2001)		
	• Under 65 standardised mortality ratio, 2009		
	• Disability free life expectancy, 2001/9		
	Cost indicators		
	• Standardised spatial wage differentials, 2007-9		
	• Land and buildings cost index, 2004		
	• Index of multiple deprivation, 2010		

v. Results

The English and Scottish NHS resource allocation formulae's estimates of the healthcare spending needs of the devolved territories are shown in Table 3.3. The results are expressed as an index relative to English per capita healthcare expenditure need of 1. Thus an index of 1.1 for a territory implies spending need of 10% above England's, whilst a score of 0.9 for a territory implies spending need 10% below England's. The results in Table 3.3 combine the results from the different components of the formulae (i.e. HCHS, Prescribing and PMS) into a unified estimate of healthcare spending needs.

The results for the English formula exclude the Health Inequalities element. This is so that the two formulae are compared on a like-for-like basis, i.e. where the objective of both formulae is to allocate resources to ensure 'equal access to healthcare for those in equal need'. The implications of including the Health Inequalities element of the English formula (which as noted above introduces a second objective of reducing avoidable health inequalities) are discussed subsequently.

Table 3.3: Estimates of the healthcare spending needs of the UK devolved territories: comparing the results of the English and Scottish NHS allocation formulae

	English NHS formula				Scottish NHS formula			
	England	Wales	Scotland	NI	England	Wales	Scotland	NI
Age	1.000	1.044	1.009	0.932	1.000	1.035	0.997	0.937
Additional need	1.000	1.086	1.086	1.134	1.000	1.055	1.110	1.104
Cost	1.000	0.968	0.977	0.969	1.000	1.002	1.007	1.006
Total need	1.000	1.097	1.072	1.024	1.000	1.095	1.114	1.039

The two formulae provide relatively similar estimates of healthcare spending need (bottom row). Wales' per capita healthcare expenditure need is estimated almost identically by the two formulae (9.5% above England according to the English formula and 9.7% above England according to the Scottish formula). NI's per capita need ranges from 2.4% to 3.9% higher than England's on the English and Scottish formulae respectively. For Scotland there is somewhat more variation between the two formulae, with needs ranging from 7-11% above England's according to the two formulae. (These assessments of relative spending needs are not wildly out of kilter with the patterns of healthcare expenditure shown in Table 3.1, but this is largely by accident rather than design).

Reasons for the divergence between the formulae results can be discerned by considering how the formulae assess the devolved territories for the three core elements of the formulae: age, additional needs, and costs.

Healthcare expenditure needs increase with age and therefore Wales, which has a slightly older population, is allocated more resources per capita under both the Scottish and English formulae (the English formula allocates Wales an additional 4.5% resources per capita on account of the age distribution of its population, whilst the Scottish formula allocates Wales an additional 3.5% per capita; the differences between the two formulae being due to slightly different weights attached to different age groups in the formulae' estimates of need). Conversely NI which has a younger population is allocated fewer resources per capita than England by both formulae. Scotland has a similar age distribution as England and thus is allocated very similar resources (around 1% more than England by the English formula and 0.3% less than England by the Scottish formula).

The additional needs index covers a PCT's need for healthcare as a result of underlying morbidity or social deprivation. The second line in Table 3.3 reveals that all three devolved territories have significantly higher 'additional' needs than England according to both formulae. However, the English formula appears more generous to Wales and NI than the Scottish formula, but the Scottish formula is more generous to Scotland than the English formula is.

The difference between the two formulae in how they score the territories relates to the indicators that the formulae use. In the Scottish formula, two indicators predominate – the standardised under- 75 mortality rate and the standardised long-term limiting illness rate explain around three quarters of the variation in HCHS additional need scores. The English formula also places significant weight on death rates by age band, but considers a wider range of socio-economic measures of deprivation (see Table 3.2).

The English formula thus includes several indicators of labour-market and wider socio-economic deprivation that tend to favour the most deprived inner-city areas. This hypothesis is confirmed by further analysis, not reported here but available from the authors, which compares the 'additional need' scores of all UK PCTs individually according to the two formulae. This analysis shows that, relative to the English formula, the Scottish formula allocates relatively fewer resources for 'additional needs' to London PCTs in particular, but also to other deprived urban areas such as Manchester and Knowsley. In contrast, the Scottish formula allocates relatively more resources to less urban areas.

The cost index for the devolved territories, shown in the third line of Table 3.3, shows minor differences between the formulae in how they assess the excess costs associated with delivering healthcare. According to the English formula, the DAs face lower per capita healthcare delivery costs than England. The Scottish formula on the other hand indicates that the DAs face marginally greater per capita delivery costs than England. The reason for this difference is that the English and Scottish

costs assessments emphasise different factors. The English formula takes into account labour and facilities costs in different parts of the country. This tends to allocate more resources to London and the southeast where such costs are higher. The effect of the Scottish cost index in contrast is largely to compensate rurally based PCTs for the higher costs of delivery in remote and sparsely populated areas.

So in terms of cost, both the English and Scottish formulae do tend to shift resources between PCTs; but at national level these adjustments tend to make relatively little difference. This is because a relatively small proportion of the population of each country are effected by such adjustments (this is particularly true of the Scottish cost element, which allocates significant additional per capita resources to very remote areas, but these adjustments can be accommodated by very minor reductions in per capita allocations to other parts of the country given how few people live in very remote areas).

The fact that the Scottish formula adjusts more resources to rural areas but the English formula adjusts more resources to urban areas is neither surprising nor alarming. Both formulae are based on evidence as to how the costs of healthcare delivery vary across different areas – the Scottish formula does not allocate many additional resources to high-cost urban areas because it has few of these areas in its territory; likewise the English formula does not adjust resources particularly to remote island PCTs because there are no such PCTs in England on which evidence for the costs associated with such areas could be assessed (the English formula makes no special allowance for the Isles of Scilly, doubtless because the Isles of Scilly represent just a small proportion of the population residing in the Cornwall PCT). Thus the differences between the two formulae seem unlikely to be due to fundamental normative differences about whether for example, the costs associated with higher labour costs in urban areas should be compensated for, and more to do with the sample of geographical areas on which the cost formulae have been derived.

The two formulae can be compared further by considering how they would each allocate resources across the regions of England as well as the territories of the UK. Figure 3.1 compares the per capita healthcare allocations made by the English formula (x-axis) with the per capita healthcare allocations made by the Scottish formula (y-axis). The dotted 45° line represents equality between the two formulae – if a region is located above this line its needs are estimated to be higher by the Scottish formula; if a region is located below the line its needs are estimated to be higher by the English formula.

Figure 3.1: Comparing the English and Scottish allocation formulae's assessment of the spending needs of UK territories and regions
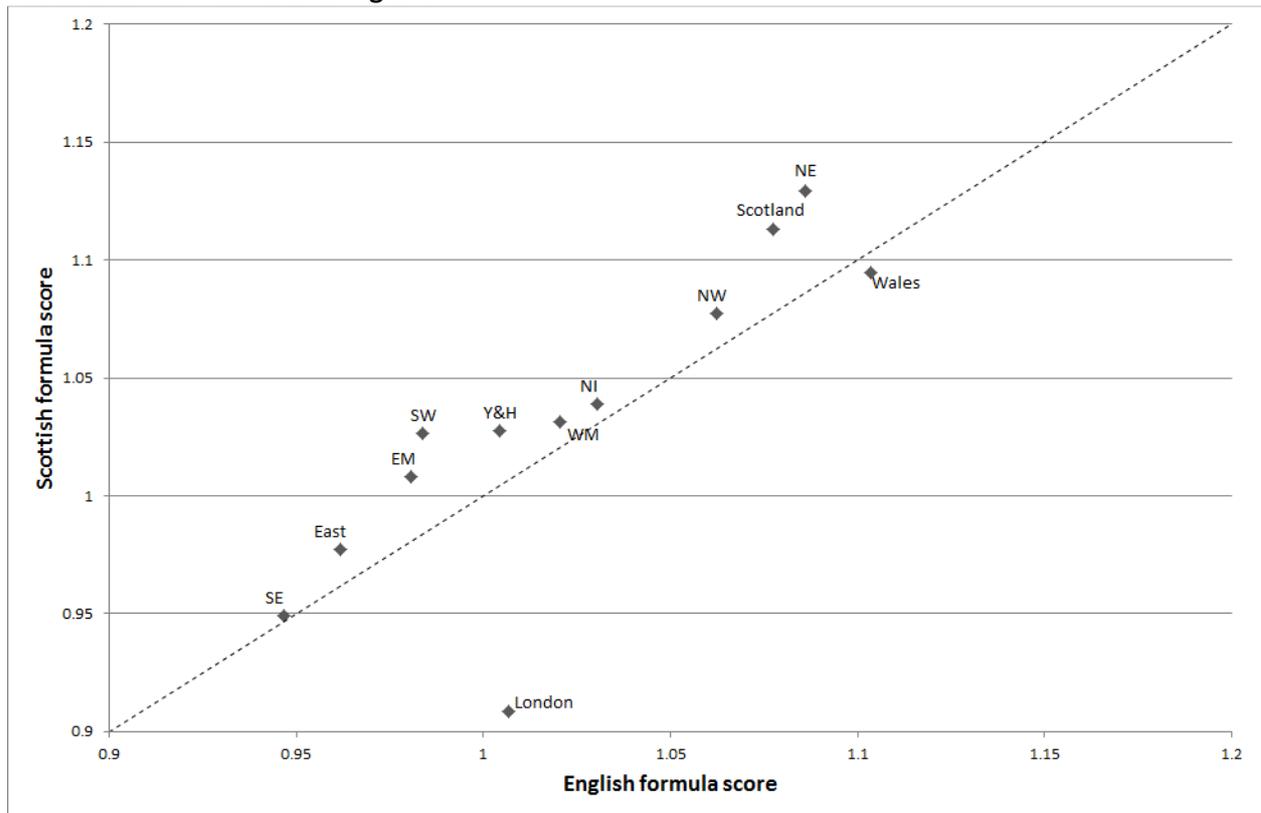
For most regions, the two formulae provide very similar estimates of spending needs. For the South East, East, East Midlands, Yorkshire and the Humber, West Midlands, Northern Ireland, the North West and Wales, the two formulae estimate needs within 3 percentage points. For Scotland and the North East of England, the Scottish formula allocates more resources than does the English formula, and this is mainly due to the fact that both these areas have relatively high rates of mortality and self-assessed morbidity, indicators which are weighted fairly heavily by the Scottish formula in estimating 'additional needs' for healthcare (the Scottish formula also does not penalise these areas so much for the fact that they have relatively lower labour and other factor costs). The healthcare spending needs of the South West region are estimated to be about five percentage points higher by the Scottish formula compared to the English formula. This results from a combination of factors, including: that the region has a relatively elderly population which the Scottish formula appears to compensate slightly more than the English; that the South West is allocated fewer 'additional needs' resources by the English formula because it performs relatively well on measures of socio-economic disadvantage; and that the region benefits to a small extent by the Scottish formula's allocation of resources to rural areas on account of the additional costs incurred in delivering healthcare in those places.

The main outlier is London. The Scottish formula estimates London's per capita needs to be 9% lower per capita than the English average, whereas the English formula estimates London's per capita needs to be 1% higher than the English average. There are two main drivers of this difference, both alluded to already. First, the English formula estimates that London has higher 'additional needs' than the Scottish formula does, because the English formula relies more heavily on indicators of labour market disadvantage and ethnic minority population, whereas the Scottish formula relies more heavily on mortality and indicators of self-assessed morbidity as proxies for 'additional need'. Second, the cost element of the English formula allocates resources to areas which experience high labour and other factor costs, whereas the extent to which the Scottish formula compensates for these types of costs is more limited. If the English formula results are calculated *excluding* the cost element, then the two formulae's estimate of the relative spending needs of the UK regions are closer (Figure 3.2).

Figure 3.2: Comparing the English and Scottish allocation formulae's assessment of the spending needs of UK territories and regions, excluding ACA from the English formula
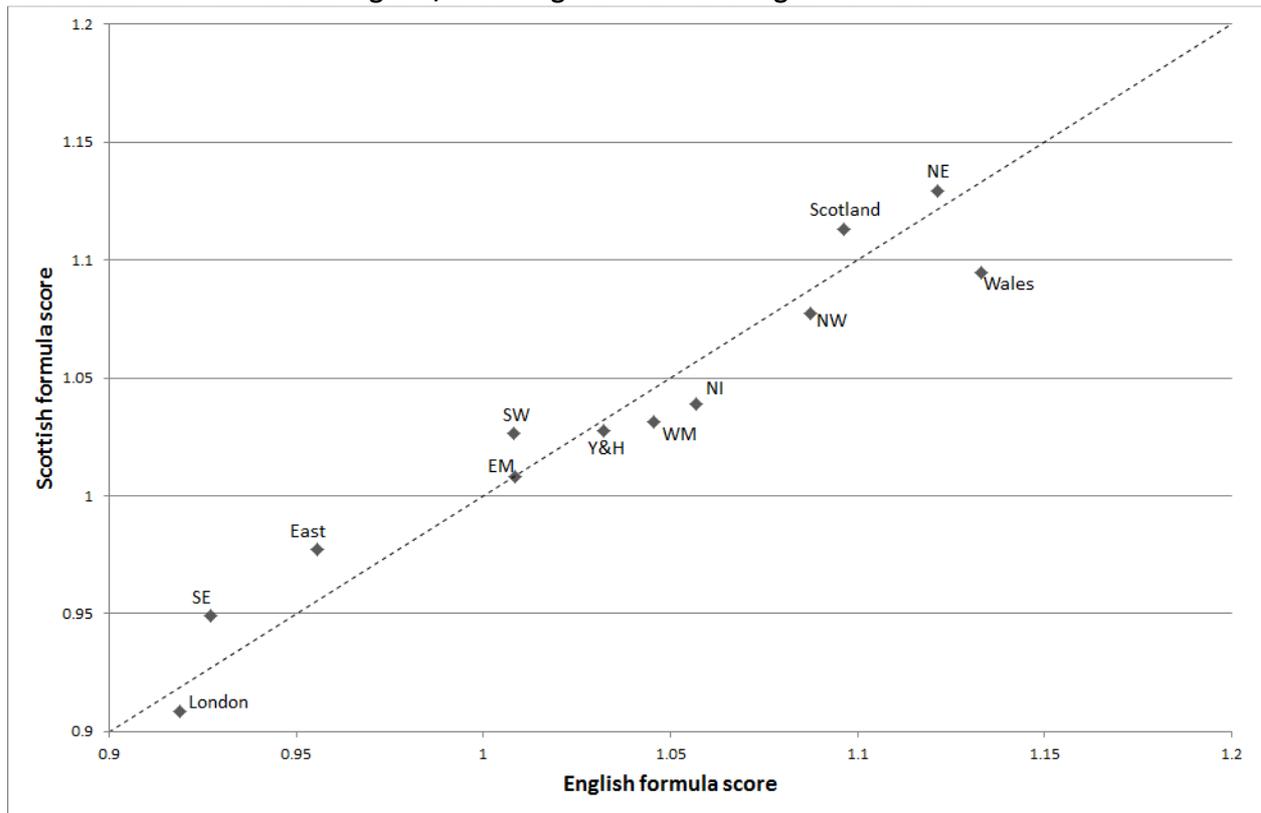
The results so far discussed assume that the objective of both formulae is to enable PCTs to commission similar levels of healthcare for populations with similar healthcare needs. As noted previously however, the English formula also contains an objective to reduce avoidable health inequalities. The health inequalities element of the English formula allocates resources towards areas that have lower Disability Free Life Expectancy (DFLE), an indicator which combines information on mortality and morbidity. The weight attached to the DFLE indicator is determined by Ministers (unlike the other elements of the formula, the weight cannot be determined through analysis of existing healthcare utilisation). It was weighted 15% in 2009/10 and 2010/11 before being reduced to 10% in 2011/12. Within England, the distributional effects of the health inequalities element are significant across PCTs, with the effect of the health inequalities element being to direct funding from southern and eastern England to the north of England and inner London (National Audit Office, 2011).

Inclusion of the health inequalities element would also have significant distributional consequences at a UK level, and specifically to additional resources being allocated to the devolved territories, relative to England. After incorporating the HI element at 15% into the English formula, the estimated relative per capita spending needs of Wales, Scotland and NI increases to 16%, 12% and

9% higher than England's, respectively (Table 3.4). Inclusion of the HI indicator in the Scottish formula (which is shown in Table 3.4 for comparison purposes only – the Scottish formula does not in practice adjust for Health Inequalities) results in similar magnitude increases in the overall scores for each devolved territory.

Table 3.4: Estimates of the healthcare spending needs of the UK devolved territories, including an adjustment for Health Inequalities

| | English NHS formula | | | | Scottish NHS formula | | | |
	England	Wales	Scotland	NI	England	Wales	Scotland	NI
Health inequalities score	1.000	1.607	1.443	1.585	1.000	1.607	1.443	1.585
Total need	1.000	1.155	1.117	1.091	1.000	1.161	1.156	1.11

There has been criticism of what is seen by some to be an arbitrary weight attached to the HI indicator, and debate around how the indicator should be weighted (House Of Commons Health Committee, 2009). This is an illustration of the normative dimension to formula funding; that there is no 'ineluctable scientific logic' to the choice of one formula over another, and policy judgements that have important issues for resource allocation are required at all stages of formula funding (Midwinter, 2001; McLean and MacMillan, 2003; Smith, 2003). This leads to much lobbying (on the part of grantees) for a particular outcome, while the government has the ability to tweak the pattern of allocations to favour their electoral ambitions. There are ample examples of such manipulation in relation to the funding formulae for local government. John and Ward (2001) for example showed that local government resources were targeted at swing districts and conservative-flagship districts during the Conservatives time in office, while Hilber et al., (2011) demonstrate that resources were targeted at Labour swing districts during Labour's 13-year period in office. Allocations to PCTs may be less politically motivated (given that there are no elections of members to PCTs), but possibly only marginally so – Carr-Hill et al (1997) describe how Conservative and Labour governments manipulated the healthcare formula to distribute relatively less (then more, respectively) resources to health trusts in inner cities.

Despite these issues that are inherent to formula funding, they do not detract from the issue that this paper seeks to address, namely how similarly the English and Scottish formulae allocate resources, and how these allocations may or may not differ from actual allocations under Barnett.

vi. Further analysis of the similarity between the health allocation formulae, and an extension to school education

To further test the hypothesis that the English and Scottish assessments of relative health spending needs differ significantly, each formula was applied in turn to all 176 health boards in the UK, and then aggregated to each of the nine English regions and three UK territories. To compare the results of the two formulae, the two formulae were used to give each of these areas a need score which relates spending need in that area to the average English spending need which is scaled to 1. For example, a score of 1.1 indicates that an area has a spending need 10% above the English average.

To test the similarity of the formulae, the relative need scores for each area according to the Scottish formula were regressed on the equivalent need scores according to the English formula, i.e.:

$$ESN_i = \alpha_0 + \alpha 1 SSN_{i} + e_i \quad (1)$$

where ESN is the per capita spending needs of area i according to the English formula, and SSN is the per capita spending needs of area i according to the Scottish formula. If the two formulae actually assessed relative needs for each area identically, then both the coefficient α_1 and R^2 would equal 1.

Another major area of DA spending is education. The Scottish Government uses a formula to allocate school education resources to local authorities (LAs) in Scotland; likewise the Westminster government uses a formula to allocate resources for school education to LAs in England (King and Eiser, 2017). As with health, there are differences between the Scottish and English education systems. These relate not only to different exam and curriculum frameworks and age of transition from primary to secondary school, but also differences in the use of performance frameworks and league tables, and the degree of school autonomy in curriculum implementation. Moreover, the allocation formulae themselves measure and weight things such as pupil disadvantage and rurality in different ways. The interesting question that arises is whether these differences in policy and allocation formulae result in different patterns of allocations across the territories. As with health, this question is examined by applying the English and Scottish education formulae in turn to the UK's local education authorities, aggregating the results to the level of the English regions and UK devolved territories, and comparing the formulae using equation (1).

Table 3.5 shows the results of applying equation (1) to both the health and education formulae. Column A considers the similarity of the education formulae. We cannot reject the hypothesis that the coefficient α_1 is equal to unity, which indicates that, on average, the two formulae assess the relative needs of the 12 regions for education spending in the same way. The R^2 of 0.70 however suggests some variation around this average relationship. The area where the formulae disagree

most is London. Its score is 1.17 on the English formula but only 1.02 on the Scottish formula. This difference arises chiefly because the English formula includes an 'Area Cost Adjustment' (ACA) which adjusts assessed spending needs in areas with high labour market and other factor costs. The Scottish education formula allows for these types of cost to a lesser extent, and is instead more likely to compensate areas for the costs associated with sparsity and rurality.

Column B repeats the regression in column A after removing the ACA element of the English formulae. The two formulae now appear very similar, with the English formula explaining 97% of the variation in the Scottish formula. The coefficient $\alpha 1$ however has increased to 1.2 (and is statistically different from unity), suggesting that the English formula is scaled slightly more 'steeply' – giving relatively more resources to the most needy regions and relatively less to the least needy.

Column C presents the results of regressing the English health need scores on the Scottish health need scores. The coefficient α_1 of 0.64 is significantly less than unity, and the R^2 of 0.68 suggests some variation about this average. As with the education formula, much of the discrepancy between the formulae arises because the English health formula contains an ACA which tends to allocate relatively more resources to London and the south east. Excluding the ACA from the English formula and re-running the regression (column D) sees the coefficient α_1 rise to 1.1 (we cannot reject the null hypothesis that α_1 is equal to one) while the R^2 increases to 0.92.

Table 3.5: Regression of regional spending needs assessed by the English formula on regional spending needs assessed by the Scottish formula

	Education formula	Education formula (excluding ACA)	Health formula	Health formula (excluding ACA)
	(A)	(B)	(C)	(D)
Coefficient on Scottish formula (α_1)	1.108 (0.227)***	1.204 (0.066)***	0.636 (0.138)***	1.06 (.088)***
R^2	0.701	0.971	0.68	0.94

*Standard errors in parentheses. *** = significant at 1% level. N=12 in all cases*

This analysis suggests that if ACAs are ignored, then the English and Scottish need formulae for both education and health appear very similar in how they estimate relative need. But if the English and Scottish need formulae produce similar estimates of regions' spending needs only when the ACAs are excluded, would England and Scotland be able to agree on a common approach to the labour market cost problems that the ACAs seek to address?

On the one hand, it could be argued that there is scope for acceptable compromise here. The Scottish health formula does allocate marginally more resources to health boards in urban areas to compensate for higher labour costs, so the principle of allocating more to areas which face higher

factor costs is enshrined within the Scottish formula. Given that Scotland is willing to make some adjustment for these costs, it seems possible that it would agree to make rather more allowance for the much larger problems that arise in and around London.

However, there is an ongoing debate in England itself about the relatively high per capita public spending in London engendered by the ACAs. People's views here depend on their beliefs about the geographical agglomeration of activity in south-east England. One view is that agglomeration increases productivity at a national level, and that public sector pay structures should reflect this; (Propper and Van Reenan, 2010) argue that current pay structures in the NHS result in wages that are too low relative to private sector wages in high-cost areas, resulting in poor NHS service quality. Another view is that agglomeration brings high external costs, which the current system of public expenditure tends to exacerbate, and that 'there is a strong economic and social case for spreading income and wealth creation more evenly' (Mackay and Williams, 2005). As Elliot et al. (2005) highlight, there is a difficult balance between efficiency and equity in the allocation of resources across regions.

vii. Conclusions

Devolution to Scotland, Wales and NI has provided the devolved governments in these territories with the opportunity to pursue different policies. This policy distinctiveness is particularly evident in health, where different decisions around structure, commissioning, and treatment availability are evident. This ability to pursue policy distinctiveness is both the intention and strength of devolution.

But the significant spending and policy autonomy of the DAs raises challenging questions about how the DAs should be funded. There has long been a widespread dissatisfaction (in England and Wales) with the way in which resources are allocated to the DAs, and repeated calls for the Barnett formula to be replaced by a 'fairer' system of needs assessment. However, a replacement for the Barnett Formula has never been seriously proposed. The reason appears to lie in a belief that the Westminster Government and the DAs would never be able to agree on what a needs-based grant allocation formula might look like, given that needs formulae require normative judgements about what constitutes need, and can be subject to political manipulation.

This paper tests this hypothesis by comparing the way in which the Scottish Government allocates resources for health within Scotland with the way in which the Westminster Government allocates resources for health within England. The rationale is to test whether the two formulae – developed by different researchers using different data, and working within different political frameworks – provide similar or very different estimates of the healthcare spending needs of the UK territories.

The two formulae considered are the allocation formulae used by the NHS in England and the NHS in Scotland to allocate resources to PCTs and Health Boards respectively. Although it would never be possible to say that one of these formulae was more accurate, or 'right' than the other, both formulae are 'trusted' in the sense that they have been used by politicians of different persuasion to guide territorial resource allocation for several decades.

In a statistical sense, results show that the two formulae do provide very similar estimates of the spending need of UK regions and territories, despite relying on a wide variety of different needs indicators. For most regions, the difference between the two formulae in the allocation of healthcare resources amounts to less than 3% per capita. The English formula tends to allocate slightly more resources to urban areas facing higher factor costs and higher levels of labour-market disadvantage. The Scottish formula tends to allocate relatively more resources to rural areas. This result is neither alarming nor surprising. Rather than being due to normative differences between the formulae in what constitutes a need, it stems from the fact that the two formulae have been developed based on evidence from territories with slightly differing characteristics (i.e. there are no remote island PCTs in England which explains why the English formula is less sensitive to the needs of Scottish island PCTs).

The results of the analysis suggest that there are more similarities than differences in the pattern of grant allocations made by the English and Scottish needs assessment formulae for health. An extension of the analysis to education services also suggests a reasonable degree of similarity between the allocative outcomes of the English and Scottish formulae. Where there are differences, these often relate to the treatment of factor costs.

This is not to say that the development of a needs based replacement for Barnett will be politically easy. Nor is it suggested that a Barnett replacement should be based on either the English or Scottish formulae, or even a hybrid of the two. Indeed, using a sub-national formula to allocate resources at national level may create perverse incentives in how DAs might design their own formulae; and Wales and NI are likely to feel institutionally disadvantaged if a Barnett replacement was based on English and Scottish funding formulae rather than their own.

Nonetheless, the analysis may be useful in establishing a starting point and negotiation positions for each DA in the event of Barnett reform. Results suggests that Scotland's spending need for health, relative to England, ranges from 7% - 11% higher per capita, depending on whether the English or

Scottish health formulae are used for assessment. For education, Scotland's spending needs are between 1-2% below England's depending on which formula is used.

Health and school education combined account for around 50% of the Scottish Government's devolved budget. It is likely to be more difficult to assess in a formal sense the two governments' priorities for allocation of many other public services, particularly where those allocations are not determined by explicit formulae. But it seems unlikely that Scotland's spending need for the remaining 50% of public services is sufficiently high to justify its 17% per capita total spending advantage over England. If this analysis had been based purely on the English allocation formulae, then a Scottish Government could argue that the English formulae did not adequately reflect Scotland's spending need, perhaps because it did not adequately reflect the influence of sparsity on the costs of public service provision. However, the Scottish Government would be in a weaker position in this regard given that their own formula have been used to estimate the broad parameters of its spending need envelope.

It could be argued however that statistical similarity is not the same as political similarity. Relatively small changes or differences in funding amounts can become political totemic within the UK's constitutional politics.

Indeed, a more fundamental barrier to the adoption of needs assessment to allocate grant to the devolved administrations is the fact that needs assessment is inherently political. Any needs-based system of grants for the DAs must identify – either implicitly or explicitly – some standard spending policies to use as the basis for determining grants. Under the Barnett Formula, the grants paid to the DAs change in line with changes in English spending, implying that English spending policies are taken as the standard. Heald (2009) describes this as policy decisions in England having 'first-mover advantage', which seems counter to the spirit of devolution.

This feature of the Barnett Formula is becoming increasingly problematic now that major differences in public service provision are arising between England and the DAs, with the DAs' preferring a more social democratic approach to public services. One example concerns the decision in England to reduce public spending on higher education, offset by the introduction of tuition fees, which resulted in lower block grants for the DAs. Another example is the trend in England, but not Scotland, toward greater use of competition and commissioning frameworks in health provision (National Audit Office, 2012).

Designing a system of needs assessment when policies diverge is potentially more challenging, as it is less clear how to define the relevant policy standard against which needs are assessed (King and

Eiser, 2016) . This policy divergence is likely to increase the difficulties associated with agreeing what constitutes a legitimate spending need. It is this fact which is likely to undermine the development of needs assessment formulae for allocating grant between the UK's relative autonomous devolved governments.

4. The labour-market resilience of UK regions to the Great Recession

David N.F. Bell and David Eiser

Abstract. London was at the vanguard of the UK's recovery from recession, recovering its pre-recession levels of output and employment more rapidly than other regions. London has continued to experience more rapid growth in employment and employment rate than other regions in the period since 2012. But it has experienced somewhat larger real wage falls in this period relative to other regions. This paper considers possible reasons for these patterns. It considers in particular the extent to which the larger proportion of migrant labour in the capital's economy might account for the observed trends.

i. Introduction

The UK Great Recession of 2008/9 was the deepest since the Second World War, with real GDP falling by over 6% (Blundell et al. 2014). The UK entered recession in Q2 2008, and although it officially exited recession in Q3 2009, output remained below the pre-recession peak for far longer than in previous recessions, taking some 5.5 years to recover its pre-recession level. Yet despite this sluggish performance on output growth, headline measures of labour market performance have consistently been stronger than expected. Employment fell by just 2% and had recovered its pre-recession level almost a year before the return to pre-recession output levels. Hours worked also fell by less than the fall in output and recovered more quickly. And the UK unemployment rate has consistently come in below the Bank of England's projections (Haldane, 2015).

The experience of this recession has thus been very different from the recessions of the 1970s, 80s and 90s, when output recovered faster than employment. But the implication is falling productivity measured both as output per worker and output per hour. Five years after the start of the recession, hourly productivity was 3% below its 2007 level (and 15% below trend) (Pessoa and Van Reenan, 2014). At the same time, real wages fell throughout the period of 'recovery', and are now some 10% lower than before the recession (Gregg et al. 2014a).

The emerging explanation for the productivity puzzle is that UK labour markets are now more flexible due to a combination of factors relating to welfare reform, demographics, declining unionisation and greater use of flexible working arrangements (e.g. zero hours contracts) (Blundell et al. 2014; Gregg et al. 2014b). Welfare reforms have increased the conditionality attached to benefits. When combined with more generous in-work support through tax credits, the unemployed have become closer substitutes to those in work. This keeps labour supply high even when demand is weak (Pessoa and Van Reenan 2013; Gregg et al. 2014b). Increased labour market participation by older workers and higher levels of immigration have also contributed to the increased labour supply. Increased supply has exerted downward pressure on reservation wages because labour market participants increasingly care about staying in work rather than securing higher wages. One manifestation of this has been a dramatic rise in the numbers of self-employed with very low incomes, a proportion of whom may be classed as 'hidden unemployed', as well as a substantial rise in the numbers employed part-time who would prefer to work full time (Bell and Blanchflower 2011; 2013).

Increased labour supply is part of the explanation for falling real wages which in turn might induce firms to substitute labour for capital. In addition, limitations on access to credit for businesses and bank forbearance of bad debtors (Riley et al. 2014), as well as some industry-specific factors

(McAfferty, 2014) has led to lower investment – and a consequent fall in the capital-labour ratio which in turn has led to lower productivity (Pessoa and Van Reenan, 2014).

The broad picture just sketched for the UK – a prolonged fall in output and productivity, stronger performance of employment, rapid growth in part-time and self-employment, and dramatic falls in real wages – has been extensively documented. But there has been relatively little work aimed at understanding the extent to which this broad national picture might vary across UK regions.

Responses to recessionary shocks do tend to differ across UK regions, with the geographical pattern of the response reflecting the nature of the shock (Martin, 2012). The fact that the 2008-9 crisis was precipitated by a financial sector crisis led some to hypothesise that London would be hit harder by the recession than other regions. On the other hand, London has tended to be more resilient to past recessions in general than other parts of the UK (Fingleton et al. 2012). Moreover, there has been a longer-run trend towards a greater spatial concentration of the UK's economic activity in the capital (Martin et al. 2015). At the same time, concern has been expressed that the Coalition Government's fiscal consolidation programme during the 2010-15 Parliament might disproportionately disadvantage those regions outside of London and the South East that are relatively more reliant on the public sector for job opportunities (Martin, 2012; Beatty and Fothergill, 2011; Jones and Green, 2009).

This paper investigates the extent to which the UK's general experience during the recession and recovery has played out at a regional level. It considers the resilience of UK regions to the recession (measured both in terms of regions' resistance to the recessionary shock and the subsequent speed and extent of recovery from recession). Resilience is considered in terms of employment levels and rates, output (GVA) and real wages. The paper explores the extent of slack in regional labour markets, measured by unemployment, underemployment and inactivity rates. The paper then examines the role that both immigration of overseas-born individuals and inter-regional migration within the UK have played in shaping the observed trends during the recession and recovery, finding that migration has played a strong role in shaping the UK's regional response to the recession. The paper also considers the extent to which the UK Government's programme of fiscal austerity may have played in determining the varying experiences of the UK regions during the recession.

Regions are of course not the only spatial unit at which one might wish to consider the geographical variation in response to the recession. It will be equally important to consider the experience of cities (Martin et al. 2014) or some particular aggregation of local authority areas (Industrial Communities Alliance, 2015). But the advantage of looking at regions is the depth and timeliness of economic data that exists for them. As with any study of spatial interaction, the difficulties of

arriving at unambiguous definitions of spatial units mean that results have to be interpreted with a degree of caution.

The remainder of this paper is structured as follows. The next section analyses headline measures of economic performance – output, employment and wages – for the UK regions; it demonstrates that London has recovered particularly quickly from the recession in output and employment terms, but that, despite this, real wages in London have fallen more than elsewhere. Section 3 assesses the extent of labour market slack in each region, drawing on analyses of underemployment, unemployment and inactivity; unemployment and underemployment are almost back to pre-recession levels in London. The capital has also experienced large falls in inactivity. Section 4 considers the role that patterns of migration have played in influencing these trends; Section 5 concludes.

ii.   London at the vanguard of the UK's recovery from recession

There are several ways of measuring or characterising the resilience of regional employment or output to recession (Martin, 2012; Fingleton et al. 2012). One of these is resistance, the vulnerability or sensitivity of a region to a recessionary shock. Another is the speed and extent of recovery from recession.

Table 4.1 considers how these two concepts of resilience have played out across UK regions[7]. London emerges clearly as the most resilient region during the Great Recession in terms of total employment (panel 1). London experienced a relatively small peak-to-trough fall in employment; the pre-recession level of employment was regained after 26 months, more rapidly than any other region; furthermore employment is now 17% above the pre-recession level, which represents a substantial increase in total employment relative to other regions.

Employment growth in London partly reflects increases in total population, a point to which we subsequently return. However, in terms of its employment rate (panels 2 and 3), London does not appear to be more resistant than other regions. But it has proved much more resilient in terms of its speed of recovery: London's 16+ employment rate returned to the pre-recession level more quickly than other region[8]., and is now 3.3 percentage points higher than the pre-recession level. This can in part be explained by the fact that the increase in the relative size of the pensioner population has

---

[7] The 12 UK regions and the abbreviations we use in this paper are: North East (NE); North West (NW); Yorkshire and the Humber (Y&H); East Midlands (EM); West Midlands (WM); Eastern (East); London (LDN); South East (SE); South West (SW); Wales (WAL); Scotland (SCO); and Northern Ireland (NI).
[8] Wales' working age employment rate returned to its pre-recession level two months more quickly than London's, but this recovery proved to be shortlived, and Wales' employment rate subsequently fell back.

been smaller in London than elsewhere. In terms of the working age employment rate London's relative performance has not been quite so strong, although it is nonetheless the case that London recovered its pre-recession employment rate faster than other regions (bar Wales) and its rate is now higher relative to the pre-recession rate than in other regions.

In the context of the employment and employment rate trends just discussed, there is something of a puzzle in terms of the behaviour of regional real wages not so much during the recession itself but in the subsequent recovery. As shown in panel 4 of Table 4.1, London experienced real wage falls that were not dissimilar from those of other regions during the recession (defined broadly as the period between median weekly earnings in 2007 and in 2010). But in the period since 2011 London has experienced particularly strong ongoing declines in real wages, relative to other regions (Figure 4.1).

This trend for London to have experienced larger wage falls than other regions in the recovery from recession is also observed if we look at hourly wages or the weekly earnings of full-time workers, suggesting that this is not fundamentally an issue about a disproportionate rise in part-time working in the capital.

It is also worth flagging up that Scotland, relative to other regions, has seen relatively poor employment performance but smaller declines in real wages in the post-recessionary period. This is potentially significant in the context of ongoing devolution of income tax powers to Scotland. This is the topic of the subsequent paper in this PhD.

The employment trends are reflected in data on hours worked and output (Figure 4.2 shows the change in various key metrics between 2007-13). By 2013, real output was 8% higher in London compared to pre-recession, whilst hours worked were 11% higher. By 2013, most other regions (other than the SE) remained below pre-recession levels of output. In terms of Gross Value Added (GVA) per capita, no region had recovered its pre-recession level, but London was the closest to having done so.

London has not however been immune to falling productivity. Productivity growth in all regions decoupled from the long-term trend in the mid to late 2000s and has stagnated or even fallen slightly since[9].

---

[9] Figure 4.2 gives the impression that productivity has increased in Scotland, but this is purely the result of the fact that productivity in Scotland increased significantly in 2008; it has stagnated since then and declined slightly in 2013.

Table 4.1. *Regional employment, employment rates and real wages in the recession and recovery*

	NE	NW	Y&H	E Mids	W Mids	East	London	SE	SW	Wales	Sco	NI
Total employment (000s)												
Pre-recession level	1132.8	3143.6	2409.5	2114.9	2490.0	2725.4	3721.6	4149.6	2496.5	1332.5	2506.6	778.0
Post-recession level	1095.7	3083.2	2324.9	2087.6	2393.0	2724.3	3682.2	4046.2	2396.8	1297.6	2438.2	743.0
Pre to post fall	-3.3%	-1.9%	-3.5%	-1.3%	-3.9%	0.0%	-1.1%	-2.5%	-4.0%	-2.6%	-2.7%	-4.5%
Months to return to pre-recession level	71	53	73	73	57	35	26	65	74	65	75	37
Latest (Dec-Feb 2017) to pre-recession	3%	5%	4%	4%	3%	6%	17%	5%	5%	4%	0%	3%
16+ employment rate (%)												
Pre-recession rate	55.8	57.8	59.4	61.1	58.7	62.0	61.6	63.8	61.4	56.5	60.7	57.7
Post-recession rate	53.7	56.4	56.9	59.8	56.1	60.8	59.5	61.8	58.9	54.3	58.9	54.4
Pre – to post fall	-2.1	-1.5	-2.5	-1.4	-2.6	-1.2	-2.0	-2.0	-2.5	-2.2	-1.8	-3.3
Months to return to pre-recession rate	92	83	104	n/a	102	88	57	n/a	88	68	n/a	97
Latest (Dec-Feb 2017) to pre-recession	0.5	1.4	-0.1	-1.3	-1.2	-0.4	3.3	-0.9	0.1	0.8	-1.8	-1.0
16-64 employment rate (%)												
Pre-recession rate	68.4	70.3	72.2	74.2	71.5	75.6	69.9	77.6	76.9	69.7	74.3	68.2
Post-recession rate	65.9	68.6	69.3	72.8	68.4	74.9	67.5	75.1	73.6	67.5	71.8	64.5
Pre – to post fall	-2.5	-1.7	-2.9	-1.5	-3.2	-0.7	-2.5	-2.6	-3.3	-2.2	-2.5	-3.7
Months to return to pre-recession rate	71	77	77	74	98	67	65	91	83	53	84	77
Latest (Dec-Feb 2017) to pre-recession	2.1	3.2	1.6	1.0	0.2	1.4	3.5	1.0	1.9	3.3	-0.9	0.6
Real median weekly wage 2007	410	439	431	438	435	479	599	506	433	415	443	406
Change real weekly wage 2010-2007	0.9%	-2.6%	-1.1%	-1.3%	-1.0%	-1.0%	-2.0%	-2.2%	-1.6%	0.1%	-0.1%	-1.6%
Change real weekly wage 2016-2007	-1.2%	-6.0%	-6.6%	-5.1%	-4.5%	-4.8%	-10.4%	-5.4%	-5.3%	-1.9%	-2.1%	-3.3%

Source: Labour Force Survey (LFS) 3-month average labour-market statistics (seasonally adjusted) and Annual Survey of Hours and Earnings (ASHE). For the employment level and rates, the pre-recession is Jan – Mar 2008 (the quarter prior to recession). The post-recessionary trough is taken as Jul-Sep 2009, the first quarter after the end of the technical recession.

Figure 4.1: Real terms change in gross weekly wage by region, 2007=1
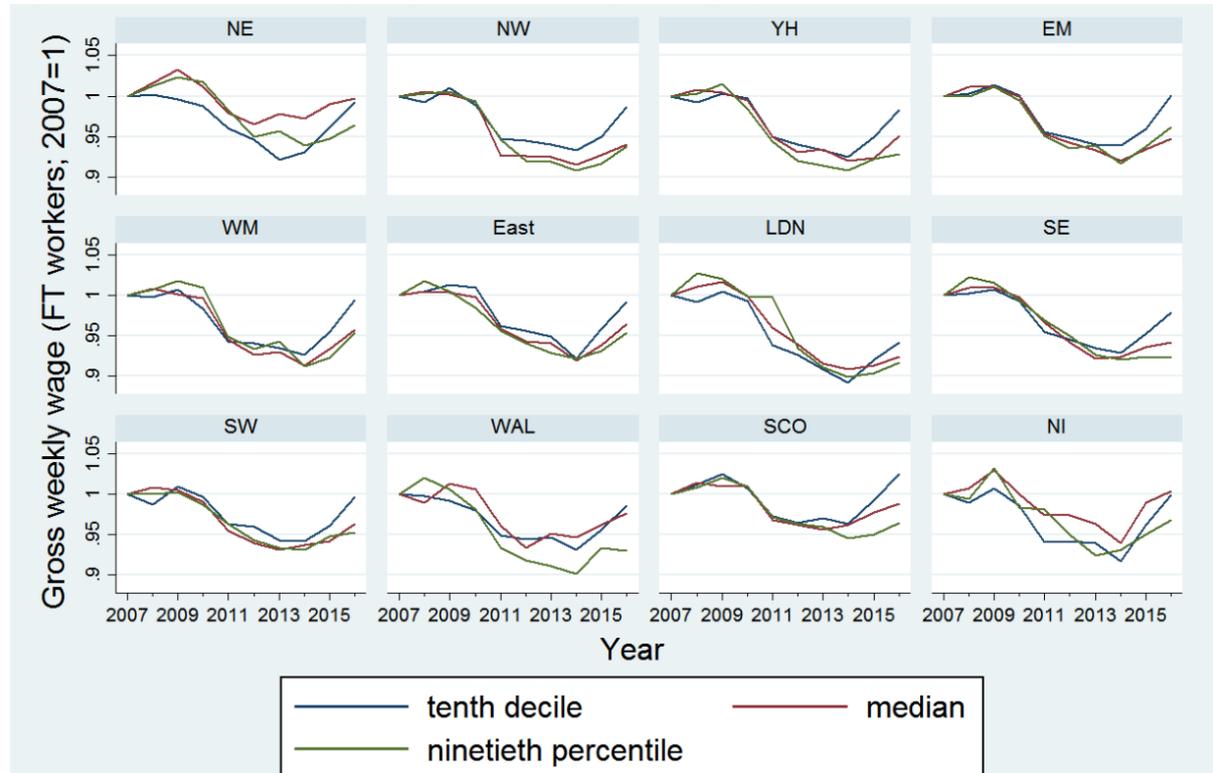
Source: Annual Survey of Hours and Earnings

Figure 4.2: **Gross Value Added (GVA), hours worked, productivity and output per capita**
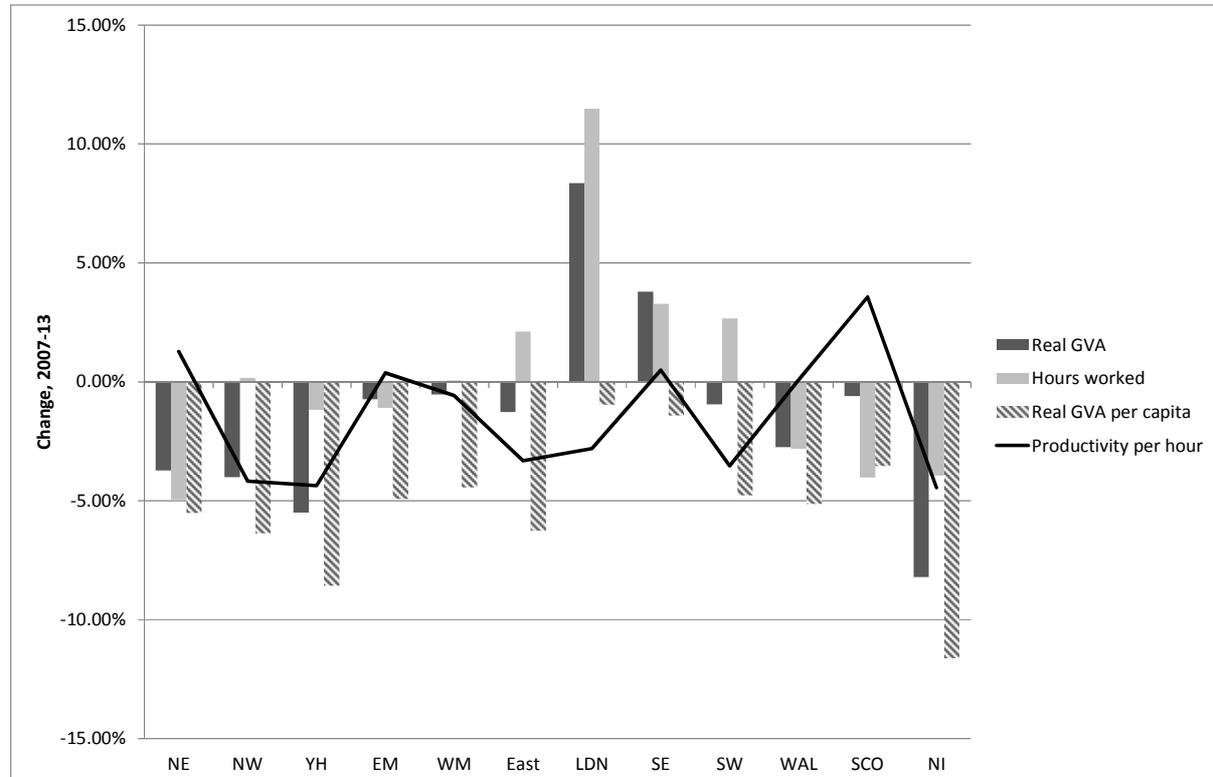
Source: ONS Regional Productivity Statistics

The picture then is one of much stronger recovery from recession of London's economy in relation to employment, output and employment rates, but combined with relatively marked wage falls in the capital.

What might explain London's relatively stronger employment recovery during the recession and subsequently? Gordon (2016) considers a range of hypotheses. These include various structural advantages (including roles for the industrial mix and the boost from the devaluation of the pound on in-bound tourism, with lesser roles for occupational structure and the growth in fraction of wealthy global elites who base themselves for at least part of the time in London) and spatially significant (public and private) policies – these include a concentration of head office and other strategic functions in London in response to general rationalisation pressures, and a greater concentration of public investment in the capital.

Gordon argues however that these traditional explanations can only partially account for London's employment record relative to the national story in this recession, given that most of these factors were also applicable in the recession of the early 1990s – but during that recession, London did not experience similarly strong employment growth relative to the national average. The missing piece of the story is argued to be the UK Government's support of the financial sector, including bail-outs, implicit subsidies, and monetary expansion (via bank rate reductions and injections of liquidity). Whilst recognising that the precise role of the financial support policies in supporting the London-centred boom is unquantifiable, the argument is that subsidies to the investment banks absorbed some of the momentum of the expected recessionary downswing, whilst the quantitative easing policies supported a strong upswing following the recession by benefitting the institutions that handled resulting financial transactions, and the shareholders of those institutions.

Notwithstanding explanations for London's strong employment performance, the question remains how this can be reconciled with sharper real wage falls in London than elsewhere, particularly since 2012. There are on the face of it a number of broad potential explanations. One is that the large rise in employment in London has resulted in major changes in the composition of employment (whether by occupation, demographics, etc.) which have served to reduce the aggregate median earnings in London faster than other regions. Another is that London might have some combination of higher levels of labour market slack, together with a labour market that is relatively more flexible in the face of that slack. It may well be for example that the composition of the London economy (including lower rates of union coverage, higher proportion of private sector employment, and a relatively younger workforce) result in wages in the capital being more responsive to a given level of labour market slack.

The next section considers whether trends in labour market slack across regions can shed any further light on this apparent puzzle.

iii. Labour-market slack

In seeking to explain differential falls in real wages across regions, a logical first step is to consider the level of labour market slack, or spare capacity in the labour market, in each region. Labour market slack is most commonly measured by the unemployment rate, and specifically as the gap between the unemployment rate and some measure of the 'natural' or 'long-term' unemployment rate.

One of the features of the Great Recession at the UK level has been the relatively small rise in unemployment compared to previous recessions. But this seems hard to square with the dramatic real wage falls observed. Indeed, since the recession both the unemployment rate and real wage growth have consistently come in under the Bank of England's projections (Haldane, 2015). As a result there has been much debate as to how much slack exists in the labour market.

In this context, it is clear that the unemployment rate is not the only gauge of labour market slack, particularly during a recession. Assessments of labour market slack need also to consider the extent of hidden unemployment (the economically inactive[10] who are not currently searching for work, but who would rejoin the labour force if the job market were stronger), and the incidence of underemployment (those working part-time who want full-time work) (Blanchflower and Levin, 2015; Green and Livanos, 2015). The remainder of this section considers how trends in three measures of labour-market slack: unemployment, underemployment, and inactivity, have evolved in each UK region during recession and recovery.

The first panel of Table 4.2 shows the evolution of working age unemployment rates in the UK regions during the recession and subsequently. Between the start and end of the recession itself, the unemployment rate in London increased by 2.3 percentage points; this represents a slightly smaller increase than observed in most other regions, although from a higher base. The unemployment rate in London returned to its pre-recession rate before any other region, and is now one percentage point below the pre-recession rate. Compared to its pre-recession unemployment rate therefore, London appears to be faring better than most other regions, although the unemployment rate in London remains higher than it does in all other regions bar the north east.

---

[10] The economically inactive are those who are not participating in the labour force, in other words those who are neither in work nor actively searching for work.

At UK level the recovery from the recession has been characterised by a disproportionate rise in part-time relative to full-time work. Many of those accepting part-time work are doing so only because full-time work is unavailable. The result is a rise in involuntary part-time working. As well as the headline rate of unemployment therefore, assessments of labour market slack need also consider the incidence of underemployment (that is, people working part-time who want a full-time job). The underemployment rate measures the extent to which workers would like to work greater (or fewer) hours at their current wage (Bell and Blanchflower, 2011; 2013). When combined with the unemployment rate it thus provides a broader measure of labour market slack, considering slack at the intensive margin as well as the extensive margin.

Figure 4.3 shows the evolution of the underemployment rate across UK regions between 2002-14. In all regions, underemployment increased by more than unemployment during the recession. Underemployment also remains further above the pre-recession level than unemployment in all regions. Similar to the story for unemployment however, the underemployment rate in London is now closer to the pre-recession level than it is in other regions, albeit remaining relatively high in London.

At national level, labour market slack is often measured in relation to the deviation from longer-term 'normal' (or 'natural') rates of unemployment or participation. On this basis, the fact that London has returned to pre-recession rates of unemployment and underemployment might indicate that there is relatively little slack in the regional labour market. This interpretation however does little to explain why wages have fallen relatively more in London than in other regions. On the other hand, unemployment and underemployment rates remain as high or higher in London than in some other regional labour markets.

Table 4.2. *Regional unemployment and inactivity rates*

	NE	NW	Y&H	E Mids	W Mids	East	London	SE	SW	Wales	Sco	NI
Working age unemployment rate												
Pre-recession rate	6.4	6.1	5.0	5.5	6.2	4.6	7.0	3.9	3.7	5.3	4.7	4.6
Post-recession rate	9.9	8.8	8.7	7.3	10.2	6.4	9.3	6.2	6.6	8.8	7.3	7.2
Pre – to post increase	3.4	2.7	3.7	1.8	4.0	1.8	2.3	2.3	2.9	3.5	2.5	2.6
Months to return to pre-recession rate	0.2	-1.5	0.2	-1.4	-0.4	-0.1	-1.0	-0.5	0.0	-0.2	-0.1	0.8
Latest (Dec-Feb 2017) to pre-recession	102	83	n/a	81	86	84	76	91	n/a	95	107	n/a
Working age inactivity rate												
Pre-recession rate	26.9	25.1	24.0	21.5	23.7	20.8	24.8	19.2	20.2	26.4	22.0	28.5
Recession peak	26.9	24.8	24.1	21.5	23.8	19.9	25.6	20.0	21.2	26.0	22.6	30.6
Difference	0.0	-0.3	0.1	0.0	0.1	-0.8	0.8	0.8	1.0	-0.4	0.6	2.0
Latest to pre-recession	-2.4	-2.2	-1.8	0.1	0.2	-1.3	-2.9	-0.6	-2.0	-3.3	1.1	-1.2
Months to return to pre-recession rate	18	18	20	74	19	18	30	64	68	18	43	33

Source: Labour Force Survey (LFS) 3-month average labour-market statistics (seasonally adjusted). The pre-recession rates are taken from Jan – Mar 2008 (the quarter prior to recession). The post-recessionary rates are from Jul-Sep 2009, the first quarter after the end of the technical recession.

Figure 4.3: Quarterly underemployment by region, 2002-14
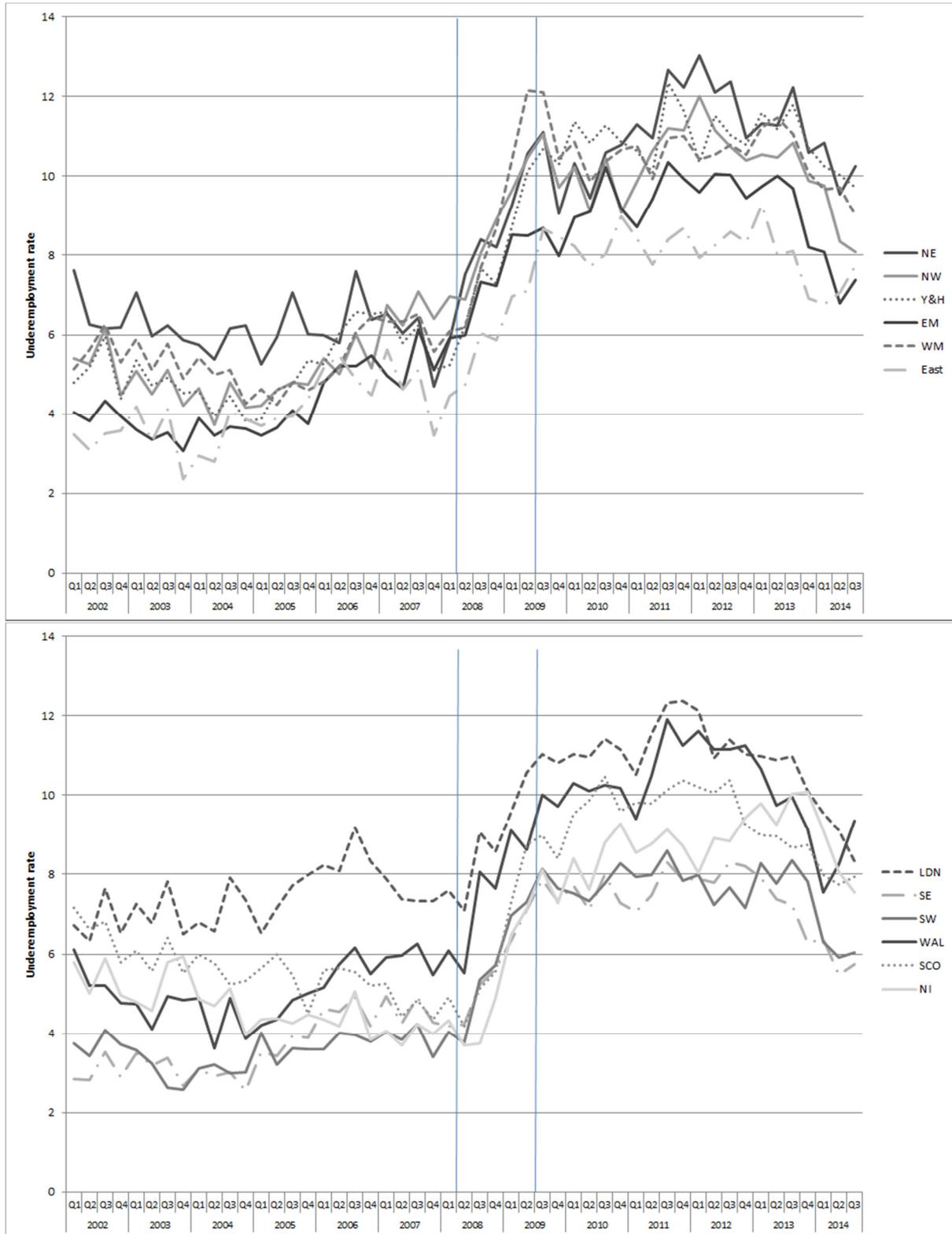
Source: Labour Force Survey. Notes: Data is quarterly, seasonally unadjusted, from Q1 2002 to Q3 2014. The two horizontal lines demarcate the start and finish of the UK recession (Q2 2008 – Q3 2009).

Another measure of labour market slack is the proportion of the working age population who are economically inactive (i.e. not participating in the labour market). Inactivity is correlated with unemployment, particularly during recessions, reflecting an element of 'hidden unemployment' (McVicar, 2006). The hidden unemployed are individuals who are not part of the official unemployment statistics, often because economic conditions dissuade them from actively searching for work, but who would return to the labour force if economic conditions improved. In this context, one way to infer labour market slack is to consider the rate of inactivity relative to some measure of the long-term or equilibrium level of participation.

The second panel of Table 4.2 shows the evolution of working age inactivity rates in the UK regions during the recession and subsequently. London (together with the south east and south west) experienced relatively larger increases in inactivity during the recession, but since the end of the recession inactivity rates have fallen reasonably markedly in London, as they have done in other regions. The inactivity rate in London is currently in line with average across UK regions.

On the basis of the evidence presented here, it is hard to argue that relatively more pronounced wage falls in London since the end of the recession can be accounted for by higher levels of labour-market slack in the London economy. Indeed if slack is measured relative to pre-recession levels of unemployment and underemployment, it could be argued that there is less slack in London's labour market than there is in most, if not all, other regions. –. London's working age inactivity rate is well below its pre-recession level, but this is also the case for several other regions.

Gauging exactly the extent to which there remains slack in the labour-market is difficult because it is less clear whether inactivity rates will continue to fall. There has been debate for example as to whether various reforms to illness and disability benefits made in the UK since 2011 may have raised the equilibrium participation rate. Indeed, in 2014 the Bank of England's Monetary Policy Committee revised upwards its estimate of the UK's equilibrium level of participation.

iv. Role of migration

If relatively greater wage falls in London cannot be explained by greater labour market slack, what else might account for it?

One of the most obvious distinguishing features of the London economy since the recession has been the substantial increase in working age population, driven largely by immigration. Between 2011-15, London's working age population increased by 4.3%, compared to just 0.7% for the UK as a

whole (after London, the region with the second highest rate of population increase in these years was east region, but its population growth rate was less than a quarter of London's).

Growth in the overseas born population explains a large part of London's recent population growth. LFS data shows that between 2011-16 London's overseas population increased by 16%, compared to growth of 3% in the UK born population. Outside of London the overseas born population increased even more rapidly − by 24% - but in aggregate terms this makes less of an impact on the rate of total population growth given that the overseas born account for a much smaller percentage of the working age population outside of London (9%) than inside (35%).

What is even more striking is how the employment rate of immigrants in London has increased in the post-recessionary period. The employment rate of the overseas born in London has increased from around 64% at the beginning of 2012 to 74% in 2016, meaning that the employment rate of the overseas born is now on a par with those who are UK-born (Figure 4.4) in the capital. Outside London, the employment rate of the overseas born has also increased substantially, but it makes less difference to the total employment rate given that the overseas born make up a smaller proportion of the labour-market.

To formalise the impact of non-UK born individuals on the overall employment rate, we disaggregate the change in the employment/ inactivity rate in each region into three parts: the part attributable to the change in the UK-born rate, the part attributable to the change in the non-UK born rate, and the part due to the population share of migrants. The disaggregation recognises that the employment (E) to population (P) ratio (i.e. the employment rate) is a function of the employment rate for UK born nationals weighted by the population share of UK nationals (θ), and the employment rate of overseas-born individuals (OB) weighted by the population share of the overseas-born (1-θ). Thus the change in the employment rate between period 0 and period 1 is given by:

$$\frac{E_1}{P_1} - \frac{E_0}{P_0} = \left(\theta^{UK} \frac{E_1^{UK}}{P_1^{UK}} + (1 - \theta^{UK}) \frac{E_1^{OB}}{P_1^{OB}} \right) - \left(\theta^{UK} \frac{E_0^{UK}}{P_0^{UK}} + (1 - \theta^{UK}) \frac{E_0^{OB}}{P_0^{OB}} \right) + e \qquad (1)$$

The term e is a residual which measures the effect of the change in the population share of the two groups. We exploit the identity in equation 1 to disaggregate the change in employment rate into the three components.

Table 4.3 shows the results of this disaggregation for the period from Q1 2012 to Q4 2016. We contrast London with the rest of the UK regions (rUK).

Table 4.3 shows that, of the 7 percentage point increase in London's employment rate between Q1 2012 and Q4 2016, 3.8 percentage points can be attributed to the increased overseas-born employment rate and 2.9 percentage points can be attributed to the increased employment rate of UK-born individuals. (The residual, -0.4 percentage points, reflects the change in the share of the overseas-born population.) For rUK, of the 4.1 percentage point rise in the employment rate, only around half a percentage point can be attributed to the overseas-born population.

Table 4.3: Decomposing aggregate working age employment rate changes into part attributable to overseas-born and part attributable to UK-born residents

	London	rUK
Employment		
Increase in employment rate Q1 12 – Q4 16	7.1%	4.1%
Change accounted for by UK-born rate	2.9%	3.7%
Change accounted for by overseas-born rate	3.8%	0.5%
Change accounted for by changing population share	-0.4%	-0.1%

Source: Labour Force Survey.

Might the disproportionate role of immigrants to the London economy since 2012 account for London's relatively steeper wage declines?

From a theoretical perspective, the impacts of immigration on the wages and employment of native workers is ambiguous. The impacts depend, in the short-run, on the extent to which immigrants' skills are complements or substitutes to the skills of existing workers (if the elasticity of substitution between immigrants and natives exceeds one, immigrants will displace native workers), and, in the longer-run, on how immigration affects the demand for labour. The impact of immigration is likely to depend on whether the economy is at full employment or if there is excess demand for labour.

Immigrants may render the labour market more flexible given that they have a lower propensity to join unions and/ or because their reservation wage is below that of native workers (Angrist and Kugler, 2003). On the other hand immigrants may increase labour demand if they are more entrepreneurial than native-born workers.

The effect of immigrants on labour markets is not restricted to wages and employment. Economies may also respond to increases in immigration by altering the output mix or by influencing the use of particular technologies. For example, suppose that immigration leads to an increase in supply of unskilled workers; this may encourage the use of production technologies that are more intensive in the use of low-skilled labour, in turn enabling the economy to absorb the additional supply of low-skilled labour without significant changes in the wage structure or output mix (Dustmann et al. 2008).

Empirical research on the labour market effects of immigration to the UK tends to find little adverse effects of immigration on wages and employment of the UK-born. Manacorda et al. (2011) find that immigrants and UK-born workers are not close substitutes on average, although the less skilled are closer substitutes for immigrants than the more highly skilled. Similarly, Dustmann et al. (2005) consider the period 1983 – 2000 and find that immigration had no statistically significant effect on the employment outcomes of UK-born workers – although immigration is more likely to have a negative impact on the employment outcomes of the UK-born who have low educational attainment.

In terms of wages, for the period 1997-2005, Dustmann et al. (2013) find that an increase in the number of migrants corresponding to 1% of the UK-born working-age population resulted in an *increase* in average wages of between 0.1% and 0.3%. Conversely Reed and Latorre (2009) consider the UK over the period 2000-2007 and find that a 1% increase in the share of migrants leads to a 0.3% *fall* in average wages of the UK-born.

Average effects conceal the aforementioned fact that immigration tends to have asymmetric effects across the skill/ wage distribution. Dustmann et al. (2013) show that a 1% increase in the share of migrants in the UK-born working age population leads to a 0.6% decline in the wages of the 5% lowest paid workers and an increase in the wages of high paid workers. Moreover the adverse wage effects of immigration are likely to be greatest for resident workers who are themselves migrants, as the skills of these 'incumbent migrant workers' are likely to be closer to the skills of new migrants than for those of UK-born workers.

Of course the effects of immigration on 'home' labour markets depend on the characteristics of migrants, including their skills and country of origin. Madsen and Andric (2017) find that, over the period 1850-2010 for countries in north and south America and Australasia, educated immigrants and those from protestant countries reduce unemployment, while poorly educated immigrants enhance unemployment. For the UK over the period 1995 – 2010, the Migration Advisory Committee found that non-EU migration was associated with a reduction in the employment of UK-born workers, but there were no statistically significant effects for EU immigration.

Given that the impact of immigrants on UK-born workers depends on the substitutability of immigrants for UK-born workers, an important strand of empirical work has considered the characteristics of recent immigrants to the UK. Clark and Drinkwater (2008) find that the labour-market outcomes of cohorts arriving between 2000-3 were markedly different to those of cohorts arriving between 2004-7, and these outcomes vary according to the migrant's area of origin. Labour-

market outcomes are influenced by migrants' human capital, language proficiency, and the length of time migrants expect to stay in the country.

Recent migrants tend to be younger than the UK-born, and those from the A8 countries tend to be employed in low-skilled sectors; as a result, the average wages of migrants fell significantly during the late 2000s (Clark and Drinkwater, 2008). Recent migrants to the UK tend to do worse on average than the native-born in terms of employment rates and earnings, particularly after controlling for education. Recent migrants from A8 countries experience the lowest returns to education, reflecting that high-skilled individuals from these countries are often employed in unskilled occupations. Clark and Drinkwater note however a distinction between short-term labour-market prospects and longer-term prospects. Many A8 migrants stay in the UK for short-term temporary or seasonal work; but the prospects for those remaining for longer is likely to depend on the extent to which workers enhance their English-language skills or acquire further human capital.

The question as to whether immigration might account for trends in the London labour market relative to other UK regions has been considered by Gordon and Kaplanis (2014) in the context of the global cities hypothesis. The global cities hypothesis posits that growth of 'service class' employment will lead to increased demand for low-paid personal services jobs and thus occupational polarisation. Gordon and Kaplanis (2014) find evidence of such polarisation in London over the period 1995-2000, but argue that this was primarily attributable to an influx of poor country migrants rather than the demands of a burgeoning global service class. Immigrants tend to be channelled initially at least into bottom-tier jobs, putting downward pressure on wages and thus allowing an expansion in employment in that tier.

But how effectively do immigrants to London's labour market subsequently 'move-up' the occupational distribution? Friedman and MacMillan (2017) find that international migrants appear to fare particularly badly in the London labour market, experiencing comparatively low upward mobility and high downward mobility. Rates of upward mobility are adversely affected by the fact that domestic migrants to London are disproportionally drawn from advantaged backgrounds, and have a tendency to 'hoard opportunities' in the capital.

The preceding discussion suggests that the increase in immigration might account for at least a part of London's relatively faster wage falls relative to other regions since 2011. Despite other regions seeing similar increases in the proportion of immigrants, the effects of immigration on London's labour market may be more marked given that immigrants make up a much larger share of those employed, and – as discussed previously – immigrants account for a greater proportion of recent increases in the employment rate in London than is the case outside the capital.

Indeed, notwithstanding whether immigration may or may not depress the wages of UK-born workers, the marked contribution of immigration to recent changes in the structure of London's labour-market might go some way to explain greater differential wage falls in the capital. In other words, greater wage falls in the capital might be accounted for simply as a result of compositional changes in the London labour market, especially given what we know about immigrants initially being channelled into relatively low-wage paying occupations.

The hypothesis that a greater proportion of immigrants employed in London might depress wages through compositional effects can be tested fairly straightforwardly with a Blinder-Oaxaca wage decomposition.

The Oaxaca-Blinder approach decomposes the difference in mean earnings between two groups into a part which can be explained by differences in composition of the two groups (i.e. what proportion of the difference in mean wages is due to differences in mean levels of education, age, etc.), and a part which remains unexplained, i.e. which reflects differences in the returns to particular characteristics in each group.

The Oaxaca-Blinder approach involves regressing the outcome variable Y (wages) on covariates X for the two groups of interest (in this case, the log real wage in 2012 and the log real wage in 2016). The difference in mean wage between the two groups $\hat{Y}_{2016} - \hat{Y}_{2012}$, can be decomposed as follows:

$$\hat{Y}_{2016} - \hat{Y}_{2012} = \hat{\beta}_{2016}(\overline{X}_{2016} - \overline{X}_{2012}) + (\hat{\beta}_{2016} - \hat{\beta}_{2012})\overline{X}_{2012}$$

The first term on the right-hand side of the equation represents the composition effect, i.e. the effect that can be explained by the difference in the mean characteristics of workers $\overline{X}_{Scot} - \overline{X}_{rUK}$, assuming that those workers received the same 'returns' as those in Scotland. The second term represents the wage structure effect, i.e. the part of the wage gap that can be explained by differences in the returns to characteristics, assuming the average characteristics of rUK workers.

In the analysis here, I run two wage decompositions: one for London, and the other for the rest of the UK. The dependent variable in each decomposition is the log of the weekly real wage (where wages are deflated using the Consumer Prices Index). The independent variable of interest is a dummy variable indicating whether a respondent was born in the UK or overseas. The hypothesis is that the endowment effect of overseas-born employees will work to depress average wages between 2012 and 2016, and that this effect will be stronger in London than it is for other regions.

Other covariates included are dummy variables for gender, highest qualification (three categories), age (four categories), whether single or in a couple, and up to 81 occupations. The analysis is constrained to full-time employees aged 16-64, and to respondents in wave 1[11].

Results are summarised in Table 4.4 below. According to the LFS data, between 2012 and 2016, average weekly wages declined by 0.76% in London and increased by 1% in rUK. These trends are clearly not as marked as the ASHE data indicates, and indeed these changes in themselves are not significant.

Nonetheless, the decomposition does indicate that the increase in the proportion of immigrants in the labour market does account for some downward movement in aggregate wages, as a result of a composition effect. In London, the increase in the proportion of immigrants can account for a 0.6% decrease in wages. Recall that this is purely a composition effect, i.e. it accounts for the fact that immigrants tend to earn less on average than UK-born workers, but says nothing about whether immigrants affect wages of UK-born workers. It is also worth mentioning that none of the other covariates is significant at the five per cent level in explaining wage changes in London between 2012 and 2016.

For rUK, the coefficient on foreign born workers is also negative and statistically significant. The coefficient for rUK is less than half the size of the coefficient on foreign born workers in London, although the London and rUK coefficients are not statistically different from one another at the five per cent level.

This analysis provides some partial evidence that increased employment of overseas born immigrants may account for greater wage falls in London between 2012 and 2016 through a composition effect. However, for this story to provide a good explanation as to why London wages fell differentially after 2012, but not during the earlier post-recessionary period, we would expect a similar analysis over the period 2008-12 to reveal no significant difference between London and rUK in terms of the extent to which immigrants account for downward pressure on wages. Running an equivalent placebo decomposition for 2008-12 suggests that the increase in migrants can also account for a slightly larger downward wage effect in the London compared to rUK labour market. In summary, the evidence here as to whether immigration to London can account for proportionately higher wage falls in the capital since 2012 is inconclusive.

---

[11] LFS respondents are interviewed over five waves, with one wave corresponding to a quarter. Thus any given respondent is included in five consecutive quarters. However, wage questions are only included in wave 1 and wave 5. Nonetheless, without conditioning on wave would result in some respondents appearing twice in the sample, which may bias results.

Table 4.4: Summary of Blinder-Oaxaca wage decomposition

	London		rUK	
difference	-0.0076	*(-0.0198)*	0.0102	*(0.0063)*
endowments	-0.0039	*(0.0113)*	-0.0011	*(0.0039)*
coefficients	-0.0049	*(0.0162)*	0.0108**	*(0.005)*
		Endowment effects		
Non-UK born	-0.0063***	*(0.0023)*	-0.00265***	*(0.000475)*
N (2012)	1847		16414	
N (2016)	1836		15546	

Figure 4.5: Working age employment rate in London by country of birth, 2007-16
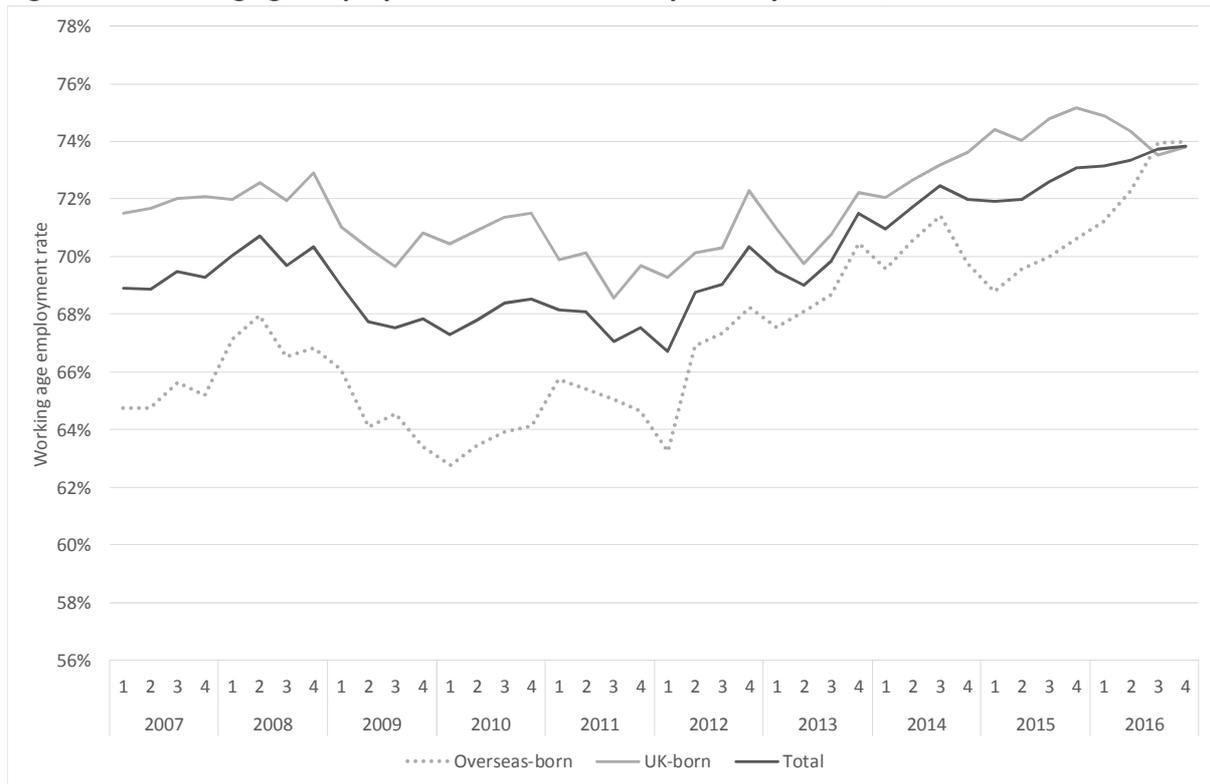
Source: Labour Force Survey.

Figure 4.5: Working age employment rate in rUK by country of birth, 2007-16
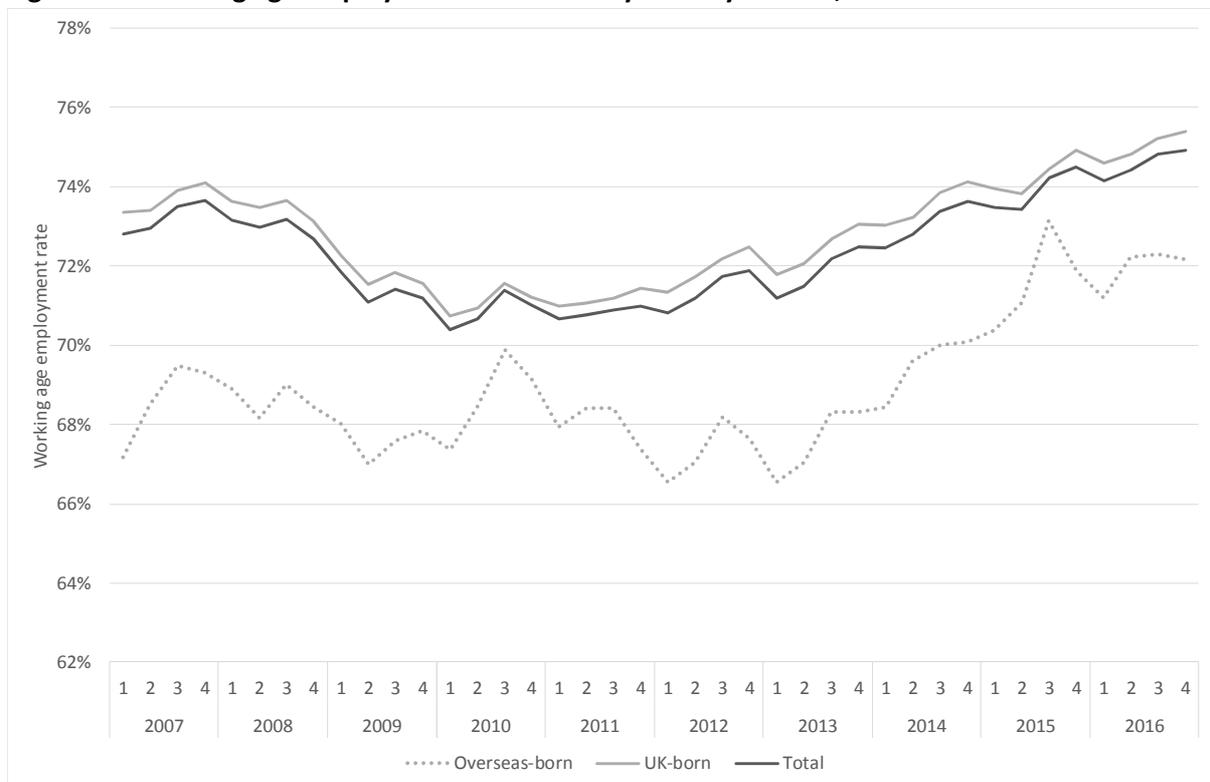
Source: Labour Force Survey.

Of course, immigration to the UK from overseas is only one part of the migration picture. Changes in patterns of internal migration may also effect relative regional wage growth over time. Classical economic theory suggests that migration is the principle mechanism for correcting regional wage disparities over time. However, given the selectivity of interregional migration by skill level, the effect of interregional migration on wages is ambiguous. Skilled migrants may compete with incumbent workers, leading to a decrease in the wages of workers who share those skills; on the other hand, skilled migrants can bolster the productivity of existing workers and firms through a variety of channels (Dickey, 2014; Faggian et al. 2017).

Data on internal migration flows between regions of the UK (regardless of place of birth) is available from National Health Service records. Effectively, individuals are assumed to have moved when they register with a new doctor in a different region. As such there are limitations with such data, as some groups of individuals are less likely to re-register when they move than others. Nonetheless, the data are the most comprehensive available.

Figure 4.7 shows net migration (net migration is in-migration minus out-migration, thus a negative number implies net out-migration) flows from London to UK regions between 2002-15, the earliest that comparable data is available. The data covers the 12 months to March in each year in the figure.

London acts as a source of net migration to other regions (principally to those in the south of England: the SW, East and SW regions each have net in-migration of around 20,000 per year, while other regions have net migration of between -6000 and 6000). But in the year to March 2009 (i.e. the height of the recession), net migration from London halved relative to the previous two years. Analysis of the inflows and outflows (not shown here but available on request) shows that the fall in net migration from London was due in part because of a rise in inflow to London from other regions (London was the only region to experience a rise in inflow in 2009), but more particularly due to a fall in outflow from London.

At UK level, the recession was characterised by a significant fall in the rate of job-to-job flows. By Q2 2014, the rate of job-to-job moves had recovered somewhat, but remained below its pre-recession peak. The fall in net migration from London during the recession and the aftermath may represent a regional manifestation of this phenomenon. Alternatively it may simply cyclical variation in demand for housing space in the capital. Indeed, to the extent that most outmigration from London is to the surrounding east and south east regions – and thus within London's extended housing and labour market area, it is not clear what these headline migration statistics might reveal about wage trends in London itself.

Figure 4.7: Net migration from London, year to March
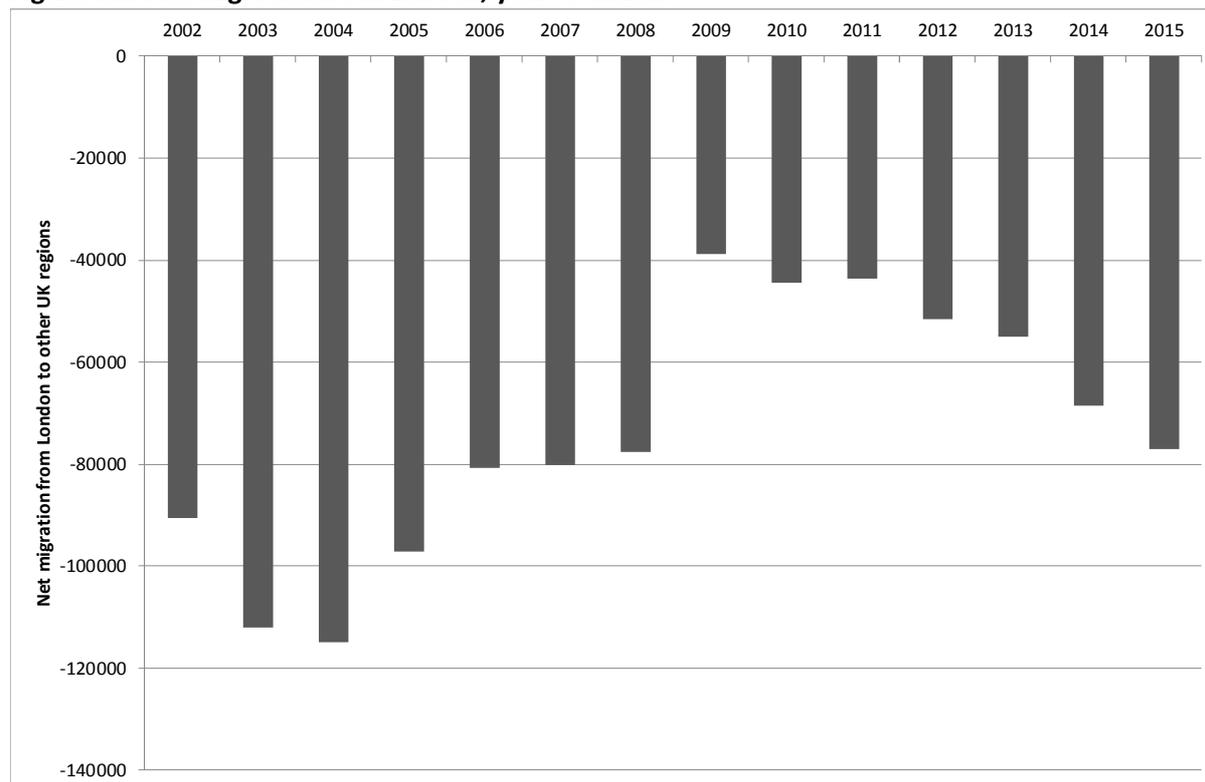
Source: Office for National Statistics, Internal Migration Statistics.

v. Conclusions

Across the UK as a whole, the response to the Great Recession has been unusual in that employment performance has been relatively strong while real wages have fallen. London has experienced an exaggerated version of this general picture, while some regions have seen a weaker employment response.

Although London did not experience particularly smaller declines in employment or output during the economic crisis of 2008-9, it has been at the vanguard of the UK's recovery since the recession. London recovered its pre-recession employment level and its pre-recession employment rate substantially before other regions. And its employment rate is now well above the pre-recessionary level, which cannot be said of any other region.

A number of reasons for London's stronger employment performance have been put forward. These include various structural advantages (including roles for the industrial mix and the boost from the devaluation of the pound on in-bound tourism, with lesser roles for occupational structure and the growth in fraction of wealthy global elites who base themselves for at least part of the time in London), spatially significant (public and private) policies, and the role of specific support to the financial sector.

But despite strong employment performance, wage growth in the capital has been relatively weak since 2012. One potential explanation relates to the role of immigration. Immigrants account for a much greater proportion of employment in London than in rUK, and the employment rate of migrants has increased substantially in recent years. Given that recent migrants to the UK earn less on average than the UK-born, such compositional changes in the labour force might explain London's relatively greater wage falls via compositional effects. However, empirical evidence presented in this paper finds inconclusive evidence of such an effect. A limitation of the analysis is that immigrants are treated as a homogenous group, which is clearly disingenuous.

Moreover, any explanation for weaker wage growth in London has to account for the fact that wage growth has been slower in London across the distribution. There is thus unlikely to be one explanation as to why wages have fallen more rapidly in London than elsewhere in recent years, despite London's stronger employment performance.

Some of the welfare reforms since 2012 might be expected to have a relatively more significant effect on increasing participation rates in London. The local housing allowance that determines Housing Benefit rates is being uprated by inflation rather than an index of local housing rents. Given that the cost of living – certainly in relation to housing costs – is increasing faster in London than the rest of the country, it might be hypothesised that these benefit cuts might have greater 'bite' in London, raising the incentive to participate in the labour force by relatively more in London compared to other parts of the country. Similarly, 49% of families effected by the benefits cap, a cap on the maximum weekly benefit that a family can receive, live in London (DWP, 2012).

Furthermore, it may be that the labour market in the capital is more flexible generally given the structural composition of the economy (lower trade union coverage, less reliance on the public sector and large employers in total employment).

But these are purely speculative hypotheses, and further research is needed to understand the causes of recent labour market trends both nationally, and in terms of regional divergence. From the perspective of fiscal devolution, one important implication of differential labour market performance is what it means for regional income tax liabilities. The subsequent paper considers the role of regional employment and wages in determining income tax liabilities, and asks to what extent devolved governments can influence the key determinants of tax liabilities within their jurisdictions.

5. Tax devolution and accountability for revenue growth: lessons from Scotland

David Eiser and David N.F. Bell

Abstract. Income tax is being devolved to the Scottish Parliament. Future growth in the Scottish Government's budget will reflect the rate of growth of income tax revenues per head in Scotland relative to those in the rest of the UK (rUK). Understanding the factors that influence the growth in income tax revenues is thus increasingly important for policy makers. This paper first considers the factors that determine relative growth of regional income tax revenues per capita, including the roles of labour market factors (wages and employment) and the role of the income tax structure. A key finding is that changes to the structure of income tax have, in recent years, influenced relative growth in regional income tax revenues by at least as much as changes to the underlying economic determinants. The paper then goes on to examine the factors determining relative wage growth at different parts of the wage distribution, given that wages (particularly at the upper end of the distribution) are a major determinant of income tax revenues. This analysis uses the Recentred Influence Function regression method of Firpo et al. (2009) to assess the role of workforce composition factors in explaining why Scotland's wage gap with rUK closed differentially between 2002/3 and 2014/15 at different points of the wage distribution. The results suggest that some compositional factors help explain wage convergence at the median, but that higher wage returns to workers in rUK have limited wage convergence at the 90th percentile.

Key words: tax devolution, wage growth, fiscal federalism

i. Introduction

As has already been discussed, tax decentralisation to Scotland – including 'full' devolution of taxes such as Stamp Duty, partial devolution of income tax, and assignment of a proportion of VAT – is motivated in part by the argument that such decentralisation increases the accountability of Scottish Ministers for the relative performance of the Scottish economy.

Under Scotland's new Fiscal Framework, an adjustment will be made to the Scottish block grant to reflect the value of revenues foregone by the UK Government as a result of transferring each revenue stream to Scotland. In the first year, this 'block grant adjustment' (BGA) will be equal to the revenues raised from the tax to be devolved by the UK government from Scottish taxpayers in the year prior to the tax being transferred. In subsequent years, this BGA will be indexed so that it grows in line with the per capita growth rate of the equivalent tax in rUK.

It follows that the Scottish budget will be better off than it would have been without tax devolution if per capita tax revenues grow more rapidly in Scotland than in rUK. This arrangement raises questions about the extent to which the Scottish Government will be able to influence the growth rate of per capita tax revenues (relative to the growth rate of equivalent rUK revenues).

One way in which the Scottish Government could influence the relative growth rate of revenues is of course to vary tax rates or thresholds differentially from those in rUK. Uncertainties about the elasticity of taxable income with respect to the tax rate means that there are always likely to be political controversies surround the likely effect of tax rate changes on tax revenues.

This paper does not consider the question of the revenue effect of policy change. Instead, it considers what scope the Scottish Government has to influence growth of the tax base through its policies more generally. In other words, the paper considers the scope for achieving differential growth in revenues per capita where the income tax structure does not vary between Scotland and rUK. A reason for choosing this 'no tax policy differential' setting for analysis is that, over the period for which we have data on UK inter-regional tax revenue outturn, no inter-regional tax policy variation has been possible. But the question of whether there has been much temporal variation in regional tax revenue performance – and the reasons for any observed variation over time – is nonetheless useful in informing the question of the extent to which a devolved government might be able to influence the relative growth rate of per capita revenues.

The paper considers this issue specifically in the context of income tax. This decision is largely made on the basis that income tax is, in revenue terms, the most significant of the taxes being transferred to Scotland. But the paper considers the extent to which the conclusions are applicable to the other taxes being transferred.

The paper essentially consists of two parts. In the first part, we consider trends in UK regional income tax revenues per capita, and we examine the possible drivers of trends observed.

In the second part, we consider more specifically the factors that might determine differential wage growth across regions. Wages are likely to be a major determinant of income tax revenue, so the scope for a regional or devolved government to influence wage growth is an important factor in determining a devolved government's ability to influence relative growth in revenues.

The empirical analysis in the second part of the paper examines specifically the factors that caused wages in Scotland to converge with those of rUK during the period since devolution in 1999.

Since establishment of the Scottish Parliament in 1999, median wages in Scotland – which were around 6% lower than the UK median wage in 1999 – have converged with those in the UK. However, Scottish income tax revenues per capita have not converged to anything like the same extent as median wages, and remain significantly below those of the UK as a whole. In essence, this is due to the fact that income tax liabilities have become increasingly concentrated on the top 10% of income earners over time, and whilst Scottish median wages have converged to the UK, wages at the 90th percentile have not converged to nearly the same extent. The analysis in the second half of the paper focusses on understanding what has caused Scottish wages to converge to those of the UK differently at different parts of the wage distribution.

The paper examines the extent to which changes in the relative composition of the Scottish and rUK workforce can explain the convergence at different parts of the wage distribution. We employ a relatively novel approach to decomposing wage differentials at different parts of the distribution, originally developed by Firpo et al. (2007). The method enables analysis of the extent to which wage convergence at different parts of the wage distribution can be explained by changes in the composition of the Scottish workforce relative to that of the rest of the UK (rUK), including demographics, qualifications, and industrial structure.

The remainder of this paper is structured as follows. Section 2 describes the UK income tax system, describes recent trends in regional income tax revenues, and considers the factors that are likely to influence growth in tax revenues per capita at regional level, drawing on both graphical and econometric analysis. Section 3 considers trends in income tax revenues, employment and wages in Scotland specifically. Section 4 describes the decomposition methodology, and Section 5 presents the results. In Section 6 we consider factors other than workforce composition which may influence relative wage differentials between Scotland and rUK. Section 7 concludes.

ii. Income tax: structure, determinants, and revenues in UK regions

Income Tax is the largest UK tax in revenue terms, accounting for 25% of UK Government revenues (£170bn) in 2015/16 (Grace et al. 2015). The primary forms of taxable income are earnings from employment, income from self-employment and unincorporated businesses, pension income, income from property, and income from savings and investments (including property, bank and building society interest, and share dividends).

The basic structure of income tax is that there is a Personal Allowance (an amount of income on which is not liable for tax) which for most people in 2015/16 is £11,000. A basic rate of 20% is applied to income up to £43,000, a higher rate of 40% applies to income between £43,000 and £150,000, and an Additional Rate of 45% applies above £150,000. A number of tax reliefs are available (e.g. on pension contributions and charitable donations). Income from savings and dividends is subject to slightly different rates.

Income tax revenues per capita vary substantially across UK regions (Figure 5.1). Across the UK as a whole, income tax raised £169bn in 2015/16, equivalent to £2,600 per person. On a regional basis, revenues per capita varied from £1,450 in Northern Ireland to £4,680 in London.

Revenues per capita were increasing until 2007/8, fell markedly until 2012/13, and have recovered only slowly since. Revenues are determined by both the tax base and tax policy. The principle determinants of the tax base are wages and salaries, and employment. Part of the explanation for the fall in total income tax revenues (and revenues per capita) across the UK relates to the trends discussed in the previous chapter – the fall in employment during the recession, and the subsequent sluggish wage growth since then.

Figure 5.1: Income tax revenues in UK regions (2016/17 prices)
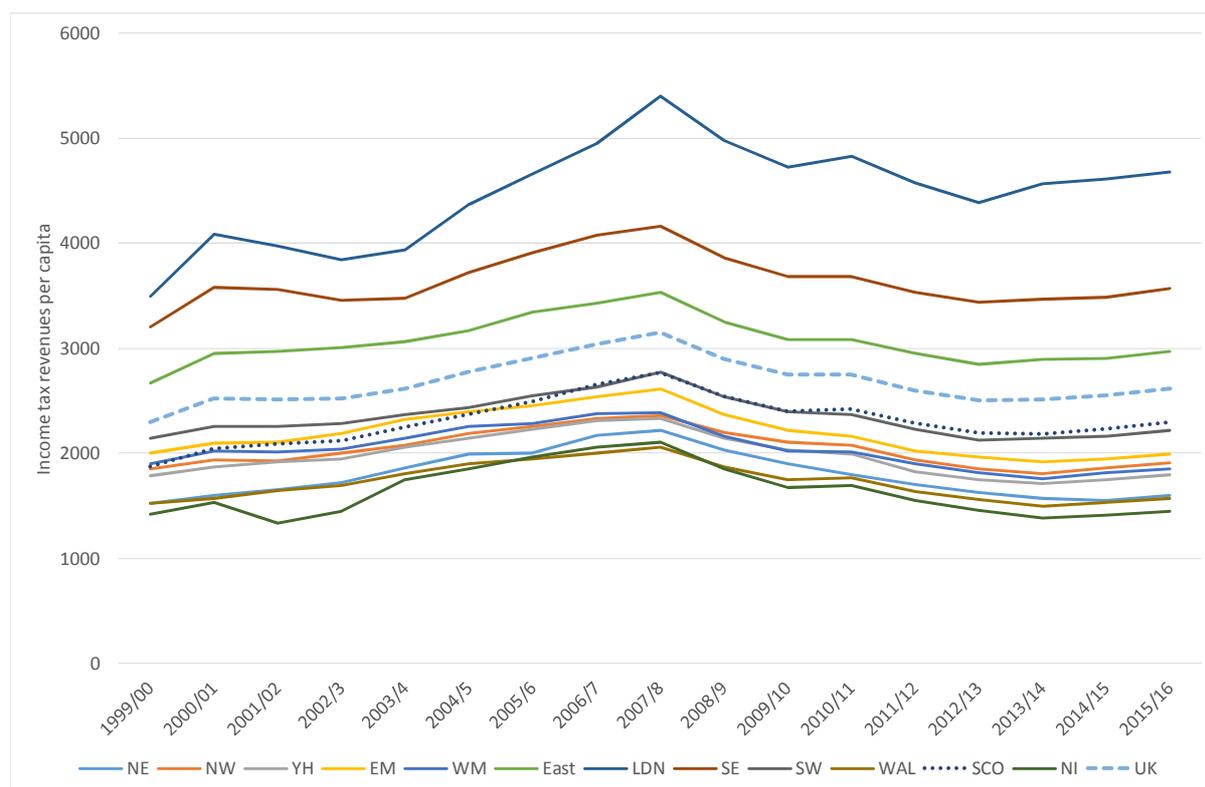
Source: ONS, Country and Regional Public Sector Revenues, May 2017

But changes in income tax policy have played a role too. Two key changes are worth noting. Since 2007/8 the Personal Allowance has increased significantly in real terms. Between 2007/8 and 2008/9 it increased by 12% in real terms, whilst between 2007/8 and 2015/17 it increased by 65% in real terms. This represents a substantial tax cut (HM Treasury's tax ready reckoner statistics suggest that each £100 increase in the personal allowance costs the Exchequer around £600m).

Alongside the real terms increase in the Personal Allowance, tax rates on the highest earners have increased. During the last parliament, the Higher Rate threshold was reduced in real terms by around 13%. Furthermore an Additional Rate of tax was introduced in 2010/11, initially at 50% before being reduced to 45% in 2013/14.

The increases in the Personal Allowance combined with reduction in Higher Rate threshold and introduction of the Additional Rate have together resulted in income tax liabilities becoming increasingly concentrated on higher earners. The proportion of the adult population who pay income tax has fallen from 66% in 2007/8 to 56% in 2015/16. The proportion of income tax paid by the top 1% of taxpayers increased from 21.3% to 27.5% between 1999/2000 and 2015/16, whilst the proportion paid by the top 10% increased from 50.3% to 58.9%.

Figure 5.2 shows that London tended to pull away from other regions in terms of income tax revenues per capita between the mid-2000s and 2014/15. This might reflect stronger employment rate growth in London post-recession, as described in the previous chapter. But it is also likely to reflect the tax policy changes which have resulted in tax liabilities becoming increasingly concentrated on higher earners.

Figure 5.3 plots the coefficient of variation across the 12 UK regions for income tax revenues per capita and three variables that are likely to influence income tax revenues: the 16+ employment rate, the median wage (for full-time workers) and the wage at the 90th percentile (also for full-time workers). The coefficient of variation for income tax revenues per capita increased by some 16 percentage points over the ten years from 2003/4 to 2013/14. Over this period there was essentially no change in the coefficient of variation associated with the economic determinants. This does seem to suggest that changes in the distribution of income tax revenues per capita over regions owes more to tax policy changes than it does to divergence in economic performance.

Figure 5.2: Index of per capita income tax revenues in UK regions (UK=1)
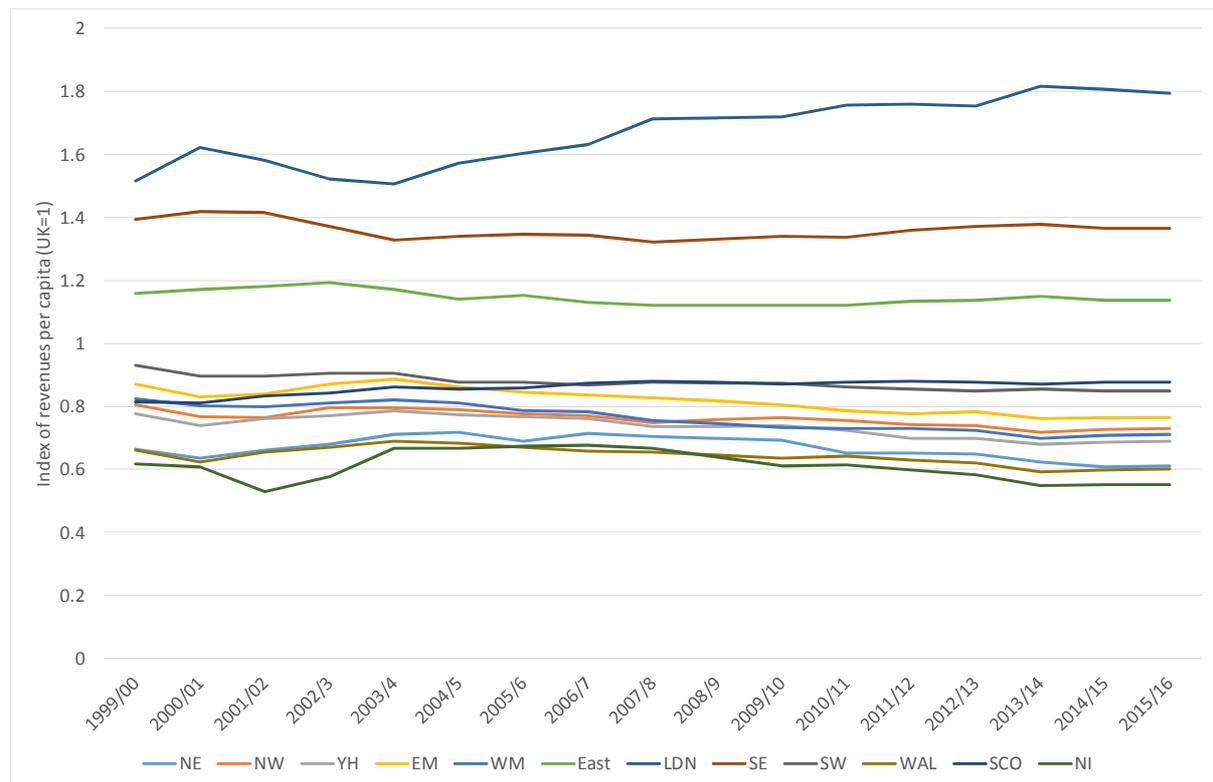
Source: ONS, Country and Regional Public Sector Revenues, May 2017

Figure 5.3: Coefficient of variation: income tax revenues per capita and determinants
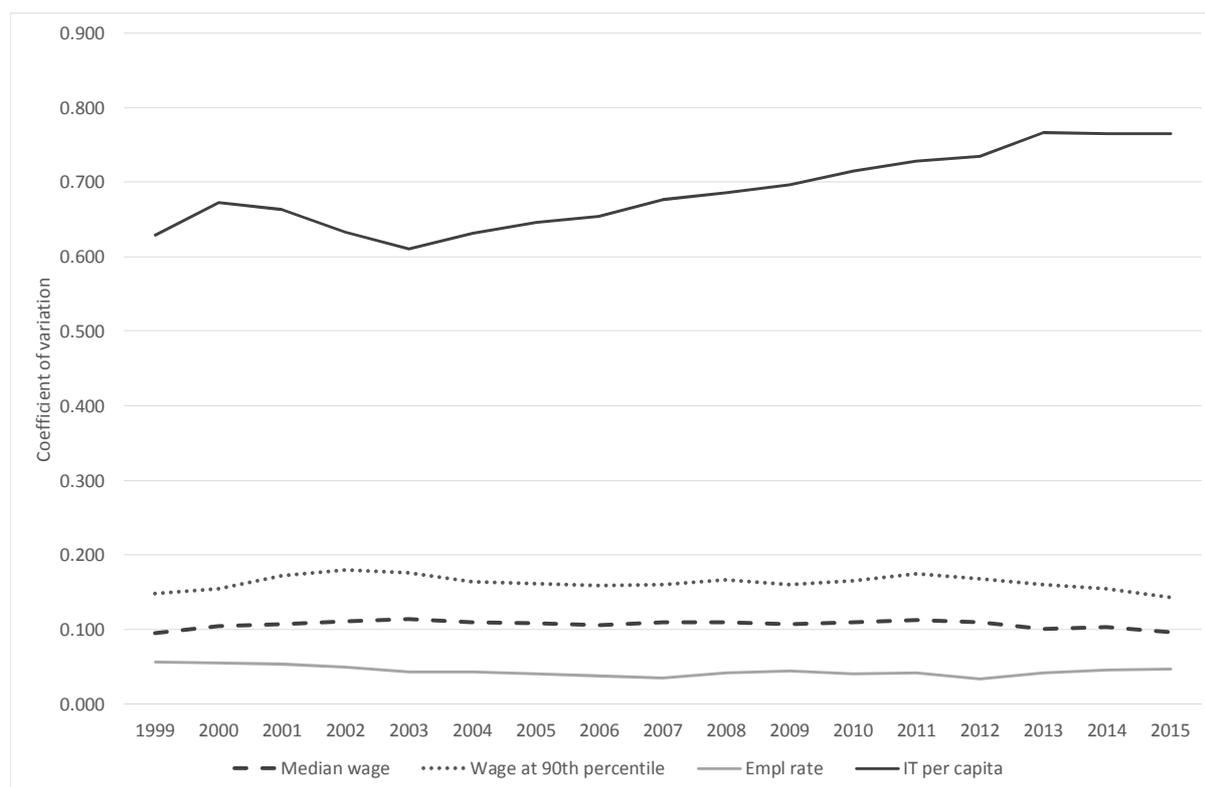
Source: ASHE, LFS, ONS (Country and Regional Public Sector Revenues)

Most of the increase in the income tax coefficient of variation is attributable to London. If London is excluded, the coefficient of variation on income tax increases by seven percentage points, as opposed to 16. Between 2003/4 and 2013/14, London's income tax revenues per capita, expressed as an index relative to the UK as a whole, increased from 150% to 181%. Yet over the same period, median wages in London fell from 129% to 126% expressed as an index, whilst wages at the 90th percentile fell from 145% to 141%. On the other hand, London's 16+ employment rate increased relative to the UK over the period, from 100% to 105%.

The substantial increase in London's relative per capita tax liability, at a time when its relative wage has been falling, further suggests a strong role for tax policy in determining relative regional growth in per capita tax revenues.

In an attempt to shed further light on the determinants of per capita income tax revenues, a panel regression was run of the form:

$$Log(IT_{rt}) = \alpha_r + \beta_1 Log(ER_{rt}) + \beta_2 Log(W50_{rt}) + \beta_3 Log(W90_{rt}) + \beta_4 t + u_{rt}$$

Where the dependent variable is the log of income tax revenues in region r and in year t, ER is the 16+ employment rate, W50 is the median wage of full-time workers, and W90 is the wage of full-time workers at the 90th percentile. α_r is a regional fixed effect; β_1, β_2 and β_3 are parameters to be estimated, t is a time trend and u is a random error. The 16+ employment rate is used, rather than the 16-64 rate, as this partially captures changes in employment due to demographic changes.

The regression is run for the 12 UK regions over the period 1999/00 – 2015/16. Results are shown in Table 5.1, with robust standard errors reported in brackets. The model is run over three periods. Model (1) includes the full 1999/00 to 2015/16 period including a time trend, where this time trend is split for periods before and after 2007/8. The time trend is intended to capture trends that are not captured within the explanatory variables themselves, potentially including for example the trend towards a higher proportion of the self-employed in total employment. The time trend is split into a pre and post 2007/8 trend, reflecting significant changes in the underlying trend of economic determinants (such as the slowing in the growth of productivity, discussed in the previous chapter) and the marked changes to the structure of income tax, as already highlighted.

Model (2) and Model (3) are run over the each of the two sub-periods distinctly, with a simple linear time trend.

In Model (1) the elasticity of tax revenue per capita with respect to the employment rate is approximately unity; a one percent increase in the employment rate is associated with slightly more than a one per cent rise in income tax revenues per capita. The coefficient on median wage is not significant, whilst the coefficient on the wage at the 90th percentile is significant at the 10 per cent level. It implies that a one percent rise in the 90th percentile of a full-time worker's weekly wage is associated with a 1.17% rise in tax revenues per capita.

When the model is run over the earlier period (Model 2), the coefficient on the employment rate increases to 1.3, whilst the coefficient on the 90th percentile wage is of a counterintuitive sign. This presumably reflects a period of strong growth in both employment and tax revenues per capita, at a time when the tax system was less dependent on high earners, given its structure. In contrast the later period sees a relatively weaker role for the employment rate and a stronger role for wage growth at the 90th percentile, although the coefficient on this variable remains significant at the 10% level only.

Table 5.1: Regional income tax revenues per capita, the employment rate and wages

	(1) 1999/00 – 2015/16	(2) 1999/00 – 2007/8	(3) 2008/9 – 20015/16
Log [ER (t)]	1.044	1.314	1.007
	(0.451)***	(0.246)***	(0.357)**
Log [W50 (t)]	0.381	0.916	-0.328
	(0.659)	(0.764)	(0.531)
Log [W90 (t)]	1.174	-1.114	1.131
	(0.608)*	(0.597)*	(0.556)*
Trend	0.034	0.043	-0.014
	(0.006)***	(0.003)***	(0.004)***
Post 2007/8 trend	-0.060		
	(0.010)***		
R^2	0.93	0.90	0.94
Sample size	204	108	96

*Notes: Robust standard errors in parentheses. *, ** and *** indicate significance at the 0.1, 0.05 and 0.01 level respectively*

iii. Looking more closely at Scotland

The discussion thus far has considered trends in income tax revenues per capita across UK regions generally, but the focus of this paper is on Scotland particularly. Between 2000/01 and 2006/7, Scotland did experience somewhat more rapid growth in its tax revenues per capita than was observed in rUK (Figure 5.4). The implication is that, had income tax been devolved to Scotland in 1999 under the terms of its current Fiscal Framework arrangements, it would have been about £800m better off than it otherwise was without tax devolution (Bell et al. 2015). Since 2007 however, Scotland's income tax revenues per capita have not converged any further with those of rUK.

Figure 5.4: Scottish income tax revenues per capita (UK=1)
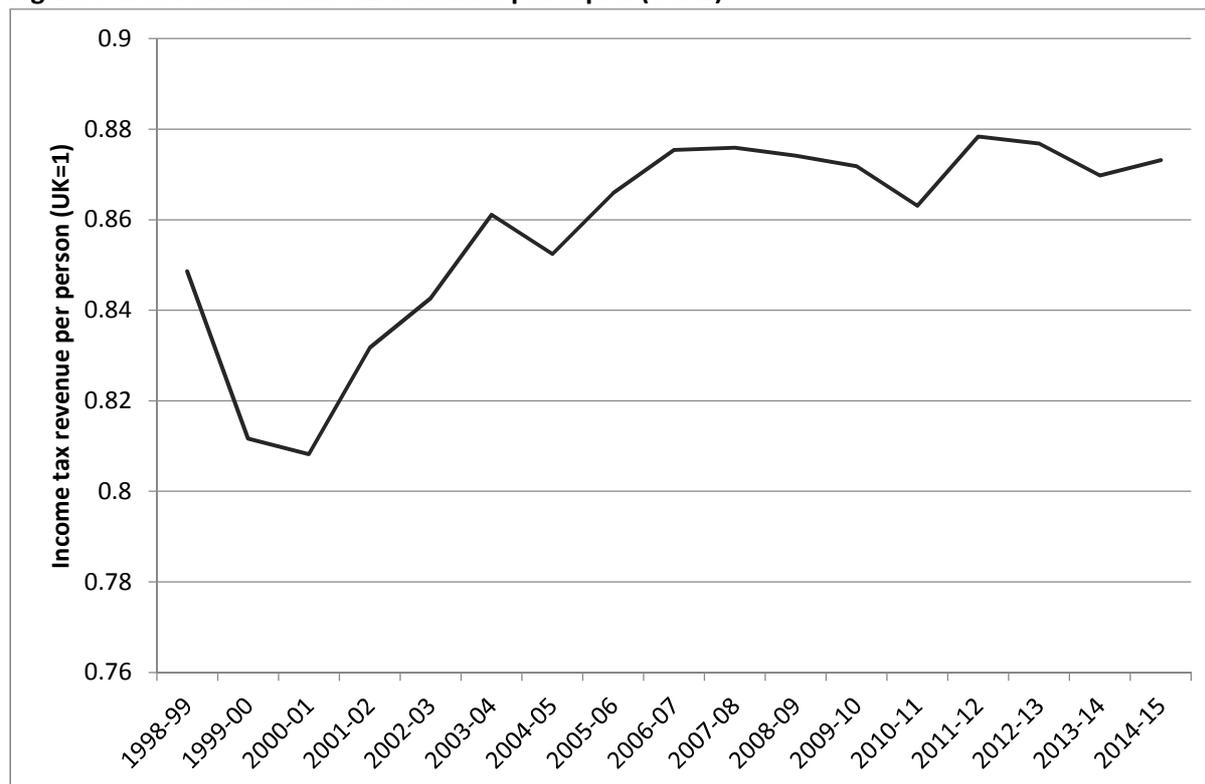
Source: Government Expenditure and Revenue Scotland

This convergence in Scottish tax revenues per capita coincides with some convergence between Scottish and rUK employment rates (between 1999 and 2009), and – since 2004 – convergence between Scottish and rUK median wages.

Lack of convergence in Scotland's income tax revenues per capita since 2006/7 seems likely to be due to a combination of some reversal in its relative employment rate position since the recession, combined with a lack of convergence in wages at the 90th percentile in Scotland with those in rUK.

Figure 5.5 shows how median weekly wages have evolved across UK regions since 1997, expressed relative to the UK median. Between 1997-2004, median wages in Scotland hovered around 93-94% of the UK median. But since 2004, median wages in Scotland have closed the gap to the median. By 2015, the median wage in Scotland was the same as the median wage in the UK.

No other UK region experienced an equivalent change in its fortunes over the period (although the North East has experienced some convergence in its earnings from a lower base, whilst East and to a slightly lesser extent the South East have experienced secular falls in their median wage relative to the UK's). London experienced rapid relative wage growth between 1999 – 2005; this relative wage growth slowed over the period to 2012, following which London's median wage fell substantially relative to that of other regions.

London is clearly an outlier in terms of its wage level. Part of this differential undoubtedly reflects the industrial and occupational structure of the city and its tendency to attract the most highly skilled and productive workers, of given quality or attributes (Gibbons et al. 2010). But a substantial literature also emphasises that workers in large urban centres earn more than workers in their non-urban counterparts, and that this premium can be substantially higher in the case of a 'global city' such as London. Explanations have included that agglomeration tends to increase the productivity of workers (Rice et al. 2006) as a result of positive urbanisation externalities including reduced transport and input costs, technology and knowledge spillovers, and proximity to consumers (Fujita et al. 2001). It has also been hypothesised that human capital accumulation may be faster in cities (Glaeser 1999), and furthermore that cities enhance the probability of better matching between firms and workers.

The narrowing of Scotland's median wage gap is evident across many segments of the labour market. Median wages for Scottish males and females closed the gap to those in the UK for weekly pay and hourly wages over the period 2004-2015, both for full-time workers only (Figure 5.6) and all workers (Figure 5.7). Moreover, the Scottish median wage gap narrowed by roughly the same extent in both the public and the private sector (Figure 5.8). (Figure 5.8 is based on data from the LFS rather than ASHE; as such it is based on fewer observations and there is consequently more noise in the data, although the convergence trend is statistically significant.)

However, although convergence between Scottish and UK median wages is quite pronounced, convergence is less pronounced at other parts of the wage distribution (Figure 5.9). Wages at the tenth percentile have tended to be higher in Scotland than in the UK as a whole, while at the 25th percentile, convergence to the UK wage has occurred to about the same extent as it has done for median wages. In the upper half of the distribution however, convergence has been less marked. Wages at the 75th percentile remain 3% lower than those in the UK, whilst wages at the 90th percentile are 6% lower than those in the UK.

This lack of convergence in wages at the top of the distribution is likely to be a major factor accounting for the fact that there has been little convergence in income tax revenues per capita between Scotland and the UK as a whole, as noted previously. Berthier (2015) shows that, whilst Scotland's share of UK income tax revenues paid on incomes up to £30,000 is proportionate to its share of UK population (8.3%), Scotland's share of income tax revenues paid on higher incomes is lower than its population share. For example, Scottish taxpayers account for only 6.2% of the UK's income tax paid on incomes between £150,000 and £200,000, and just 3.8% of the UK's income tax paid on incomes above £200,000. Taxpayers earning over £200,000 accounted for 20% of UK income

tax revenues in 2012/13, but only 11% of Scottish income tax revenues (Berthier, 2015). Income tax paid at the Additional Rate accounted for 15.2% of UK revenues (HMRC, 2016) but just 7.7% of Scottish revenues in 2012/13 (Berthier, 2015).

In the remainder of this paper, we focus more specifically on the question of Scotland's wage growth relative to rUK. In particular, we consider what might account for the convergence in Scotland's median wage with that of rUK between 2004/5 and 2014/15, and why similar wage convergence has not occurred at the top end of the wage distribution.

The rationale for this focus on relative wage growth is because of the link from wage growth to income tax revenues, which are now devolved to the Scottish Parliament. Despite the fact that the empirical evidence presented thus far does not suggest a very strong link between wages and per capita tax liabilities, it seems likely that this apparent anomaly can be accounted for to a large extent by tax policy changes. Indeed, the forecast of wage growth remains arguably the key economic determinant in the OBR's fiscal forecasts; and it is widely acknowledged that comparatively large errors in recent years in forecasting UK income tax revenues are largely attributable to consistently overestimating future UK wage growth (OBR 2016).

Figure 5.5: Median weekly wage in UK regions relative to UK median
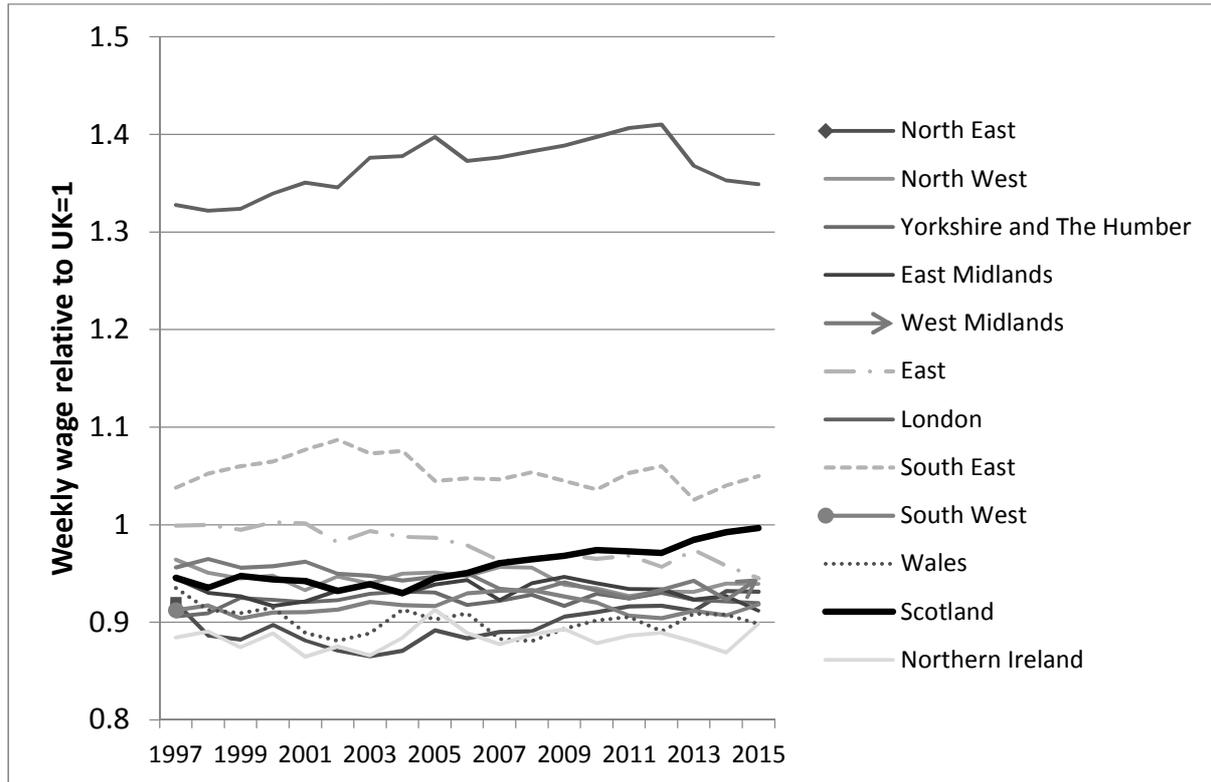
Source: Annual Survey of Hours and Earnings

Figure 5.6: Scottish median wage relative to UK equivalent (full-time workers)
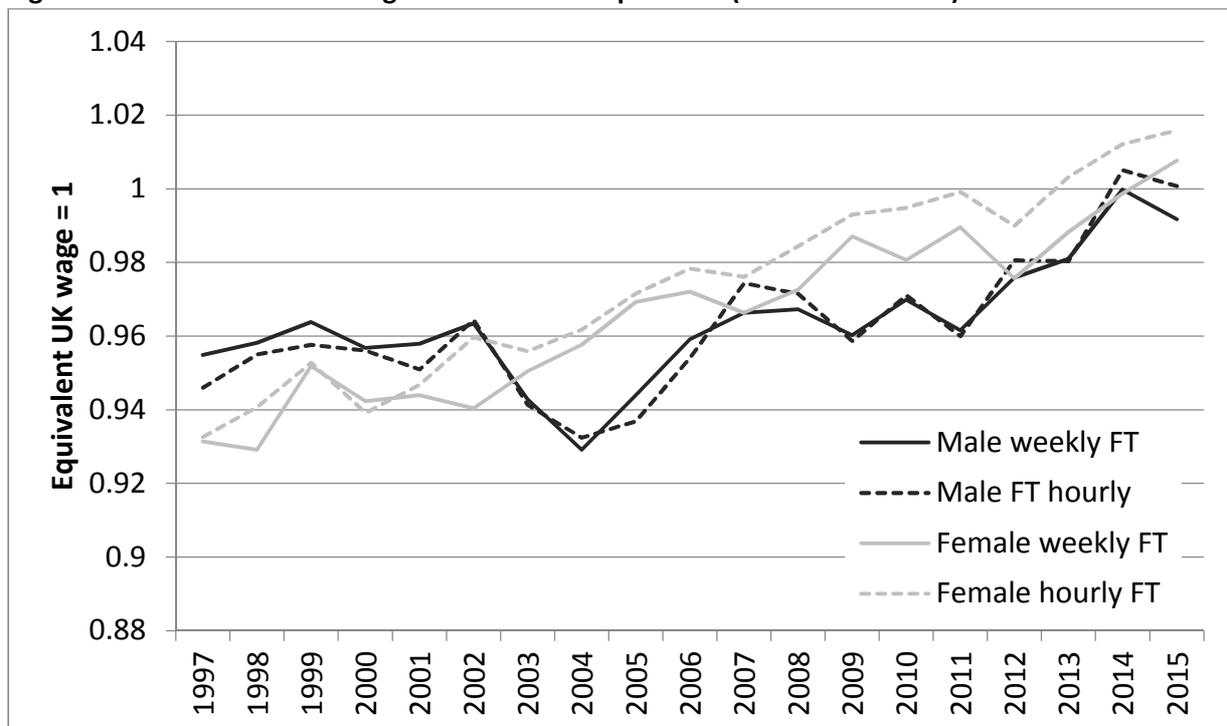
Source: Annual Survey of Hours and Earnings

Figure 5.7: Scottish median wage relative to UK equivalent (all workers)
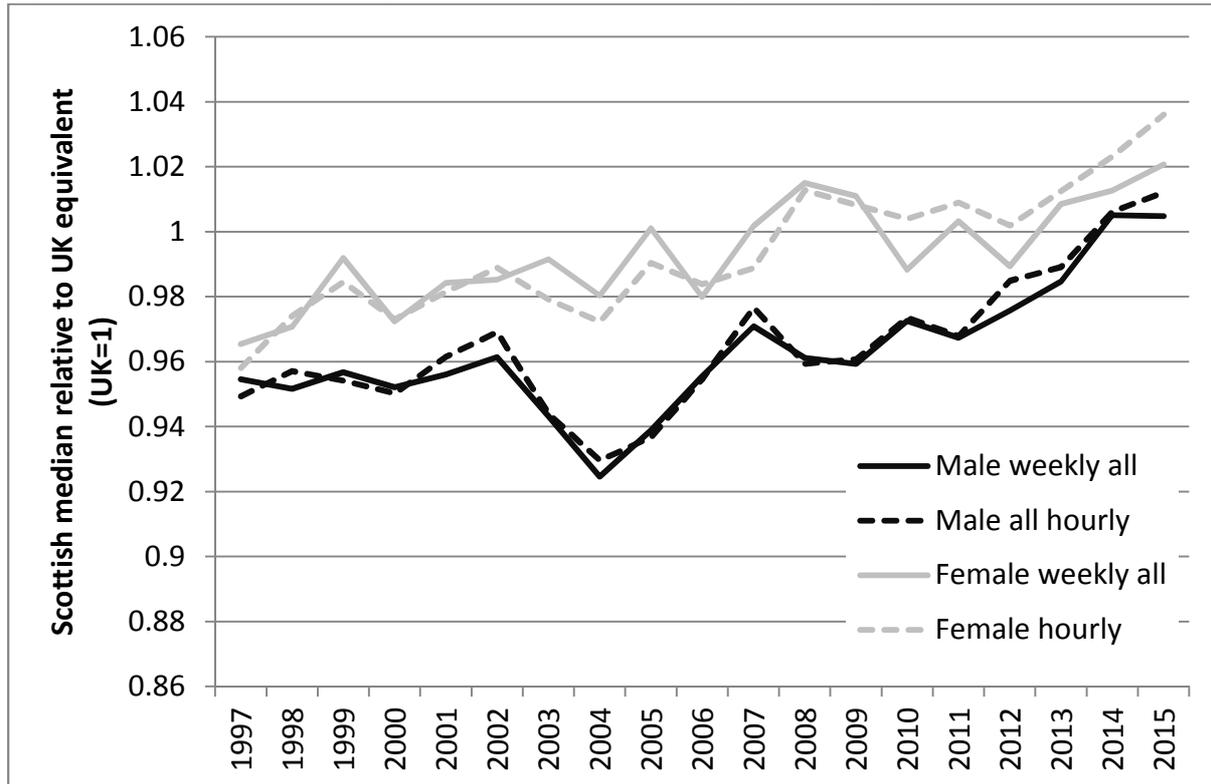
Source: Annual Survey of Hours and Earnings

Figure 5.8: Scottish FT weekly median wage in public and private sector relative to UK equivalent
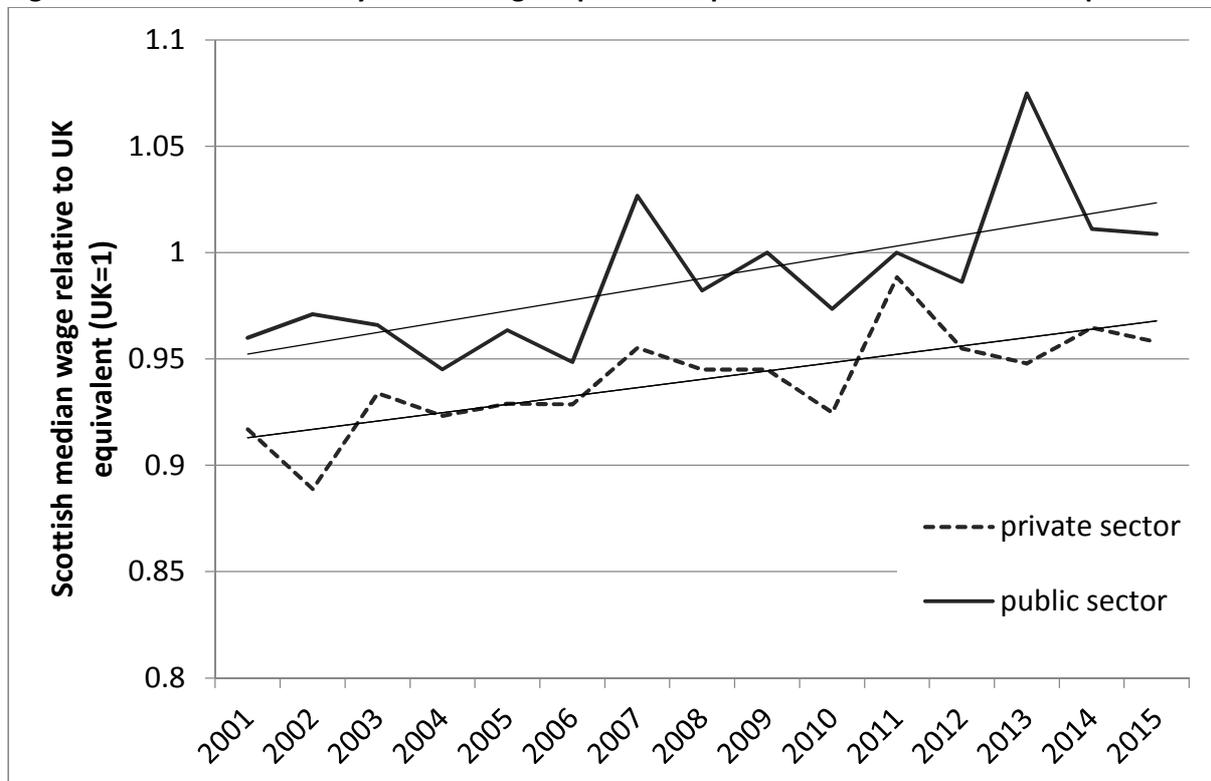
Source: Annual Survey of Hours and Earnings

Figure 5.9: Scottish FT weekly wage relative to rUK equivalent at five percentiles
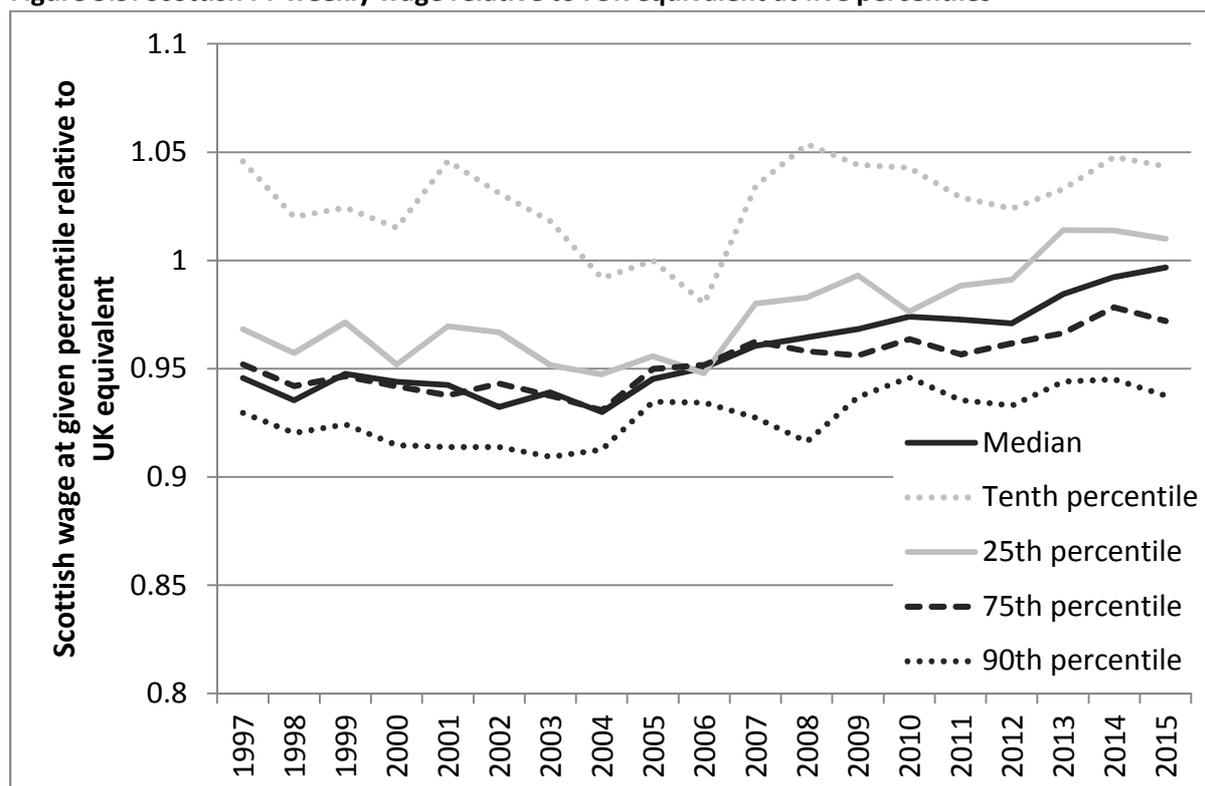
Source: Annual Survey of Hours and Earnings. Note: the average coefficient of variation (the ratio of the standard error to the estimate) over the 1997-2015 period is 0.8% at the tenth percentile, 0.7% at the 25th percentile, 1% at the median, 1.2% at the 75th percentile, and 3.2% at the 90th percentile.

iv. Approach

We now turn to the main empirical focus of this paper: understanding the extent to which compositional factors might account for wage convergence at different points in the wage distribution.

Attempts to explain how much of a difference in mean wages between two groups can be explained by the composition of those groups (by age, qualification, industry, etc.) often relies on the Oaxaca-Blinder decomposition. The Oaxaca-Blinder approach decomposes the difference in mean earnings between two groups into a part which can be explained by differences in composition of the two groups (i.e. what proportion of the difference in mean wages is due to differences in mean levels of education, age, etc.), and a part which remains unexplained, i.e. which reflects differences in the returns to particular characteristics in each group.

The Oaxaca-Blinder approach involves regressing the outcome variable Y (wages) on covariates X for the two groups of interest (Scottish workers and rUK workers). The difference in mean wage between the two groups $\hat{Y}_{Scot} - \hat{Y}_{rUK}$, can be decomposed as follows:

$$\hat{Y}_{Scot} - \hat{Y}_{rUK} = \hat{\beta}_{Scot}(\overline{X}_{Scot} - \overline{X}_{rUK}) + (\hat{\beta}_{Scot} - \hat{\beta}_{rUK})\overline{X}_{rUK}$$

The first term on the right-hand side of the equation represents the composition effect, i.e. the effect that can be explained by the difference in the mean characteristics of workers $\overline{X}_{Scot} - \overline{X}_{rUK}$, assuming that those workers received the same 'returns' as those in Scotland. The second term represents the wage structure effect, i.e. the part of the wage gap that can be explained by differences in the returns to characteristics, assuming the average characteristics of rUK workers.

The limitation of the Oaxaca-Blinder decomposition is that it can only be used to analyse differences in mean earnings between two groups. As noted above, median wages in Scotland have closed the gap to those in rUK much more than has been the case for wages at the 90th percentile. What we would like to do is to assess the extent to which compositional differences in the labour force between Scotland and rUK can explain differences in wages at different parts of the wage distribution.

To do so, we make use of the technique developed by Firpo et al. (2007). This approach involves two broad stages. In the first, unconditional quantile regressions are estimated to identify the influence of explanatory variables X on the quantile q_τ of the outcome variable Y (which in our case is the log of the weekly wage). The Recentred Influence Function (RIF) regression approach proposed by Firpo et al. (2009) is used for this stage.

In the second stage, the raw earnings differential between Scotland and rUK for different quantiles is decomposed into a composition effect and a wage structure effect, using the approach proposed by Firpo et al. (2007). The strength of the Firpo et al. method is that the composition effect and returns effect can then be further decomposed into the contribution of individual covariates. This allows us to identify the specific characteristics that drive differences in Scottish and rUK wages at different parts of the wage distribution. The Firpo approach has been applied to examine issues associated with the male-female wage gap in China (Chi and Li, 2008) and regional wage differences in Portugal (Galego and Pereira, 2013).

Stage 1: Unconditional quantile regression
In this paper we are concerned with estimated unconditional quantile effects rather than the more commonly estimated conditional quantile effects. We now offer some brief discussion of the distinction between the two.

Conditional quantile regressions focus on the conditional quantile of an individual, which is that individuals' position in a virtual distribution in which all individuals are assumed to have the same observed characteristics. For example, if individuals would differ only with respect to their education

level, the conditional quantile of a low-educated person would be his/her earnings quantile among all low-educated individuals, whereas the conditional quantile of a highly-educated person would be his/her earnings quantile among all highly-educated persons.

Conditional quantile regressions provide an estimate of the return to a certain characteristic, where the return varies across individuals based on the conditional quantile into which they fall. For example, the conditional quantile regression approach can be used to estimate the impact on an individual's earnings of gaining a particular qualification, holding all other factors constant. Specifically, the conditional quantile approach assumes that the conditional quantile of the individual remains the same when the characteristic changes, i.e. what is the impact of moving from the tenth percentile of individuals without a degree to the tenth percentile of individuals who have a degree?

However, we may be interested in understanding what the impact of gaining a qualification is not just on a subgroup with specific covariates (conditional effects) but on a given population with individuals with different characteristics. Unconditional quantile regressions focus on the unconditional quantile of an individual, which is his/her earnings quantile in the overall earnings distribution, abstracting from observed and unobserved characteristics.

Unconditional quantile regressions estimate the effect of a small change in workers' characteristics on a given quantile of the overall distribution. The method can thus be used to explore, for example, the impact of raising the proportion of the workforce with a given qualification on a particular quantile of the overall wage distribution. Specifically, the coefficients from an unconditional quantile regression represent the approximate marginal effects of the explanatory variable on a given quantile, q_τ. Thus whereas a conditional quantile regression captures only the within-group effect of a given covariate (i.e. what is the effect of gaining a qualification on the wages of an individual at a given quantile), unconditional quantile regressions also capture the between group effect which recognises that an increase in the proportion of the workforce with qualifications will influence the overall distribution of wages.

Given that the objective of this paper is to compare the evolution of the overall wage distribution in Scotland relative to rUK, the unconditional quantile regression approach is clearly the most appropriate.

We estimate unconditional quantile regressions for workers in Scotland and rUK using the two-step procedure proposed by Firpo et al. (2009). The first step involves the estimation of a so-called recentered influence function (RIF) for each individual.

An influence function represents the influence of an individual observation on a distributional statistic of interest, such as a quantile, in response to a change in the underlying distribution. The influence function is 'recentred' so that the mean of the recentred influence function corresponds to the statistic of interest. The RIF for a given quantile q_τ is given by the expression:

$$RIF\ (Y;\ q_\tau, F_Y) = q_\tau + \left(\tau - \frac{I\{Y \leq\ q_\tau\}}{f_Y(q_\tau)} \right)$$

Where $f_Y(q_\tau)$ is the marginal density of Y at the point q_τ, estimated by kernel methods; q_τ is the sample quantile; $I\{Y \leq q_\tau\}$ is an indicator function indicating whether the value of the outcome variable is below q_τ.

In the second step, the RIF is regressed on the explanatory variables X (i.e. the dependent variable Y is replaced by the corresponding RIF). Firpo et al. (2007) show that the estimated coefficients represent the approximate marginal effects of the explanatory variables on the wage quantile q_τ.

Stage 2: Decomposing earnings differences for a given quantile

The second stage involves decomposing earnings differences at a given quantile into a composition effect and a wage structure effect. This is done by estimating a counterfactual earnings distribution, which represents earnings that Scots workers would have received had they received the same returns to labour market characteristics as rUK workers.

Then, the composition and structure effects are further decomposed into the contribution of each individual covariate, making use of the RIF unconditional quantile regressions. In a traditional Blinder-Oaxaca decomposition, this counterfactual would be estimated by applying rUK wage returns to covariates mean for Scotland, $\overline{X}_{Scot}\hat{\beta}_{rUK}$ (or vice versa). However, when the conditional expectation of wages is non-linear, another approach is needed.

The approach proposed by Firpo et al. is to re-weight the sample and apply the RIF regression to the re-weighted sample. The re-weighting factor is defined as:

$$\varphi_i = \left[\frac{1-p(X_i)}{p(X_i)}\right] \cdot \left[\frac{p}{1-p}\right]$$

where p(X) is the probability of a worker being a Scot given individual attributes X. This probability is calculated by pooling data from both regions and estimating a logit or probit model.

The estimated re-weighting factor is then applied to the data in rUK to calculate a counterfactual wage distribution. The RIF regression is estimated for this counterfactual distribution. Having estimated RIF regressions for Scotland, rUK and the counterfactual distribution, it is possible to obtain a wage decomposition for any unconditional quantile (τ):

$$\hat{q}_\tau(Y_{Scot}) - \hat{q}_\tau(Y_{rUK}) = \left\{ \overline{X}_{rUK}\left(\hat{\beta}_c - \hat{\beta}_{rUK}\right) + \hat{R}_\tau^s \right\} + \left\{ \left(\overline{X}_{Scot}\hat{\beta}_{Scot} - \overline{X}_{rUK}\hat{\beta}_c\right) + \hat{R}_\tau^c \right\}$$

Where $\hat{q}_\tau(Y_{Scot}) - \hat{q}_\tau(Y_{rUK})$ represents the raw gender earnings differential at the τth quantile and $\overline{X}$ represents the vector of covariate averages. Since is from the counterfactual distribution which assumes the Scottish returns to labour market characteristics for rUK workers, $\hat{\beta}_c - \hat{\beta}_{rUK}$ measures the wage difference in the returns to labour market characteristics, and thus $\overline{X}_{rUK}\left(\hat{\beta}_c - \hat{\beta}_{rUK}\right)$ represents the wage structure effect. $\left(\overline{X}_{Scot}\hat{\beta}_{Scot} - \overline{X}_{rUK}\hat{\beta}_c\right)$ represents the composition effect, i.e. the wage difference at the τth quantile due to differences in workforce composition. $\hat{R}_\tau^s$ and $\hat{R}_\tau^c$ are the estimates of the approximation errors corresponding to the structure and composition effect. These errors results from the linear specification assumed by the RIF regression functions.

In order to test whether the effects of covariates are significant for the different elements of both the composition effect and wage structure effect, standard errors can be estimated by bootstrapping.

Data

To analyse differences in earnings between Scotland and rUK, we use data from the Labour Force Survey (LFS). The LFS is a quarterly survey of around 100,000 individuals (of whom around 8,000 are Scottish residents). For those individuals in employment, the LFS asks respondents to provide their gross weekly wage. The survey also collects detailed information on individual background (age, qualifications, family composition, etc.) and employment type (industry, occupation, etc.).

We analyse differences between Scottish and rUK wages for full-time workers, for males and females separately, at both the median, and the 90th percentile. We compare differences between Scottish and rUK wages in two periods: 2002/3, and 2014/15. We combine two years of data (eight quarters) to maximise sample size.

Note that LFS respondents are interviewed over five waves, with one wave corresponding to a quarter. Thus any given respondent is included in five consecutive quarters. However, wage questions are only included in wave 1 and wave 5, and so by definition, all individuals in our sample are drawn from one of these two waves. The implication is that there is some duplication in the pooled sample. Specifically, there is a risk that some individuals interviewed in wave 1 during the first four quarters of our two-year subsample may be re-interviewed in wave 5 in the second four quarters of the two-year subsample (depending on sample attrition). But individuals interviewed in wave 5 during the first four quarters, or in wave 1 during the second four quarters, will not be duplicated in the sample. This duplication may bias downwards the standard errors.

The dependent variable in the regressions is the log of the real weekly wage, where wages are deflated using the Consumer Price Index. Explanatory variables include four dummy variables for age (excluded category 'Aged 40-49'); five dummy variables for qualification (excluded category 'GCSE or equivalent'); dummy variable equal to one if born outside the UK; dependent variable equal to one if there are dependent children in the household; dummy variable equal to one if respondent is single; four dummy variables for occupation[12] (excluded category 'professional and managerial occupations'); 13 dummies for industry[13] (excluded category 'public administration'); and two dummies for workplace size (excluded category 'between 12-249 employees'). The mean value of each variable for Scotland and rUK in each time period is shown in Table 5.2.

The change in the occupational classification during the study period is clearly problematic, as it results in a large number of occupations being classified under a few very broad headings. There is no straightforward solution to this issue. Occupations could be grouped by wage decile, but this is unlikely to be particularly informative given that the distribution of occupations within deciles has changed over time.

---

[12] We take occupation defined at the 2-digit level in the Standard Occupational Classification (SOC) and group occupations into one of five broad categories: professional and managerial occupations; intermediate occupations; routine occupations; manual occupations; and service occupations, following the classification strategy of Holmes and Mayhew (2012). There was a change in the SOC between our two periods, with SOC2010 replacing SOC2000. The main implication of the revision was to reclassify many 'managerial' jobs as 'professional' occupations. By combining professional and managerial occupations into one group, we thus aim to minimise the effect of changes in occupational definitions between the two categories. Furthermore, it should be remembered that our main objective is not to compare explanations for wage changes over time, but to examine what factors caused wages to differ between Scotland and rUK in each of the two periods separately.

[13] There was a change in the Standard Industrial Classification in 2007, from SIC1992 to SIC2007. However, LFS data enables us to classify an individual's industry of work post 2007 into the 1992 classification. Thus we use a consistent set of SIC1992 industrial classifications in both time periods.

Table 5.2: Mean value of independent variables

	Males				Females			
	2002/3		**2014/15**		**2002/3**		**2014/15**	
	rUK	**Scot**	**rUK**	**Scot**	**rUK**	**Scot**	**rUK**	**Scot**
Age:								
Aged 16-29	20.2%	20.6%	17.4%	15.8%	26.3%	24.9%	21.5%	18.2%
Aged 30-39	29.4%	27.4%	24.9%	23.0%	26.9%	26.3%	23.1%	21.0%
Aged 40-49	26.1%	28.3%	27.3%	26.2%	25.8%	28.3%	26.7%	27.5%
Aged 50-59	19.9%	18.6%	22.8%	27.4%	19.1%	18.7%	23.9%	28.3%
Aged over 60	4.5%	5.1%	7.7%	7.6%	1.8%	1.8%	4.8%	5.0%
Qualification:								
Degree	21.3%	20.6%	33.3%	30.6%	23.6%	23.8%	40.5%	41.4%
Higher Education	8.9%	13.5%	9.3%	15.6%	12.1%	19.1%	10.8%	17.9%
A Level or equiv.	29.8%	37.3%	25.0%	28.1%	18.1%	19.6%	19.7%	16.7%
GVSE or quiv.	17.9%	12.3%	18.6%	13.9%	26.2%	19.0%	19.8%	14.9%
Other quals	13.3%	8.8%	9.1%	6.7%	12.2%	8.8%	6.0%	4.4%
No quals	8.7%	7.5%	4.7%	5.1%	7.7%	9.6%	3.2%	4.7%
Other personal characteristics:								
Non-UK born	7.7%	3.3%	14.5%	7.8%	9.3%	3.5%	16.0%	8.1%
Dependent child	43.2%	42.4%	43.5%	41.5%	32.7%	32.6%	35.4%	32.5%
Single	26.5%	26.3%	25.0%	25.2%	36.2%	38.0%	35.1%	37.7%
Occupation:								
Mangerial/ professional	34.4%	30.2%	36.2%	32.2%	29.1%	28.2%	36.6%	37.1%
Intermediate occ	16.2%	17.1%	18.2%	17.6%	22.5%	22.6%	17.9%	17.3%
Routine Occ	38.7%	40.4%	34.5%	36.6%	30.9%	31.9%	24.8%	25.1%
Service Occ	5.4%	6.3%	7.3%	8.1%	17.4%	17.1%	20.6%	20.3%
Manual Occ	5.2%	6.0%	3.9%	5.6%	0.1%	0.1%	0.1%	0.2%
Industry:								
A:Agriculture & forestry	1.0%	1.5%	1.1%	1.6%	0.3%	0.6%	0.4%	0.6%
B:Fishing	0.0%	0.3%	0.0%	0.3%	0.0%	0.1%	0.0%	0.1%
C:Mining, quarrying	0.5%	3.2%	0.7%	4.4%	0.1%	0.6%	0.1%	1.0%
D:Manufacturing	26.0%	21.6%	19.1%	14.5%	11.7%	9.9%	7.9%	7.5%
E:Electricity & utilities	1.4%	1.8%	1.5%	2.2%	0.6%	0.6%	0.6%	0.9%
F:Construction	9.1%	11.6%	8.1%	10.6%	1.5%	1.7%	1.9%	2.2%
G:Wholesale & retail	12.3%	10.1%	11.9%	10.0%	12.1%	10.2%	9.8%	9.3%
H:Hotels & restaurants	2.1%	2.5%	3.1%	3.1%	3.4%	4.8%	3.5%	4.1%
I:Transport & comms	10.4%	9.9%	9.5%	9.3%	4.5%	3.7%	4.0%	3.2%
J:Financial	4.5%	4.7%	4.6%	4.3%	6.3%	8.0%	4.7%	4.8%
K:Real estate & business	11.0%	8.3%	15.7%	13.5%	10.8%	8.7%	11.8%	9.3%
L: Public Admin	8.5%	9.9%	7.7%	9.4%	10.0%	11.9%	8.8%	9.0%
M:Educ	5.0%	5.4%	6.3%	6.2%	14.2%	14.5%	18.0%	15.4%
N:Health & social work	4.3%	5.0%	6.2%	5.9%	19.7%	20.6%	24.0%	28.2%
Workplace size:								
Small workplace (1-10)	15.0%	15.6%	16.0%	15.3%	15.0%	16.1%	13.3%	15.7%
Medium workplace (11-249)	52.5%	54.6%	50.0%	55.2%	53.0%	54.8%	54.1%	55.3%
Large workplace (250+)	32.4%	29.8%	34.0%	29.6%	32.1%	29.1%	32.6%	29.1%

v. Results

RIF unconditional quantile regression estimates

The results of the RIF unconditional quantile regressions at the median are shown in Table 5.3. The results show for example that an increase in the proportion of workers aged 16-29 would push down wages at the median (negative coefficient), whereas an increase in the proportion of workers with a degree would push up wages at the median (positive coefficient).

The interpretation of the coefficients is as follows. A ten per cent increase in the proportion of the workforce with a degree would tend to raise the median wage of Scottish males by 3% (0.1 multiplied by the coefficient 0.304). A higher coefficient for Scotland relative to rUK might reflect either higher returns to that characteristic in Scotland, or a higher proportion of the workforce with that characteristic.

The coefficients are largely of the expected sign. Increases in the proportion of low-wage occupations (intermediate, manual, service, routine) relative to professional and managerial occupations will tend to push down wages; increases in the proportion of employees in higher-wage sectors such as financial services will tend raises median wages.

Note how the coefficient on being born overseas was either positive or insignificant in 2002/3 but had become negative by 2014/15. This reflects in part the dramatic changes in the composition of immigrants and the jobs that they do following the accession of the eastern European countries to the EU in 2004. (In the context of recent debates on the effects of immigrants on the labour market, it is important to be clear about what this means. Given that overseas born workers tend to earn less than others, an increase in the proportion of them in the workforce will tend to reduce the median wage, purely as a result of the composition effect. The analysis says nothing about the effect of an increase in the proportion of immigrant workers on the wages of non-migrants.)

The results of the RIF unconditional quantile regressions at the ninetieth percentile are shown in Table 5.4. Some of the regression coefficients are quite different for the unconditional quantile regression of wages at the median. As might be expected, having a degree raises the 90th percentile of wages more than it raises wages at the median. As such, the coefficients tell us that a ten percent increase in the proportion of the rUK male workforce with a degree in 2014/15 would increase the 90/50 ratio of wage inequality by 1.9 percentage points[14].

---

[14] This result is obtained by multiplying the change in the increase in the proportion with a degree (0.1) by the difference between the coefficient at the 90th percentile and the coefficient at the median, i.e. 0.1*(0.466 − 0.278).

Table 5.3: Unconditional quantile regressions at median

	Males				Females			
	2002/3		2014/15		2002/3		2014/15	
	rUK	Scot	rUK	Scot	rUK	Scot	rUK	Scot
Aged 16-29	-0.314***	-0.366***	-0.373***	-0.304***	-0.206***	-0.255***	-0.302***	-0.320***
	(0.0082)	(0.0280)	(0.0102)	(0.0347)	(0.0089)	(0.0265)	(0.0108)	(0.0332)
Aged 30-39	-0.054***	-0.073***	-0.097***	-0.077***	0.011	-0.018	-0.053***	-0.127***
	(0.0069)	(0.0234)	(0.0087)	(0.0292)	(0.0083)	(0.0245)	(0.0100)	(0.0302)
Aged 50-59	-0.056***	-0.055**	0.001	0.024	-0.031***	0.004	-0.032***	-0.046
	(0.0079)	(0.0268)	(0.0092)	(0.0285)	(0.0096)	(0.0276)	(0.0104)	(0.0282)
Aged over 60	-0.171***	-0.222***	-0.056***	-0.037	-0.044*	-0.146**	-0.063***	-0.023
	(0.0134)	(0.0434)	(0.0139)	(0.0456)	(0.0253)	(0.0598)	(0.0187)	(0.0557)
Degree	0.245***	0.304***	0.278***	0.295***	0.327***	0.349***	0.313***	0.308***
	(0.0089)	(0.0349)	(0.0105)	(0.0375)	(0.0103)	(0.0333)	(0.0121)	(0.0392)
Higher Education	0.149***	0.219***	0.198***	0.176***	0.288***	0.231***	0.192***	0.160***
	(0.0110)	(0.0357)	(0.0133)	(0.0401)	(0.0118)	(0.0330)	(0.0148)	(0.0404)
A Level or equiv.	0.072***	0.124***	0.114***	0.122***	0.045***	0.094***	0.026**	0.054
	(0.0079)	(0.0292)	(0.0100)	(0.0346)	(0.0101)	(0.0321)	(0.0118)	(0.0394)
Other quals	-0.053***	0.031	-0.067***	0.018	-0.056***	-0.094**	-0.074***	-0.109**
	(0.0095)	(0.0384)	(0.0129)	(0.0493)	(0.0113)	(0.0374)	(0.0167)	(0.0518)
No quals	-0.137***	-0.135***	-0.135***	-0.115**	-0.174***	-0.136***	-0.171***	-0.122**
	(0.0103)	(0.0401)	(0.0150)	(0.0484)	(0.0124)	(0.0365)	(0.0193)	(0.0504)
Non-UK born	0.050***	-0.011	-0.083***	-0.097***	0.091***	0.093**	-0.028***	-0.089**
	(0.0095)	(0.0452)	(0.0089)	(0.0340)	(0.0109)	(0.0439)	(0.0098)	(0.0373)
Dependent child	0.035***	0.036*	0.042***	0.062***	-0.073***	-0.069***	-0.038***	0.008
	(0.0058)	(0.0201)	(0.0072)	(0.0236)	(0.0068)	(0.0203)	(0.0078)	(0.0231)
Single	-0.107***	-0.125***	-0.116***	-0.122***	-0.020***	-0.022	-0.029***	-0.051**
	(0.0067)	(0.0233)	(0.0083)	(0.0266)	(0.0066)	(0.0188)	(0.0076)	(0.0221)
Intermediate occ	-0.193***	-0.254***	-0.190***	-0.213***	-0.188***	-0.171***	-0.245***	-0.286***
	(0.0086)	(0.0298)	(0.0099)	(0.0341)	(0.0094)	(0.0283)	(0.0120)	(0.0361)
Routine Occ	-0.531***	-0.539***	-0.518***	-0.528***	-0.523***	-0.502***	-0.588***	-0.591***
	(0.0075)	(0.0264)	(0.0096)	(0.0309)	(0.0100)	(0.0296)	(0.0121)	(0.0354)
Service Occ	-0.652***	-0.664***	-0.648***	-0.655***	-0.650***	-0.571***	-0.715***	-0.685***
	(0.0108)	(0.0377)	(0.0123)	(0.0403)	(0.0106)	(0.0327)	(0.0114)	(0.0341)
Manual Occ	-0.573***	-0.634***	-0.531***	-0.599***	-0.697***	-0.741***	-0.660***	-0.123
	(0.0138)	(0.0463)	(0.0193)	(0.0551)	(0.0757)	(0.1432)	(0.1024)	(0.3173)
A:Agric & forestry	-0.033	-0.106*	-0.034	0.007	-0.059	0.056	0.013	-0.081
	(0.0250)	(0.0559)	(0.0290)	(0.0820)	(0.0535)	(0.1441)	(0.0628)	(0.1551)
B:Fishing	-0.087	0.396**	0.254	0.343	-0.680***	-0.053	-0.885***	-0.515***
	(0.2784)	(0.1715)	(0.2291)	(0.2438)	(0.0141)	(0.8368)	(0.0879)	(0.0986)
C:Mining, quarrying	0.270***	0.475***	0.330***	0.487***	0.252*	0.281***	0.132	0.324***
	(0.0340)	(0.0491)	(0.0348)	(0.0511)	(0.1300)	(0.1064)	(0.1168)	(0.0972)
D:Manufacturing	0.047***	0.100***	0.064***	0.092**	0.025*	0.096**	0.036**	0.068
	(0.0094)	(0.0324)	(0.0123)	(0.0414)	(0.0132)	(0.0405)	(0.0174)	(0.0507)
E:Elec & utilities	0.140***	0.207***	0.287***	0.287***	0.075*	0.327***	0.146***	0.230**
	(0.0240)	(0.0728)	(0.0285)	(0.0783)	(0.0456)	(0.1150)	(0.0534)	(0.0983)
F:Construction	0.164***	0.192***	0.158***	0.154***	0.034	0.033	0.099***	0.215***
	(0.0122)	(0.0390)	(0.0160)	(0.0468)	(0.0286)	(0.0789)	(0.0315)	(0.0811)
G:Wholes. & retail	-0.053***	-0.063	-0.063***	-0.109**	-0.081***	-0.154***	-0.108***	-0.124***
	(0.0108)	(0.0389)	(0.0134)	(0.0442)	(0.0127)	(0.0378)	(0.0158)	(0.0436)
H:Hotels & rest.	-0.169***	-0.125**	-0.155***	-0.208***	-0.192***	-0.172***	-0.183***	-0.211***
	(0.0174)	(0.0511)	(0.0184)	(0.0575)	(0.0175)	(0.0437)	(0.0202)	(0.0610)
I:Trans. & comms	0.067***	0.021	0.095***	0.093**	0.103***	0.065	0.094***	0.168**
	(0.0115)	(0.0391)	(0.0146)	(0.0467)	(0.0183)	(0.0605)	(0.0223)	(0.0778)
J:Financial	0.122***	0.002	0.147***	0.073	0.168***	0.118***	0.134***	0.139**
	(0.0131)	(0.0440)	(0.0162)	(0.0525)	(0.0159)	(0.0433)	(0.0199)	(0.0583)
K:Real est. & bus.	0.066***	0.072*	0.066***	0.045	0.116***	0.122***	0.090***	0.009
	(0.0106)	(0.0388)	(0.0124)	(0.0404)	(0.0132)	(0.0397)	(0.0154)	(0.0480)
M:Educ	-0.054***	0.038	-0.103***	0.028	-0.029**	0.033	-0.100***	0.029
	(0.0128)	(0.0424)	(0.0157)	(0.0447)	(0.0117)	(0.0333)	(0.0138)	(0.0394)
N:Health & soc.work	-0.130***	-0.084*	-0.109***	-0.165***	0.003	0.044	-0.071***	0.017

	Males				Females			
	2002/3		2014/15		2002/3		2014/15	
	rUK	Scot	rUK	Scot	rUK	Scot	rUK	Scot
	(0.0134)	(0.0427)	(0.0152)	(0.0514)	(0.0110)	(0.0319)	(0.0130)	(0.0348)
Small workplace	-0.098***	-0.104***	-0.140***	-0.070**	-0.105***	-0.100***	-0.120***	-0.074**
	(0.0076)	(0.0263)	(0.0094)	(0.0321)	(0.0094)	(0.0254)	(0.0117)	(0.0303)
Large workplace	0.074***	0.026	0.098***	0.046**	0.063***	0.077***	0.108***	0.079***
	(0.0057)	(0.0201)	(0.0070)	(0.0232)	(0.0071)	(0.0217)	(0.0082)	(0.0244)
Constant	6.600***	6.559***	6.523***	6.469***	6.299***	6.186***	6.404***	6.351***
	(0.0125)	(0.0444)	(0.0156)	(0.0516)	(0.0142)	(0.0438)	(0.0178)	(0.0539)
Observations	**44,285**	**4,406**	**33,742**	**3,048**	**28,422**	**3,085**	**23,859**	**2,298**
R-squared	**0.357**	**0.355**	**0.354**	**0.365**	**0.393**	**0.406**	**0.409**	**0.465**

Table 5.4: Unconditional quantile regressions at 90th percentile

	Males				Females			
	2002/3		**2014/15**		**2002/3**		**2014/15**	
	rUK	**Scot**	**rUK**	**Scot**	**rUK**	**Scot**	**rUK**	**Scot**
Aged 16-29	-0.274***	-0.286***	-0.354***	-0.321***	-0.263***	-0.231***	-0.346***	-0.324***
	(0.0116)	(0.0324)	(0.0156)	(0.0492)	(0.0118)	(0.0305)	(0.0139)	(0.0389)
Aged 30-39	-0.121***	-0.143***	-0.258***	-0.258***	-0.014	-0.086***	-0.167***	-0.186***
	(0.0124)	(0.0355)	(0.0177)	(0.0549)	(0.0134)	(0.0318)	(0.0170)	(0.0487)
Aged 50-59	0.019	0.019	0.064***	-0.019	0.007	-0.045	-0.015	0.019
	(0.0140)	(0.0417)	(0.0199)	(0.0562)	(0.0141)	(0.0353)	(0.0166)	(0.0485)
Aged over 60	-0.031	-0.015	0.052**	-0.027	-0.064**	-0.097	0.028	-0.059
	(0.0197)	(0.0541)	(0.0263)	(0.0750)	(0.0273)	(0.0602)	(0.0265)	(0.0695)
Degree	0.443***	0.427***	0.466***	0.304***	0.428***	0.399***	0.340***	0.334***
	(0.0177)	(0.0525)	(0.0198)	(0.0638)	(0.0165)	(0.0405)	(0.0159)	(0.0451)
Higher Education	0.018	0.050	0.046**	0.082	0.115***	0.083***	0.031*	0.028
	(0.0177)	(0.0425)	(0.0221)	(0.0594)	(0.0152)	(0.0283)	(0.0170)	(0.0356)
A Level or equiv.	-0.000	-0.017	0.058***	-0.012	0.070***	0.048**	0.072***	0.052
	(0.0101)	(0.0283)	(0.0136)	(0.0428)	(0.0100)	(0.0223)	(0.0114)	(0.0345)
Other quals	-0.018*	0.035	-0.028*	-0.021	0.003	0.020	-0.026*	-0.011
	(0.0108)	(0.0380)	(0.0148)	(0.0644)	(0.0106)	(0.0267)	(0.0155)	(0.0367)
No quals	-0.055***	-0.108***	-0.052***	-0.071	-0.062***	-0.034*	-0.074***	-0.017
	(0.0109)	(0.0277)	(0.0160)	(0.0582)	(0.0095)	(0.0193)	(0.0162)	(0.0374)
Non-UK born	0.092***	0.054	0.013	0.074	0.101***	0.151*	0.019	0.051
	(0.0180)	(0.0698)	(0.0187)	(0.0803)	(0.0172)	(0.0824)	(0.0161)	(0.0603)
Dependent child	0.111***	0.110***	0.129***	0.124***	-0.011	-0.020	0.046***	0.068*
	(0.0097)	(0.0283)	(0.0141)	(0.0411)	(0.0101)	(0.0247)	(0.0122)	(0.0383)
Single	-0.025***	-0.041	-0.058***	-0.051	-0.014	-0.010	-0.031***	-0.038
	(0.0096)	(0.0277)	(0.0133)	(0.0389)	(0.0090)	(0.0223)	(0.0105)	(0.0292)
Intermediate occ	-0.405***	-0.357***	-0.377***	-0.255***	-0.467***	-0.416***	-0.323***	-0.326***
	(0.0155)	(0.0474)	(0.0215)	(0.0660)	(0.0165)	(0.0409)	(0.0191)	(0.0548)
Routine Occ	-0.506***	-0.485***	-0.562***	-0.388***	-0.547***	-0.451***	-0.494***	-0.417***
	(0.0124)	(0.0386)	(0.0171)	(0.0558)	(0.0145)	(0.0357)	(0.0157)	(0.0475)
Service Occ	-0.495***	-0.441***	-0.539***	-0.405***	-0.452***	-0.405***	-0.341***	-0.289***
	(0.0138)	(0.0410)	(0.0186)	(0.0547)	(0.0138)	(0.0348)	(0.0141)	(0.0394)
Manual Occ	-0.433***	-0.456***	-0.471***	-0.292***	-0.490***	-0.432***	-0.498***	-0.097
	(0.0164)	(0.0471)	(0.0220)	(0.0715)	(0.0354)	(0.1305)	(0.0352)	(0.4717)
A:Agric & forestry	0.063**	0.091*	0.122***	0.064	-0.046	0.045	0.009	0.021
	(0.0271)	(0.0514)	(0.0355)	(0.0679)	(0.0401)	(0.1218)	(0.0407)	(0.1841)
B:Fishing	-0.082	-0.199**	-0.016	0.199	-0.060***	-0.392	-0.248**	-0.011
	(0.2129)	(0.0977)	(0.1885)	(0.4599)	(0.0171)	(0.3160)	(0.1160)	(0.0731)
C:Mining, quarrying	0.446***	1.000***	1.134***	1.395***	-0.022	0.531**	0.393**	0.594***
	(0.0772)	(0.1144)	(0.1194)	(0.1543)	(0.1587)	(0.2122)	(0.1932)	(0.2107)
D:Manufacturing	0.073***	0.163***	0.162***	0.153**	0.065***	0.078*	0.214***	0.092
	(0.0150)	(0.0434)	(0.0223)	(0.0626)	(0.0166)	(0.0439)	(0.0257)	(0.0712)
E:Elec & utilities	0.105**	0.122	0.266***	0.139	0.070	0.248	0.101	0.157
	(0.0411)	(0.1006)	(0.0572)	(0.1157)	(0.0602)	(0.1762)	(0.0763)	(0.1882)
F:Construction	0.125***	0.189***	0.166***	0.113*	-0.014	0.046	0.062	0.276**
	(0.0191)	(0.0508)	(0.0275)	(0.0647)	(0.0296)	(0.0728)	(0.0387)	(0.1190)
G:Wholes. & retail	0.084***	0.040	0.175***	0.135**	0.003	-0.061*	0.036*	-0.089*
	(0.0161)	(0.0414)	(0.0228)	(0.0556)	(0.0153)	(0.0340)	(0.0201)	(0.0521)
H:Hotels & rest.	0.026	0.159***	0.152***	0.123	-0.008	-0.022	0.101***	-0.054
	(0.0225)	(0.0581)	(0.0268)	(0.0779)	(0.0190)	(0.0376)	(0.0241)	(0.0490)
I:Trans. & comms	0.100***	0.140***	0.217***	0.302***	0.081***	0.060	0.174***	0.185*
	(0.0174)	(0.0499)	(0.0265)	(0.0778)	(0.0224)	(0.0504)	(0.0316)	(0.0952)
J:Financial	0.543***	0.291***	0.816***	0.469***	0.190***	0.033	0.443***	0.380***
	(0.0316)	(0.0814)	(0.0453)	(0.1234)	(0.0230)	(0.0464)	(0.0354)	(0.0990)
K:Real est. & bus.	0.280***	0.181***	0.327***	0.360***	0.201***	0.111**	0.271***	0.095
	(0.0212)	(0.0623)	(0.0261)	(0.0757)	(0.0200)	(0.0474)	(0.0242)	(0.0705)
M:Educ	-0.389***	-0.328***	-0.395***	-0.107	-0.102***	0.002	-0.159***	-0.110*
	(0.0243)	(0.0691)	(0.0283)	(0.0949)	(0.0199)	(0.0499)	(0.0213)	(0.0660)
N:Health & soc.work	0.068***	0.062	0.039	0.065	-0.078***	-0.015	-0.146***	-0.208***

	Males				Females			
	2002/3		**2014/15**		**2002/3**		**2014/15**	
	rUK	Scot	rUK	Scot	rUK	Scot	rUK	Scot
	(0.0244)	(0.0630)	(0.0305)	(0.0810)	(0.0143)	(0.0348)	(0.0184)	(0.0520)
Small workplace	-0.078***	-0.031	-0.085***	-0.050	-0.061***	-0.043*	-0.050***	0.039
	(0.0111)	(0.0311)	(0.0160)	(0.0441)	(0.0115)	(0.0254)	(0.0140)	(0.0383)
Large workplace	0.108***	0.139***	0.165***	0.159***	0.096***	0.067**	0.141***	0.128***
	(0.0100)	(0.0299)	(0.0143)	(0.0458)	(0.0103)	(0.0269)	(0.0131)	(0.0379)
Constant	7.138***	7.062***	7.084***	6.977***	6.952***	6.851***	6.922***	6.848***
	(0.0205)	(0.0618)	(0.0287)	(0.0815)	(0.0206)	(0.0493)	(0.0252)	(0.0783)
Observations	**44,285**	**4,406**	**33,742**	**3,048**	**28,422**	**3,085**	**23,859**	**2,298**
R-squared	**0.187**	**0.217**	**0.176**	**0.177**	**0.203**	**0.240**	**0.166**	**0.189**

Decomposing the wage gaps

Estimates from the RIF regressions are used to calculate the wage structure and composition effects for males and females. A summary of the results is presented in Table 5.5.

The upper panel compares the wages of Scottish workers with those of rUK. Scottish median wages for males were 3.5% lower than in rUK in 2002/3, and 1.9% lower in 2014/15. Some of this wage convergence can be explained by compositional effects (the coefficient on the explained component is negative in 2002/3, implying that the composition of the Scottish workforce contributed to its negative wage gap, whereas by 2014/15 the coefficient was positive, implying compositional factors were now in Scotland's favour). However, the coefficient on compositional effects is not statistically significant in either period. Unexplained wage structure effects account for around 3% of the difference in wages in both periods.

For females, wage convergence has been more marked, with Scottish wages being 6.9% lower in 2002/3 but just 1.3% lower by 2014/15. Some of this convergence can be explained by compositional effects, which were not significant in explaining wage differences in 2002/3, but which by 2014/15 would have led, on their own, to Scottish median wages being 1.9% higher than rUK. But the convergence is largely due to a fall in the unexplained wage structure effects (i.e. a convergence in the returns to characteristics).

At the 90th percentile the gap between Scottish and rUK wages has narrowed for males and females, but only slightly. For both males and females however, compositional effects have worked in Scotland's favour, and should have reduced the wage gap. But the wage structure effect has increased for males, and fallen only slightly for females.

One of the issues that arises when decomposing the difference in wages between Scotland and rUK into an 'explained' and unexplained' component is that there is a substantial wage premium associated with London that has nothing to do with workforce composition. The possible explanations for this London premium were discussed in the previous section.

In order to shed light on the significance of compositional factors, in the lower panel of Table 5.5 we present results of the same analysis but excluding those who work in London from the analysis (so rUK is the UK minus Scotland, and London).

Relative to rUK excluding London (rUKexLon), Scottish median wages for males have increased by around 1.4 percentage points (from 1 to 1.24% higher than rUKexLon wages) between 2002/3 and 2014/15, explained entirely by compositional factors (effectively none of the wage gap is

unexplained). The story is similar at the 90th percentile, with a rise in relative Scottish male wages driven largely by compositional factors.

Median wages for females were 1.5% lower in Scotland than rUKexLon in 2002/3, but 4.2% higher by 2014/15. Compositional factors have moved in Scotland's favour – they explained 2.7 percentage points of the gap in 2002/3 and 4.2 percentage points of the gap in 2014/15. But there has also been a significant fall in the wage structure effect, which reduced Scottish median wages by 4% relative to rUKexLon wages in 2002/3, but which had become insignificant by 2014/15. The story at the 90th percentile is similar, with the convergence in Scottish wages being driven in part through compositional factors, but in particular by a fall in the unexplained wage premium that female workers in rUKexLon receive over their Scottish counterparts.

Table 5.5: Summary of decomposition results

	2002/3				2014/15			
	Difference	**Explained**	**Unexplained**	**Observations**	**Difference**	**Explained**	**Unexplained**	**Observations**
Scottish wages relative to rUK								
Males - median	-0.0348	-0.0044	-0.0304***	48,691	-0.0185	0.0097	-0.0282***	36,790
		(0.006)	(0.009)			(0.008)	(0.011)	
Females - median	-0.0691	0.0093	-0.0784***	31,507	-0.0134	0.0187*	-0.0321***	26,157
		(0.008)	(0.010)			(0.010)	(0.011)	
Males - 90th percentile	-0.0531	-0.0252***	-0.0279**	48,691	-0.0652	0.0017	-0.0669***	36,790
		(0.007)	(0.013)			(0.011)	(0.020)	
Females - 90th percentile	-0.0907	-0.0063	-0.0844***	31,507	-0.0718	0.0016	-0.0734***	26,157
		(0.007)	(0.012)			(0.008)	(0.016)	
Scottish wages relative to rUK (excluding London)								
Males - median	0.0105	0.0138**	-0.0033	42,953	0.0242	0.0281***	-0.0039	32,016
		(0.006)	(0.009)			(0.008)	(0.011)	
Females - median	-0.0149	0.0268***	-0.0417***	27,610	0.0438	0.0417***	0.0021	22,600
		(0.007)	(0.010)			(0.010)	(0.011)	
Males - 90th percentile	0.0155	0.0004	0.0151	42,953	0.0407	0.0286***	0.0121	32,016
		(0.007)	(0.013)			(0.010)	(0.020)	
Females - 90th percentile	-0.0474	0.0168**	-0.0642***	27,610	0.0054	0.0233***	-0.0179	22,600
		(0.007)	(0.012)			(0.008)	(0.016)	

Although the composition effects in aggregate aren't significant in explaining wage differences between Scotland and rUK (i.e. including London), individual covariates can be significant in explaining wage differences. Table 5.6 presents the detailed results of a decomposition of median wages between Scotland and rUK in 2002/3 and 2014/15, for both males and females.

There is some evidence that demographic changes in the workforce are becoming more significant in explaining wage differences. The coefficient on the dummy variable for those aged 16-29 was insignificant in 2002/3 for males and females, but by 2014/15, the lower proportion of younger workers employed in Scotland would be expected to raise Scottish median wages by 0.5 percentage points for males and a full percentage point for females, relative to those of rUK.

A much larger proportion of the Scottish workforce has 'higher education', and this is capable of explaining a one percentage point wedge in wages (in Scotland's favour) in 2014/15 (the scale of effect has risen slightly and significantly for men but fallen somewhat for women).

In discussing the results from the unconditional quantile regressions, it was remarked that the coefficient on non-UK born workers had changed sign between 2002/3 and 2014/15. The effects of this feed through to the decompositions. The fact that Scotland has a lower proportion of foreign born workers pushed down on wages in 2002/3, but now has a slight upward effect on wages at the median.

Differences in occupational structure have some effect on the wage gap for males. That Scotland has a higher proportion of routine and manual occupations pushes down on wages slightly. The effect has grown very slightly (but not statistically significantly) since 2002/3.

In terms of industrial composition, Scotland's slightly lower proportion of jobs in real estate and business pushes down wages slightly for males and females, although the scale of this effect has not changed statistically significantly between the two periods. In terms of industrial composition, Scotland's higher proportion of employment in mining and quarrying (which includes the offshore sector) has a significant wage effect for males. The higher concentration of employment in this sector pushed up Scottish wages by 0.7 percentage points in 2002/3, rising to 1.2 percentage points in 2014/15, with this difference between coefficients in the two periods being statistically significant.

Finally, Scotland's somewhat lower density of jobs in large workplaces pushes down median wages somewhat for males and females, and this negative influence on Scottish relative wages has tended to increase slightly (but not statistically significantly) between the two periods.

The decomposition results at the 90th percentile are shown in Table 5.7. Recall from Table 5.5 that the gap between Scottish and rUK wages actually increased between 2002/3 and 2014/15, but this

marked a fall in the overall role of the composition effect, and an increase in the role of the unexplained wage structure effect in explaining the gap.

Looking at the coefficients on the explained wage structure effects in Table 5.7, it can be seen that a number of compositional effects did indeed move in Scotland's favour between the two periods. The coefficients on a number of age variables became positive and significant, reflecting the effect of Scotland's relatively ageing population on wages. The coefficient on non-UK born workers, which was negative in 2002/3, was insignificant in 2014/15 (reflecting the fact that non-UK born workers, who are underrepresented in Scotland, tended to raise wages at the 90th percentile in the earlier period, but had an insignificant effect in 2014/15). A further concentration of Scottish employment in the mining and quarrying sector also helps explain why overall composition effects moved in Scotland's favour.

Of course some compositional factors worked against Scotland. Most significantly, the coefficient on 'degree' is insignificant in 2002/3, but is significant and negative by 2014/15. Indeed, Table 5.2 shows that the proportion of the Scottish workforce educated to degree level was only slightly below that in rUK in 2002/3, but by 2014/15 the gap had grown somewhat. Note also how, for males, the unexplained coefficient on degree is significantly negative in 2014/15, implying that there are higher returns to a degree at the 90th percentile in rUK than Scotland.

For females, compositional effects in aggregate were not significant in explaining Scotland's wage gap at the 90th percentile. However, this hides some movement in the significance of individual coefficients. Demographic effects moved in Scotland's favour. The coefficient on the 'explained' dummy variable for females aged 16-29 is insignificant in 2002/3, but is significant and raises Scottish wages relative to rUK wages by one percentage point in 2014/15. Against this, some changes in industry structure worked to reduce the relative wages of Scottish females, including Scotland's slower growth in employment in business services, and relatively faster growth in health and social care, for example. Scotland also has a lower proportion of large workplaces, which tends to reduce its relative wages for both males and females, with the role of this coefficient becoming larger (but not statistically so) during the two periods.

130

Table 5.6: Decomposition of differences in median wages, detailed results

	Males				Females			
	2002/3		2014/15		2002/3		2014/15	
	Explained	Unexpl.	Explained	Unexpl.	Explained	Unexpl.	Explained	Unexpl.
Aged 16-29	-0.0011	-0.0107*	0.0058**	0.0110*	0.0030*	-0.0122*	0.0100***	-0.0033
	(0.002)	(0.006)	(0.003)	(0.006)	(0.002)	(0.007)	(0.003)	(0.006)
Aged 30-39	0.0010**	-0.0053	0.0019**	0.0045	-0.0001	-0.0077	0.0011**	-0.0157**
	(0.000)	(0.007)	(0.001)	(0.007)	(0.000)	(0.007)	(0.001)	(0.007)
Aged 50-59	0.0007**	0.0002	0.0000	0.0062	0.0001	0.0065	-0.0014**	-0.0040
	(0.000)	(0.005)	(0.000)	(0.008)	(0.000)	(0.005)	(0.001)	(0.009)
Aged over 60	-0.0011*	-0.0026	0.0000	0.0014	-0.0000	-0.0019	-0.0001	0.0020
	(0.001)	(0.002)	(0.000)	(0.004)	(0.000)	(0.001)	(0.000)	(0.003)
Degree	-0.0017	0.0120	-0.0076***	0.0052	0.0006	0.0055	0.0030	-0.0017
	(0.002)	(0.007)	(0.002)	(0.012)	(0.003)	(0.008)	(0.003)	(0.017)
Higher Education	0.0068***	0.0094*	0.0125***	-0.0035	0.0201***	-0.0109	0.0136***	-0.0056
	(0.001)	(0.005)	(0.002)	(0.007)	(0.002)	(0.007)	(0.002)	(0.008)
A Level or equiv.	0.0054***	0.0194*	0.0035***	0.0020	0.0007*	0.0096	-0.0008*	0.0047
	(0.001)	(0.011)	(0.001)	(0.010)	(0.000)	(0.007)	(0.000)	(0.007)
Other quals	0.0024***	0.0075**	0.0016***	0.0057*	0.0019***	-0.0033	0.0012***	-0.0015
	(0.000)	(0.004)	(0.000)	(0.003)	(0.001)	(0.003)	(0.000)	(0.002)
No quals	0.0016***	0.0002	-0.0006	0.0010	-0.0033***	0.0037	-0.0024***	0.0023
	(0.001)	(0.003)	(0.001)	(0.003)	(0.001)	(0.004)	(0.001)	(0.003)
Non-UK born	-0.0022***	-0.0020	0.0056***	-0.0011	-0.0053***	0.0000	0.0022***	-0.0049
	(0.000)	(0.002)	(0.001)	(0.003)	(0.001)	(0.002)	(0.001)	(0.003)
Dependent child	-0.0003	0.0003	-0.0008**	0.0082	0.0001	0.0015	0.0011**	0.0149*
	(0.000)	(0.009)	(0.000)	(0.010)	(0.001)	(0.007)	(0.000)	(0.008)
Single	0.0002	-0.0047	-0.0003	-0.0016	-0.0004	-0.0009	-0.0008**	-0.0083
	(0.001)	(0.006)	(0.001)	(0.007)	(0.000)	(0.008)	(0.000)	(0.009)
Intermediate occ	-0.0017	-0.0105**	0.0011	-0.0040	-0.0002	0.0039	0.0014	-0.0072
	(0.001)	(0.005)	(0.001)	(0.006)	(0.001)	(0.007)	(0.002)	(0.007)
Routine Occ	-0.0087**	-0.0030	-0.0110**	-0.0036	-0.0055	0.0066	-0.0016	-0.0006
	(0.004)	(0.011)	(0.005)	(0.012)	(0.005)	(0.010)	(0.006)	(0.009)
Service Occ	-0.0059**	-0.0008	-0.0052	-0.0005	0.0023	0.0136**	0.0019	0.0060
	(0.002)	(0.002)	(0.003)	(0.003)	(0.005)	(0.006)	(0.006)	(0.007)
Manual Occ	-0.0044**	-0.0037	-0.0090***	-0.0038	-0.0002	-0.0001	-0.0003	0.0009
	(0.002)	(0.003)	(0.002)	(0.003)	(0.000)	(0.000)	(0.001)	(0.001)
A:Agric & forestry	-0.0002	-0.0011	-0.0001	0.0006	-0.0001	0.0006	0.0000	-0.0005
	(0.000)	(0.001)	(0.000)	(0.001)	(0.000)	(0.001)	(0.000)	(0.001)
B:Fishing	-0.0002	0.0013	0.0006	0.0002	-0.0004	0.0004	-0.0007	0.0003
	(0.001)	(0.001)	(0.001)	(0.001)	(0.000)	(0.001)	(0.001)	(0.000)
C:Mining, quarrying	0.0073***	0.0066***	0.0123***	0.0069**	0.0012*	0.0002	0.0012	0.0020
	(0.001)	(0.002)	(0.002)	(0.003)	(0.001)	(0.001)	(0.001)	(0.002)
D:Manufacturing	-0.0021***	0.0114	-0.0029***	0.0041	-0.0005	0.0070*	-0.0002	0.0024
	(0.001)	(0.007)	(0.001)	(0.006)	(0.000)	(0.004)	(0.000)	(0.004)
E:Elec & utilities	0.0006**	0.0012	0.0019**	-0.0000	0.0001	0.0016*	0.0005	0.0007
	(0.000)	(0.001)	(0.001)	(0.002)	(0.000)	(0.001)	(0.000)	(0.001)

131

	Males				Females			
	2002/3		2014/15		2002/3		2014/15	
	Explained	Unexpl.	Explained	Unexpl.	Explained	Unexpl.	Explained	Unexpl.
F:Construction	0.0041***	0.0033	0.0040***	-0.0004	0.0001	-0.0000	0.0003	0.0026
	(0.001)	(0.005)	(0.001)	(0.005)	(0.000)	(0.001)	(0.000)	(0.002)
G:Wholes. & retail	0.0012***	-0.0011	0.0012***	-0.0046	0.0015***	-0.0074*	0.0005	-0.0015
	(0.000)	(0.004)	(0.000)	(0.005)	(0.001)	(0.004)	(0.001)	(0.004)
H:Hotels & rest.	-0.0007	0.0011	0.0001	-0.0016	-0.0028***	0.0010	-0.0011	-0.0011
	(0.000)	(0.001)	(0.001)	(0.002)	(0.001)	(0.002)	(0.001)	(0.003)
I:Trans. & comms	-0.0004	-0.0046	-0.0002	-0.0002	-0.0009**	-0.0014	-0.0008*	0.0024
	(0.000)	(0.004)	(0.001)	(0.005)	(0.000)	(0.002)	(0.000)	(0.003)
J:Financial	0.0002	-0.0056**	-0.0004	-0.0032	0.0029***	-0.0040	0.0001	0.0002
	(0.000)	(0.002)	(0.001)	(0.002)	(0.001)	(0.004)	(0.001)	(0.003)
K:Real est. & bus.	-0.0018***	0.0005	-0.0015***	-0.0028	-0.0025***	0.0005	-0.0023***	-0.0075
	(0.000)	(0.003)	(0.001)	(0.006)	(0.001)	(0.004)	(0.001)	(0.005)
M:Educ	-0.0002	0.0049**	0.0002	0.0081***	-0.0001	0.0091*	0.0025***	0.0199***
	(0.000)	(0.002)	(0.000)	(0.003)	(0.000)	(0.005)	(0.001)	(0.007)
N:Health & soc.work	-0.0009**	0.0023	0.0003	-0.0033	0.0000	0.0084	-0.0029***	0.0248**
	(0.000)	(0.002)	(0.000)	(0.003)	(0.000)	(0.007)	(0.001)	(0.010)
Small workplace	-0.0005	-0.0010	0.0011	0.0108**	-0.0012*	0.0008	-0.0028***	0.0073
	(0.001)	(0.004)	(0.001)	(0.005)	(0.001)	(0.004)	(0.001)	(0.005)
Large workplace	-0.0019***	-0.0143**	-0.0044***	-0.0154**	-0.0019***	0.0040	-0.0038***	-0.0087
	(0.001)	(0.006)	(0.001)	(0.007)	(0.001)	(0.007)	(0.001)	(0.007)
Total	-0.0044	-0.0304***	0.0097	-0.0282***	0.0093	-0.0784***	0.0187*	-0.0321***
	(0.006)	(0.009)	(0.008)	(0.011)	(0.008)	(0.010)	(0.010)	(0.011)
Constant		-0.0413		-0.0546		-0.1131**		-0.0532
		(0.046)		(0.054)		(0.046)		(0.057)
Observations	48,691	48,691	36,790	36,790	31,507	31,507	26,157	26,157

Standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1

Table 5.7: Decomposition of differences at 90th percentile, detailed results

	Males				Females			
	2002/3		2014/15		2002/3		2014/15	
	Explained	Unexpl.	Explained	Unexpl.	Explained	Unexpl.	Explained	Unexpl.
Aged 16-29	-0.0009	-0.0024	0.0055**	0.0053	0.0038*	0.0078	0.0115***	0.0040
	(0.002)	(0.007)	(0.002)	(0.008)	(0.002)	(0.008)	(0.003)	(0.008)
Aged 30-39	0.0023***	-0.0060	0.0050**	0.0000	0.0001	-0.0187**	0.0034**	-0.0041
	(0.001)	(0.010)	(0.002)	(0.013)	(0.000)	(0.009)	(0.002)	(0.011)
Aged 50-59	-0.0002	0.0001	0.0029***	-0.0225	-0.0000	-0.0097	-0.0006	0.0096
	(0.000)	(0.008)	(0.001)	(0.016)	(0.000)	(0.007)	(0.001)	(0.014)
Aged over 60	-0.0002	0.0008	-0.0000	-0.0060	-0.0000	-0.0006	0.0000	-0.0043
	(0.000)	(0.003)	(0.000)	(0.006)	(0.000)	(0.001)	(0.000)	(0.004)
Degree	-0.0032	-0.0032	-0.0128***	-0.0498**	0.0008	-0.0069	0.0033	-0.0025
	(0.003)	(0.011)	(0.004)	(0.021)	(0.003)	(0.010)	(0.004)	(0.020)
Higher Education	0.0008	0.0043	0.0029**	0.0056	0.0080***	-0.0060	0.0022*	-0.0004
	(0.001)	(0.006)	(0.001)	(0.010)	(0.001)	(0.006)	(0.001)	(0.007)
A Level or equiv.	-0.0000	-0.0061	0.0018***	-0.0198	0.0011**	-0.0044	-0.0022***	-0.0034
	(0.001)	(0.011)	(0.001)	(0.013)	(0.001)	(0.005)	(0.001)	(0.006)
Other quals	0.0008*	0.0047	0.0007*	0.0004	-0.0001	0.0015	0.0004	0.0007
	(0.000)	(0.004)	(0.000)	(0.004)	(0.000)	(0.003)	(0.000)	(0.002)
No quals	0.0006**	-0.0040*	-0.0002	-0.0009	-0.0012***	0.0027	-0.0011**	0.0027
	(0.000)	(0.002)	(0.000)	(0.003)	(0.000)	(0.002)	(0.000)	(0.002)
Non-UK born	-0.0041***	-0.0013	-0.0009	0.0047	-0.0058***	0.0018	-0.0015	0.0025
	(0.001)	(0.002)	(0.001)	(0.006)	(0.001)	(0.003)	(0.001)	(0.005)
Dependent child	-0.0008	-0.0004	-0.0026**	-0.0022	0.0000	-0.0029	-0.0013**	0.0074
	(0.001)	(0.013)	(0.001)	(0.018)	(0.000)	(0.009)	(0.001)	(0.013)
Single	0.0000	-0.0043	-0.0001	0.0017	-0.0002	0.0017	-0.0008*	-0.0026
	(0.000)	(0.008)	(0.000)	(0.010)	(0.000)	(0.009)	(0.000)	(0.012)
Intermediate occ	-0.0036	0.0081	0.0021	0.0215*	-0.0006	0.0114	0.0018	-0.0005
	(0.002)	(0.009)	(0.003)	(0.012)	(0.004)	(0.010)	(0.003)	(0.010)
Routine Occ	-0.0083**	0.0085	-0.0120**	0.0640***	-0.0058	0.0305**	-0.0013	0.0191
	(0.004)	(0.016)	(0.005)	(0.021)	(0.005)	(0.012)	(0.005)	(0.013)
Service Occ	-0.0045**	0.0034	-0.0043	0.0108**	0.0016	0.0079	0.0009	0.0106
	(0.002)	(0.003)	(0.003)	(0.005)	(0.003)	(0.006)	(0.003)	(0.009)
Manual Occ	-0.0034**	-0.0014	-0.0080***	0.0100**	-0.0001	0.0001	-0.0002	0.0007
	(0.002)	(0.003)	(0.002)	(0.004)	(0.000)	(0.000)	(0.000)	(0.001)
A:Agric & forestry	0.0003*	0.0004	0.0005	-0.0009	-0.0001	0.0005	0.0000	0.0001
	(0.000)	(0.001)	(0.000)	(0.001)	(0.000)	(0.001)	(0.000)	(0.001)
B:Fishing	-0.0002	-0.0003	-0.0000	0.0006	-0.0000	-0.0002	-0.0002	0.0002
	(0.001)	(0.001)	(0.000)	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)
C:Mining, quarrying	0.0120***	0.0178***	0.0424***	0.0115	-0.0001	0.0030*	0.0035*	0.0021
	(0.002)	(0.005)	(0.006)	(0.009)	(0.001)	(0.002)	(0.002)	(0.003)
D:Manufacturing	-0.0032***	0.0196**	-0.0074***	-0.0014	-0.0012**	0.0012	-0.0010	-0.0092
	(0.001)	(0.010)	(0.001)	(0.010)	(0.000)	(0.005)	(0.001)	(0.006)
E:Elec & utilities	0.0005	0.0003	0.0018**	-0.0028	0.0001	0.0012	0.0003	0.0005
	(0.000)	(0.002)	(0.001)	(0.003)	(0.000)	(0.001)	(0.000)	(0.002)

	Males				Females			
	2002/3		2014/15		2002/3		2014/15	
	Explained	Unexpl.	Explained	Unexpl.	Explained	Unexpl.	Explained	Unexpl.
F:Construction	0.0031***	0.0074	0.0042***	-0.0056	-0.0000	0.0010	0.0002	0.0047*
	(0.001)	(0.006)	(0.001)	(0.007)	(0.000)	(0.001)	(0.000)	(0.003)
G:Wholes. & retail	-0.0019***	-0.0045	-0.0034***	-0.0039	-0.0001	-0.0065*	-0.0002	-0.0116**
	(0.001)	(0.004)	(0.001)	(0.006)	(0.000)	(0.004)	(0.000)	(0.005)
H:Hotels & rest.	0.0001	0.0034**	-0.0001	-0.0009	-0.0001	-0.0007	0.0006	-0.0064***
	(0.000)	(0.002)	(0.001)	(0.003)	(0.000)	(0.002)	(0.000)	(0.002)
I:Trans. & comms	-0.0005	0.0040	-0.0005	0.0078	-0.0007*	-0.0008	-0.0015**	0.0004
	(0.000)	(0.005)	(0.001)	(0.008)	(0.000)	(0.002)	(0.001)	(0.003)
J:Financial	0.0010	-0.0117***	-0.0024	-0.0150**	0.0033***	-0.0126***	0.0002	-0.0030
	(0.002)	(0.004)	(0.003)	(0.006)	(0.001)	(0.004)	(0.002)	(0.005)
K:Real est. & bus.	-0.0075***	-0.0082	-0.0073***	0.0044	-0.0042***	-0.0078*	-0.0068***	-0.0163**
	(0.001)	(0.005)	(0.002)	(0.011)	(0.001)	(0.004)	(0.002)	(0.007)
M:Educ	-0.0015	0.0032	0.0006	0.0178***	-0.0004	0.0151*	0.0040***	0.0076
	(0.001)	(0.004)	(0.002)	(0.006)	(0.001)	(0.008)	(0.001)	(0.011)
N:Health & soc.work	0.0005	-0.0003	-0.0001	0.0015	-0.0006	0.0130*	-0.0060***	-0.0177
	(0.000)	(0.003)	(0.000)	(0.005)	(0.001)	(0.008)	(0.002)	(0.016)
Small workplace	-0.0004	0.0073	0.0006	0.0054	-0.0007	0.0029	-0.0012**	0.0140**
	(0.000)	(0.005)	(0.001)	(0.007)	(0.000)	(0.005)	(0.001)	(0.006)
Large workplace	-0.0028***	0.0092	-0.0074***	-0.0018	-0.0029***	-0.0085	-0.0049***	-0.0038
	(0.001)	(0.009)	(0.002)	(0.014)	(0.001)	(0.008)	(0.001)	(0.012)
Total	-0.0252***	-0.0279**	0.0017	-0.0669***	-0.0063	-0.0844***	0.0016	-0.0734***
	(0.007)	(0.013)	(0.011)	(0.020)	(0.007)	(0.012)	(0.008)	(0.016)
Constant		-0.0763		-0.1063		-0.1013*		-0.0746
		(0.065)		(0.086)		(0.053)		(0.082)
Observations	48,691	48,691	36,790	36,790	31,507	31,507	26,157	26,157

Standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1

vi. Explaining the unexplained

The previous section showed that Scottish median wages for males and females did close the gap to those of rUK (including London) between 2002/3 and 2014/15, but this was largely due to a reduction in the unexplained wage structure effect. Some compositional factors did move in Scotland's favour and thus help to narrow the gap (notably demographic effects, changes in the composition of the immigrant workforce, and for males, employment in the offshore sector), but other compositional effects worked in the opposite direction. In this section we consider what factors might have caused the wage structure effect to have become less important in effecting the wage gap.

One potential explanation for higher wage growth in Scotland is that unemployment fell slightly more rapidly in Scotland than the UK as a whole during the early and mid 2000s. Economists have long accepted that there is some relationship between unemployment and wages, although the precise nature of theorised and empirical relationship has evolved over time (Gregg et al. 2014).

The literature on the 'wage curve' (Blanchflower and Oswald, 1995) posits a relationship between local or regional unemployment, and wages. This literature uses individual wage data as the explanatory variable in a regression of wages on individual characteristics and local unemployment. The wage curve 'rule of thumb', based on empirical studies across a large number of countries and time periods, is that the unemployment elasticity of pay is about -0.1, i.e. a doubling of the unemployment rate reduces wages by 10%.

More recently for the UK, Gregg et al. (2014) explore the relationship between unemployment and real wages using panel data for UK regions. They argue that wages have become more sensitive to unemployment in the period since 2003 than prior to this, although their empirical estimates of the unemployment elasticity of pay range from -0.05 to -0.12 i.e. a doubling of the unemployment rate reduces wages by between 5%-12%.

As shown in Figure 5.10, the UK unemployment rate varied very little between 2001-2008, hovering around the 5% that is sometimes associated with full employment (the corresponding employment rate is shown in Figure 5.11). At the start of this period, the Scottish unemployment rate was around 1.5 percentage points higher, at 6.5%, and it subsequently descended secularly to below 5% as the recession took hold.

Thus whilst it could be argued that UK unemployment remained largely unchanged during 2001-2008, it declined by around one third in Scotland. Taking as given an employment elasticity of pay of -0.1, this fall in Scottish unemployment might explain a 3% rise in Scottish real wages. Between

2002-2009 (allowing for a one year lag in the effect of unemployment on wages), the Scottish real median weekly wage increased by 9%, compared to an increase of 5% in the UK.

This simplistic analysis suggests that the relatively faster fall in unemployment in Scotland may explain some of Scotland's faster wage growth during the mid to late 2000s. But unless the lag between unemployment and wages is particularly long, it is not clear that this can provide a full explanation as to why Scottish wages continued to converge to the UK median over the period 2010-2015. Indeed, the long-term trend in unemployment rates does not really look so different that it could explain the secular convergence between Scottish and UK median wages since 2004. Since 2010, UK and Scottish unemployment rates have tracked each other closely.

Furthermore, whilst wages might be sensitive to unemployment at the lower end of the wage distribution, the unemployed are often not close substitutes for those already in employment, especially those in the upper part of the wage distribution. As we saw in the previous section, a declining 'wage structure' effect explains a greater part of the convergence in median wages than it does for wages at the 90th percentile, which is consistent with the notion that falling unemployment has a greater influence on wages in the bottom half relative to the top half.

We also saw in the previous section however that whilst wage structure effects are significant in explaining the gap between Scottish and rUK wages, they are not significant in explaining the gap between Scottish and rUKexLon wages. Thus perhaps the more pertinent question to ask is what causes London's unexplained wage premium be so high, and why might this have changed over time? This is outwith the scope of this paper, but was discussed to an extent in the preceding chapter.

Figure 5.10: Quarterly 16-64 unemployment rate, UK and Scotland
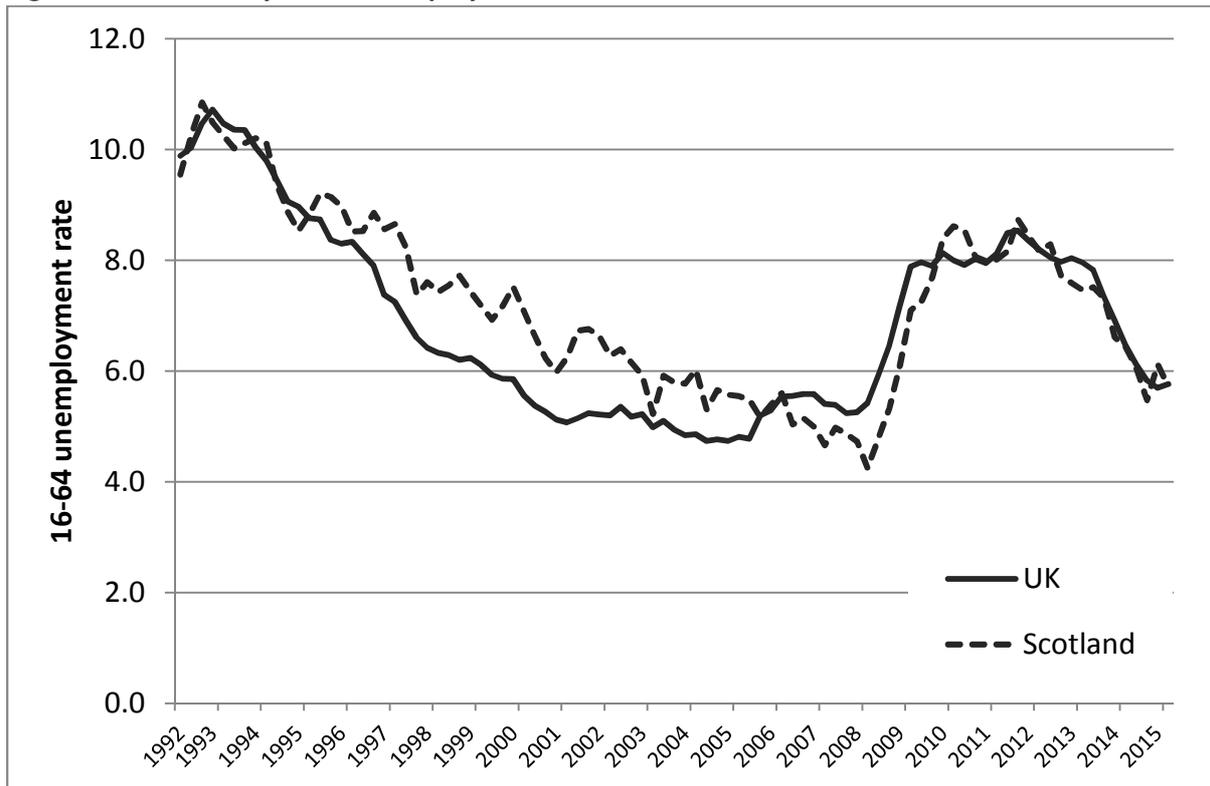
Source: ONS

Figure 5.11: Quarterly 16-64 employment rate, UK and Scotland
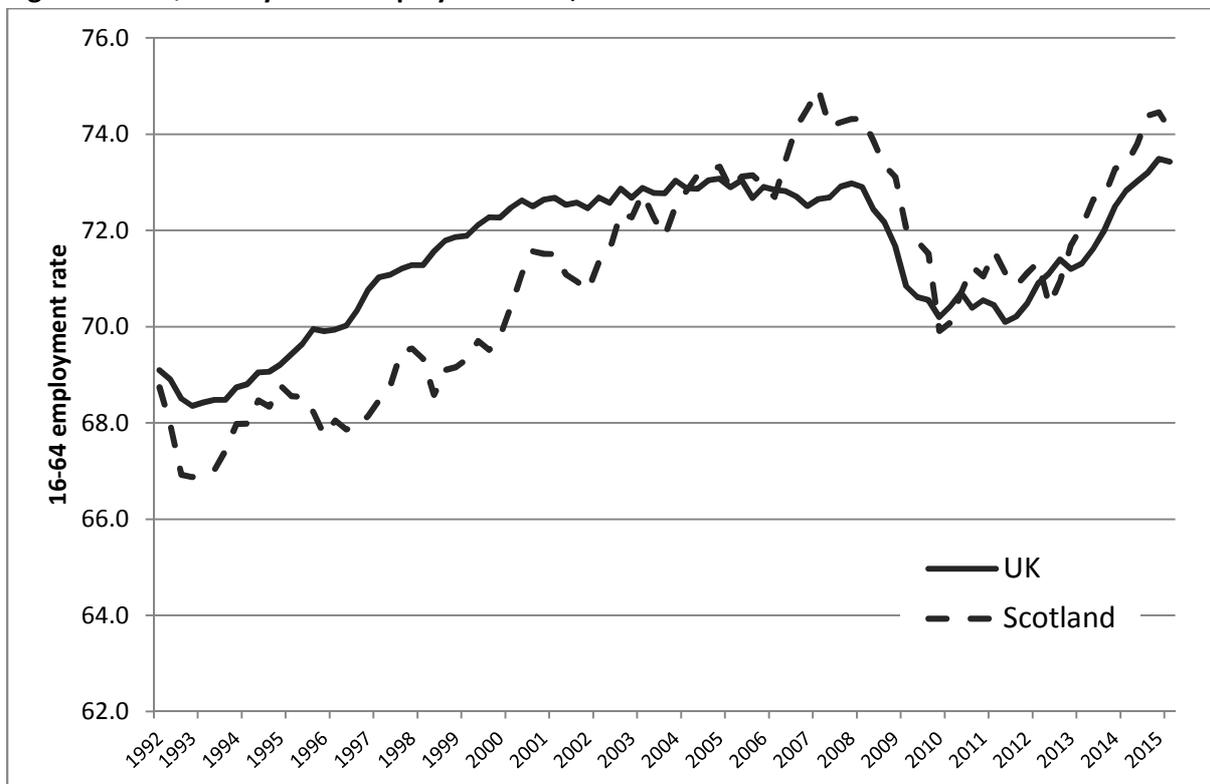
vii. Conclusions

Devolution of income tax to Scotland is motivated by the sense that it will improve the financial accountability of Scottish politicians for the policy decisions that they make. In future, the devolved Scottish budget will bear the risk in full that Scottish income tax revenues per capita grow more slowly than those in rUK; conversely, the Scottish budget will reap the rewards in full of any per capita revenue growth that is faster than that achieved in rUK.

Income tax revenues per capita are correlated with the employment to population ratio, and the wage rate; and, more specifically, on the distribution of wages. But the first part of this paper has demonstrated that changes to the income tax structure also have the potential to affect the growth rate of tax revenues, and more importantly the relative growth rate of tax revenues across UK regions. In the context of a devolved tax for which some elements (e.g. the Personal Allowance) continue to be set centrally, this is an important point. If the Personal Allowance were to continue to increase in real terms, then it is feasible, given the different distribution of income taxpayers in Scotland relative to rUK, that Scottish revenues per capita may grow less quickly than those in rUK, even if employment and wages grow at the same rates in Scotland and in rUK. Such an outcome might undermine one of the motivations for tax decentralisation.

The fact that the UK Government defines the income tax base may give rise to similar asymmetric budgetary effects under other circumstances. For example, changes to the taxation of pension income could potentially have asymmetric budgetary effects if pension income formed a different proportion of total income tax revenues in Scotland relative to rUK.

Notwithstanding these issues, what can a devolved government do to encourage wage growth and hence growth of tax revenues per capita? The second part of this paper considers this question somewhat indirectly, by assessing the extent to which wage convergence between Scotland and rUK in the recent past can be accounted for by compositional factors.

The paper finds that some of the wage convergence between Scottish and rUK wages from the early 2000s to 2015 can be accounted for by changes in the relative composition of the Scottish and rUK workforces. An explanatory role can be found for variables relating to demographics (the relatively more rapidly ageing Scottish workforce plus the changing role of migrant workers in the economy), qualifications, and some occupations and industrial sectors.

An implication is that there are things that the Scottish Government can do to stimulate wage growth, notably around improving skills and educational attainment, and promoting business growth especially in high-skill, high-wage sectors. But many of the compositional factors that appear to have influenced wage convergence are likely to be only tangentially related to devolved policy actions.

Changes in relative demographic structure is an obvious example, but the fortunes of most sectors are likely to be at least partially, if not largely, influenced by global economic developments.

The analysis in this paper suggests that Scotland's proportionately higher employment in the mining and quarrying sector, including offshore oil and gas extractive activities, can account for part of its wage convergence with rUK. The dramatic collapse of global oil prices since 2015 clearly owes nothing to Scottish (or UK) government policy, yet its effects are already showing up in terms of higher unemployment and slower house price growth in the North East of Scotland. If the oil price remains low, it seems inevitable that some of the compositional changes that have worked in Scotland's favour in recent years may reverse.

As well as the composition of the workforce, some of Scotland's wage convergence since 2002/3 is attributable to a decline in the unexplained 'wage structure' effect. In other words, there has been a convergence in wage returns. This might be accounted for in part as the result of convergence in unemployment rates between Scotland and rUK. The policy recommendation that would follow from this hypothesis is that the Scottish Government should maximise employment participation and seek to maintain a tight labour market, which should in theory stimulate both wage growth and employers' incentives to invest in labour productivity-enhancing activities. The Scottish Government controls some policy levers relating to this agenda (skills and employability), but many remain reserved (labour market regulation, out-of-work benefits, the overall fiscal stance).

Despite the convergence in median wages, there has been relatively little convergence between Scottish and rUK wages at the 90th percentile. As a result, there has been relatively little convergence in tax revenues per capita, given the growing importance of high earners in determining total tax revenues. This absence of convergence at the 90th percentile seems largely the result of the fact that there has been little or no convergence in unexplained wage returns in the upper part of the wage distribution.

The empirical analysis does suffer from a number of limitations. The sample size is not particularly large given what is being asked of the data, and indeed the LFS data on which the analysis is based provides a substantially different picture on wage trends than that provided by the ASHE. The robustness of estimated standard errors may be biased given some respondent duplication in the sample. Survey changes during the study period limit the extent to which explanatory variables (in particular on occupation) can provide sufficient sensitivity about workforce composition changes. These issues are accentuated when comparing Scotland with rUK, given differences in education systems and the scope for uncertainties around the equivalence of particular qualifications.

Nonetheless the general message – that the devolved Scottish Government has limited ability to influence some elements of the wage distribution via workforce composition – is likely to hold.

The obvious risk of Scotland's new devolution settlement is that, should relative income tax revenue growth fall as a result of purely exogenous factors, the settlement contains no insurance mechanism and the Scottish budget bears the full effect of the relative revenue fall. Whether this risk can be justified in the context of the Scottish Government's policy levers remains to be seen. But the less appreciated risk is that even if Scottish median wage and employment growth keep pace with that in rUK, Scotland may experience a fall in its relative budget if wages at the top of the distribution fall further behind those in rUK. Anything that weakens link between local economic performance and devolved revenue is likely to cause tension between governments, and may weaken the perceived link between tax devolution and political accountability.

6. Conclusions

Until recently, the Scottish Government has been reliant on a block grant from Westminster to finance the majority of its expenditure. This arrangement has long been criticised because it is seen to weaken the financial incentives faced by Scottish politicians and the accountability they face for their decisions, a view that chimes with fiscal federalism theory.

Recent tax decentralisation to the Scottish Parliament can be seen as the outcome of two somewhat different processes. The first round of tax decentralisation, operationalised through the Scotland Act 2012 following the Calman Commission report, was motivated very much by the objective to narrow the Scottish Parliament's VFI and thus improve the accountability of Scottish politicians for budgetary decisions. The second round of tax decentralisation following the 2014 independence referendum was partly motivated by these same objectives, but was motivated more immediately by a desire on the part of the UK Government to assuage a secessionary risk.

The first paper in this PhD (which was written in the lead-up to the 2014 independence referendum) considers issues including: which taxes are more and less suitable for decentralisation; what might be the constraints in exercising devolved fiscal powers; what might be the implications for design of block grant arrangements, and whether tax decentralisation might accentuate or assuage the secession threat.

Fiscal federalism theory provides few categorical conclusions around which taxes should be devolved to a sub-national government. The decision to devolve income tax to the Scottish Parliament might be seen as surprising given its key redistributive role and the mobility of the income tax base across UK regions. But this decision must be seen in the context of the fact that few taxes meet all the criteria for devolution, and income tax does have the advantage of being visible to the electorate and a reasonably stable source of revenue. The social security levers being devolved are closely linked to existing devolved functions in health and social care, and do not include benefits that are either very cyclically sensitive or that are particularly effective for redistributive purposes. In this perspective the fiscal powers that have been devolved to Scotland under the Scotland Act 2016 broadly resonate with the proposals that had been put forward by the various parties in the lead-up to the 2014 Referendum, and were discussed in Chapter 2.

The paper also highlights studies of strategic tax setting between governments in a federal setting which suggest that the relatively high mobility of capital and labour between Scotland and rUK may mitigate the extent to which the Scottish Government feels able to exercise new fiscal powers. Indeed, the Scottish Government has stated explicitly that it will match UK Government policy on landfill tax (to avoid waste being transported across the UK for disposal), it has expressed reluctance

to vary rates of Scottish income tax citing the risk of behavioural response if rates differ from those in England, and it mirrored the UK Government's recent introduction of a second homes levy in the Stamp Duty Land Tax. On the other hand, the Scottish Government has announced an intention to reduce and then eliminate Air Passenger Duty; if this policy is implemented, the UK Government is likely to face pressure from airports in the north of England to follow.

Fiscal federalism theory suggests that tax devolution can have positive external benefits on policy development across a federation via policy experimentation and yardstick competition. There has thus far been one clear example of tax devolution to Scotland resulting in a positive outcome from a tax design perspective. Stamp Duty in the UK has long been criticised for its 'slab structure' – the fact that properties above a certain price threshold experienced large increases in Stamp Duty liability, as higher rates above the thresholds were applied to the full value of the property rather than the value of the property above the threshold. Despite the inefficient structure, successive UK Governments had chosen not to reform the tax. Following devolution of Stamp Duty to Scotland, the Scottish Government replaced Stamp Duty with the Land and Buildings Transactions Tax (LBTT). LBTT removed the slab structure of Stamp Duty, and applies higher rates only to the proportion of property value above the relevant threshold. It was this reform that catalysed the UK Government to reform the structure of Stamp Duty to remove the slab structure.

Fiscal federalism theory provides few definitive conclusions about how the supporting institutional framework surrounding tax decentralisation should be designed. Grants to sub-central government should provide insurance against fiscal risks and support redistribution between richer and poorer regions, but without undermining the very fiscal incentives that tax decentralisation is designed to engender. The decision on where to strike this balance is ultimately a political one.

As well as criticism about the reliance of the Scottish Government on block grant to finance its expenditure, there has also long been criticism of the way in which this block grant is itself allocated. The Barnett Formula has been criticised because it takes no account of the spending needs of the Scottish Government (or the devolved administrations of Wales and NI); it has often been seen as being relatively generous to Scotland and NI but less so to Wales. There have long been calls to replace the Barnett Formula with some form of spending needs assessment. But, given that spending needs are contestable, would the UK and Scottish Governments ever be able to agree on how such a spending needs assessment would work?

The second paper in this PhD explored this issue in relation to health expenditure. The paper exploits the fact that both the UK and Scottish Governments have developed detailed funding formulae for allocating resources to territorial health boards within their respective territories. The paper takes

the formula used by the UK Government to allocate NHS resources to Primary Care Trusts in England, and applies it to Health Boards in Scotland. Effectively, the paper asks what level of resource those Health Boards would receive if the English formulae was applied on a UK-wide basis. Then, the formula used by the Scottish Government to allocate NHS funding to Scottish Health Boards is applied to PCTs in England.

The paper concludes that English and Scottish NHS healthcare allocation formulae do both provide similar estimates of the devolved territories healthcare spending needs. Subsequent analysis compared the English and Scottish approaches to allocating schools funding to local government, and came to similar conclusions.

It is also worth noting that, since paper 2 was published, the UK and Welsh Government's have agreed to underwrite the Welsh Government's block grant with a funding floor. This effectively means that the UK Government guarantees to protect Welsh spending on devolved functions from falling below 115% of equivalent English spending (HM Government, 2016). This 115% figure was arrived at through a needs assessment exercise that effectively takes spending allocations to English local authorities as the standard by which to assess Welsh spending needs. There is thus now some precedent within the UK for using needs assessment as part of the devolved governments' funding arrangements.

However, there would remain a number of significant challenges to replacing the Barnett Formula with a spending needs assessment. One is that both the health and education allocation formulae are complex, whilst most commentators argue that a UK-wide system of needs assessment should be relatively simple for transparency. But more importantly, needs assessment might become impractical where devolved governments pursue very different spending policies. Implicitly, needs assessment seems to require a broad consensus between governments about the size and scope of the public sector. This issue is becoming increasingly problematic now that major differences in public service provision are arising between England and Scotland, with Scotland preferring a more social democratic approach to public services. One example concerns the decision in England to reduce public spending on higher education, offset by the introduction of tuition fees. Another example is the trend in England, but not Scotland, toward greater use of competition and commissioning frameworks in health provision (National Audit Office, 2012).

Subsequent analysis (King and Eiser, 2016) examines how a system of spending needs assessment might work if devolved governments pursue very different policy priorities. This work concludes that, provided the DAs have a reasonable degree of tax-raising autonomy, it is theoretically possible

to design a grant allocation scheme that takes into account the spending decisions of each UK territory in determining each territory's grant.

Critics would argue that a focus on what is 'theoretically' possible ignores the inherently political nature of discussions about block grant design. But as recent experience in developing Scotland's Fiscal Framework showed, the Scottish and UK Governments can 'agree' on the nature of a fiscal settlement when they have to; but any agreement is the outcome of political bargaining positions and wider trade-offs played out across the policy spectrum, at least as much as it is about a technical exercise.

Indeed, Scotland's newly agreed Fiscal Framework retains the Barnett Formula, reflecting the 'vow' made by unionist politicians in the lead-up to the 2014 independence referendum. But the Barnett Formula determined block grant will be adjusted to reflect the revenues foregone by the UK Government when revenues are transferred to the Scottish Parliament. In future years, this block grant adjustment will grow at the same rate as the growth in comparable revenues per capita in the rest of the UK (rUK).

The implication is that the Scottish Government bears the risks in full that its devolved and assigned revenues grow relatively more slowly (per capita) than the equivalent revenues in rUK. The emphasis of this set-up is very much focussed on incentivising the Scottish Government to implement policies which grow its tax base. It thus broadly accords with principles of 'market-preserving federalism'.

But the implicit assumption underpinning the Fiscal Framework is that the Scottish Government can directly influence the scale of its revenues relative to those of rUK. The danger however is that a large part of the variation in relative tax revenue performance may be due to exogenous factors – or be the outcome of political decisions taken several years previously. Indeed, even at a national level it is clear that a large component of economic performance is exogenous to government policy (IMF, 2017), and this is likely to be even more the case at subnational level.

Papers 3 and 4 of this PhD have considered the factors that determine regional economic performance and thus devolved revenue performance. Paper 3 (Chapter 4) provides preliminary analysis, examining the question of differential regional labour market across the UK regions since the Great Recession of 2008/9.

The UK Great Recession of 2008/9 was the deepest since the Second World War, with real GDP falling by over 6%, and output remaining below its pre-recession peak for far longer than in previous recessions. Despite the sluggish performance on output growth, headline measures of labour market

performance have consistently been stronger than expected although productivity, and hence wage growth, have remained unusually slow.

The third paper in this PhD investigates the extent to which these broader national trends have played out at a regional level in the UK. It finds that there has been a strong regional dimension to the UK story. London has been at the vanguard of the recovery from recession. It regained its pre-recession employment rates far more quickly than other regions, and has now substantially surpassed them. London's superior recovery from recession runs counter to some expectations at the start of the recession given that the financial sector was at the epicentre of the crisis. But as well as experiencing a more rapid employment recovery, London has also more recently experienced larger real wage falls than other regions. It has therefore experienced an exaggerated version of the general story.

It is not entirely clear what may explain these trends. Explanations for London's relatively stronger employment growth have included structural advantages, spatially significant policies of the UK Government, and specific support to the financial sector. Neither is it clear why London has also experienced somewhat slower wage growth in recent years, (the analysis in the paper finds an inconclusive role for increased migrant employment in accounting for slower wage growth via a compositional effect). Nonetheless, what evidence there is suggests that recent differential regional labour market performance owes much to inherent regional characteristics, the effects of national (as opposed to regional or local) policy, and a variety of exogenous factors.

Differential labour market performance – and what determines it – is relevant in a devolved context because of what it implies for the devolved Scottish budget via the devolution of income tax in particular and the design of Scotland's fiscal framework.

The fourth paper in this PhD considers the factors that influence differential growth in regional income tax revenues. This analysis concludes that, as well as economic indicators such as regional wages and employment rates, differential growth in regional income tax revenues is also influenced by the way in which the income tax base is defined at UK level. The decision of successive UK Governments to increase in real terms the Personal Allowance, whilst reducing in real terms the thresholds at which the Higher and Additional Rates of income tax become payable, higher income earners account for a growing share of income tax liabilities over time. This has implications for the relative growth of regional revenues per capita.

Notwithstanding this important issue, the paper goes on to consider the compositional factors that might influence relative growth in regional wages. Wage growth is a major determinant of income

145

tax revenues, and thus understanding the factors that influence this becomes an important question for devolved policy making. Paper 4 examines empirically the factors that have contributed to convergence of Scottish wages with rUK wages between 1999 and 2016.

The results suggest that changes in the relative composition of the Scottish and rUK workforces can explain some of the wage convergence at the median. An explanatory role can be found for variables relating to demographics (the relatively more rapidly ageing Scottish workforce plus the changing role of migrant workers in the economy), qualifications, and some occupations and industrial sectors.

The lesson is that there are things that the Scottish Government can do to stimulate relative median wage growth, notably around improving skills and educational attainment, and promoting business growth especially in high-skill, high-wage sectors. But many of the compositional factors that appear to have influenced wage convergence are likely to be only tangentially related to devolved policy actions. Changes in relative demographic structure is an obvious example, but the fortunes of most sectors are likely to be at least partially, if not largely, influenced by global economic developments.

In this context it is perhaps surprising that the Scottish Government's new fiscal arrangements contain no insurance against asymmetric shocks or downside risks. Indeed at a time when the Scottish economy shows some signs of weakness (due in major part to the effect of the low global oil price on the oil and gas sector and its supply chain), the downside risks to the Scottish budget are arguably higher than the upside risks.

Under the agreed Fiscal Framework the Scottish budget will capture in full the revenue effects of relatively faster growth in revenues per capita. There is an implicit presumption in the Fiscal Framework that differential growth in revenues per capita can be entirely attributed to Scottish Government policy decisions (or at least that the Scottish Government should be held fully accountable for differential revenue growth), regardless of whether that differential revenue growth is the result of devolved policy decisions or some exogenous factor.

This outcome is somewhat different to the principle established by the Smith Commission, namely that the Scottish Government should bear the risks in full of policy decisions, rather than of revenue growth more generally. It would of course be impossible in practice to determine what proportion of differential revenue growth is attributable to devolved policy decisions (whether by a current or a previous devolved administration).

But once it is recognised that differential revenue growth may only partially be the outcome of devolved policy decisions, it is perhaps strange that the Fiscal Framework negotiations contained no

discussion as to the possibility of risk-sharing insurance against idiosyncratic asymmetric shocks. Differential revenue growth could arise as the result of a boom in the property market in south east England (which would raise revenues from stamp duty taxes in rUK relative to those in Scotland) or the decline in a sector on which the Scottish economy is disproportionately reliant. Most fiscal federal systems contain some insurance to devolved budgets against the risk of asymmetric shocks.

There may be reasons as to why protection against asymmetric shocks were not felt to be necessary. Even after the Scotland Act 2016 powers have been transferred in full, devolved and assigned revenues will still only account for half the Scottish budget, with the remainder coming from the block grant (a block grant which, as noted in Chapters 2 and 3) is sometimes perceived as being relatively generous to Scotland). Moreover, with most of the cyclically sensitive welfare benefits associated with unemployment and low-income remaining reserved to the UK Government, as well as some aspects of personal taxation (National Insurance Contributions), much of the fiscal effect of an asymmetric Scottish shock will continue to be absorbed by the UK Government. Indeed, Poghosyan et al. (2015) show that, whilst there is evidence of significant fiscal risk-sharing across state economies in the US, Australia and Canada, most of this fiscal risk sharing is channelled through the retention of cyclically sensitive revenue and spending functions at federal level; transfers to state budgets have a weak impact on cushioning regional shocks.

Protection against asymmetric shocks might also be perceived to be of limited value given that, in the recent past, the annual growth of per capita revenues raised in Scotland from the taxes to be devolved has been highly correlated with the annual growth rate of per capita revenues from the equivalent revenues in rUK (Eiser, 2017). As noted previously, the fiscal federal literature argues that the more closely correlated are regional business cycles, the weaker the case for inter-regional risk-sharing.

So, in answer to the question as to whether there is a trade-off between the objective to promote growth incentives and an objective to secure inter-regional equity in public good provision (for a given tax effort) the answer is undoubtedly yes. But the peculiarities of the Scottish case (its relatively generous existing settlement), combined with its fiscal framework arrangements (which protect the Scottish budget from its lower initial tax capacity, and only expose the Scottish budget to differential growth in tax revenues per capita, as opposed to tax revenues per se), go a long way towards mitigating the starkness of this trade-off.

But even if the risk or potential impact of asymmetric shocks to the Scottish budget is low, what scope does the Scottish Government have to grow revenues per capita relative to rUK? The presumption underpinning the rationale for tax decentralisation (particularly in the case of assigned

VAT) is that the Scottish Government can grow its budget by implementing policies to grow the economy and the tax base more quickly than the equivalent tax base in rUK. But it is far from obvious which policies might achieve this, particularly in the short-run. Whilst the Scottish Government does determine policy in a number of areas relevant to economic development – including education and skills, many elements of business support, infrastructure development, housing and planning policy – it has no remit in relation to monetary policy, product and capital market regulation, international trade, labour-market regulation, nor most areas of business taxation.

The Scottish Government could grow its revenues by raising tax rates where it has the ability to do so, although it is likely to feel constrained in this respect by the risk that tax differentials may trigger behavioural responses that effectively lower the marginal impacts on the budget of a given tax change. The jury remains out as to how strong such behavioural responses might be (Eiser, 2016), but the risk is likely to be used by the Scottish Government to justify policy conservatism – in much the same way as it has in Switzerland (Brülhart and Parchet, 2014). Indeed, whilst there is some evidence that the Scottish Government is prepared to use its new fiscal autonomy to differentiate tax policy in Scotland from rUK, assuaging some pessimism that it would feel politically too constrained to do so, this has been marginal so far (the Scottish Government setting a somewhat lower threshold for higher rate income tax in 2017/18 relative to the threshold set in rUK).

A further objective of fiscal decentralisation to Scotland, as set out by the Smith Commission, was to enhance the accountability and fiscal responsibility of the Scottish parliamentarians for their policy decisions. This objective has resonance with the idea from the second generation fiscal federalism theory that decentralisation can enhance accountability, where accountability refers to the links between policy decisions and outcomes. In this sense, there is a connection between accountability and the extent to which the workings of a devolved fiscal settlement are transparent to the electorate. However the complexity of the arrangements, including the interaction between Scottish and rUK revenue performance in determining the size of the Scottish budget, the possibility of perverse asymmetric effects arising from policy change, and the long lags between budget setting and reconciliation of outturn[15] (Eiser, 2017), all risk undermining this transparency.

---

[15] When setting its budget for a given financial year, the Scottish Government will rely on a forecast of Scottish income tax revenues that will be made by the Scottish Fiscal Commission. The UK Government will make available to the Scottish Government resources equivalent to the SFC's income tax forecast to draw down throughout the year (whilst deducting from the block grant the forecast of the block grant adjustment for income tax, which in turn will be based on OBR forecasts for rUK income tax). Scottish income tax revenues will then be collected throughout the year by HMRC. Once final outturn data is available, the forecast for Scottish income tax and the forecast for the income tax BGA will be reconciled with that outturn data. This

Will Scotland's new fiscal settlements assuage demands for Scottish independence? This question, which was raised in the first paper in this PhD, remains a legitimate question to ask given that Scotland's new fiscal settlement represents a response to promises made by unionist politicians during the independence campaign. There are arguably more important or visible influencers of the independence case. The long-term outlook for tax revenues from Scotland's offshore oil and gas sector will continue to form an important part of the independence narrative at least over the short-term, but has little direct influence on the operation of Scotland's existing fiscal settlement. Similarly the outcome of the Brexit negotiations and how these are framed by unionist and secessionist politicians will be of critical significance. The fact that the Scottish electorate voted decisevly to remain in the EU is undoubtedly creating further tension between the UK and Scottish Governments, but the issue does not yet appear to have decisively shifted public support for independence in aggregate.

But the fiscal settlement and how this is perceived to be working will also prove important. If Scotland's relative revenue performance does underperform that of rUK, this may help unionist politicians to crystalise the fiscal risks of independence. On the other hand, the constraints of the new settlement in terms of what the Scottish Parliament cannot change or influence are likely to form a key part of seccessionists' motivating arguments for independence, or at least greater fiscal autonomy. In short, the secession threat may depend on whether geopolitical and economic changes strengthen or weaken the case for fiscal sharing within the union, and whether a Scottish Government's fiscal policy constraints are perceived as being more pronounced if Scotland is within rather than outside the UK union.

In summary, fiscal devolution to the Scottish Parliament is intended to enhance the fiscal incentives of Scottish politicians, and enhance the accountability of Scottish politicians for budgetary decisions, in line with fiscal federalism theory. It is also motivated in part to assuage independence demands.

The extent to which fiscal decentralisation will achieve these aims remains unclear. On the one hand, the fact that Scottish parliamentarians are now responsible for setting income tax rates and bands resulted in significantly greater media attention and public debate around the setting of the 2017/18 Scottish budget than had hitherto been the case.

---

reconciliation process may result in increases or decreases in the Scottish budget in subsequent years. But the key point is that, because this final outturn data is not available until 15 months after the end of a financial year and reconciliation occurs in the subsequent financial year, the budgetary effects of revenues raised in 2017/18 will not materialise until budget 2020/21 (i.e. tax revenue outturn for 2017/18 will be available in July 2019, and thus be reconciled in the 2020/21 budget).

On the other hand, it remains unclear to what extent the Scottish Government will feel constrained in its ability to use its new fiscal powers, both because of the risk of tax base mobility, but also because of uncertain appetite among the electorate to see substantial policy differentiation. More generally, it is unclear whether the new fiscal responsibilities will fundamentally alter the types of policy interventions that will be pursued in Scotland, given uncertainties around how policies affect economic performance generally and tax revenues specifically.

What is clear is that the UK's devolved funding arrangements tend to evolve in response to political events and considerations, rather than on the basis of any clear principles. The recent agreement between the Conservatives and DUP is the latest example of this, and will add further weight to arguments that the UK's territorial distribution of funding is inequitable (although in this case, the blame cannot be laid at the door of the Barnett Formula). As a result, the UK's funding of the devolved territories is becoming increasingly asymmetric, with a different basket of taxes devolved in each case, an element of needs assessment being brought into the Welsh funding formula, and different mechanisms being used to adjust the Welsh block grant in light of fiscal devolution than are being used in Scotland.

Scotland's Fiscal Framework – including the mechanism for adjusting the Scottish block grant – is due to be renegotiated in 2021, and these negotiations are likely to both influence and be influenced by the political backdrop including the strength of secessionist demands at that point. It seems the only constant in Scotland's devolved fiscal settlement is change.

7. References

Alesina, Alberto and Enrico Spolaore. 1997. —On the Number and Size of Nations, *Quarterly Journal of Economics* 112 (4): 1027–56.

All Party Parliamentary Group on Taxation (2013) *Achieving autonomy: what the independence referendum means for Scotland's Fiscal Future.* House of Commons, Westminster.

Angrist, J. D. and Kugler, A. D. (2003) Protective or counter-productive? Labour market institutions and the effect of immigration on EU natives. *The Economic Journal* **113**, F302 -31.

Ashcroft B. (2013) *Has Scotland already spent its oil fund?*
URL: http://www.scottisheconomywatch.com/brian-ashcrofts-scottish/2013/07/has-scotland-already-spent-its-oil-fund.html.

Asthana S. and Gibson A. (2008) Health care equity, health equity and resource allocation: towards a normative approach to achieving the core principles of the NHS. *Radical Statistics* **96**.

Bahl R. and Cyan M. (2011) Tax assignment: does the practice match the theory?*, Environment and Planning C: Government and Policy* **29**, 264-80.

Ball R., King D. and Eiser D. (2011) *Assessing the relative healthcare spending needs of the UK's devolved territories: an English perspective.* University of Stirling Working Paper
http://www.barnett-or-needs.stir.ac.uk/Publications.html

Ball R., King D. and Eiser D. (2012) Assesing the relative healthcare needs of the UK's devolved territories: a Scottish perspective. *Environment and Planning C: Government and Policy* **30**(2), 322-346.

Ball R., Eiser D. & King D. (2015) Assessing relative spending needs of devolved government: the case of healthcare spending in the UK, *Regional Studies*, **49** (2), pp. 323-336.

Beatty C. and Fothergill S. (2011) The prospects for worklessness in Britain's weaker local economies*, Cambridge Journal of Regions, Economy and Society* **4**, 401-417.

Bell D.N.F. (2013) The Block Grant Adjustment*, Paper to Finance Committee, Scottish Parliament Accessed at:*
http://www.scottish.parliament.uk/S4_FinanceCommittee/Scotland_Act_12.pdf.

Bell D.N.F. and Blanchflower D. G. (2011) Underemployment in the UK in the Great Recession, *National Institute Economic Review* **215**, R23-33.

Bell D.N.F. and Blanchflower D. G. (2013) Underemployment in the UK Revisited, *National Institute Economic Review* **224**, F8-F22.

Bell D.N.F., Comerford D. & Eiser D. (2014) Funding pensions in Scotland: would independence matter? *National Institute Economic Review*, **227** (1), pp. R21-R31.

Bell D.N.F. and Eiser D. (2013) Inequality in Scotland: trends, drivers, and implications for the independence debate. University of Stirling Working Paper. http://www.centreonconstitutionalchange.ac.uk/publications/research-briefings/inequality-scotland-trends-drivers-and-implications-independence

Bell D.N.F. & Eiser D. (2015) The Economic Case for devolution to Scotland *National Institute Economic Review*, **233** (1), pp. R27-R36.

Bell D.N.F. & Eiser, D. (2016). Migration and fiscal policy as factors explaining the labour-market resilience of UK regions to the great recession. *Cambridge Journal of Regions, Economy and Society,* **9** (1), 197-215.

Bell D.N.F., Eiser D. & Beckmann K., (eds.) (2014) *The Economic Consequences of Scottish Independence*. Hamburg, Germany: Helmut-Schmidt-Universitat/Universitat der Bundeswehr Hamburg. http://www.scotecon.org/pdf/SES-Indy-ebook.pdfBell D., Eiser D. and Phillips, D. (2015) *Adjusting Scotland's Block Grant for new Tax and Welfare Powers: Assessing the Options* Institute for Fiscal Studies, London.

Bell D., Eiser D. and Phillips, D. (2016) *Scotland's Fiscal Framework: Assessing the Agreement*. Institute for Fiscal Studies, London. http://www.ifs.org.uk/publications/8213

Berthier A. (2015) Income tax in Scotland. Briefing 15/72. Scottish Parliament Information Centre, Edinburgh.

Besley, T., & Coate, S. (2003). Centralized versus decentralized provision of local public goods: a political economy approach. *Journal of public economics*,*87*(12), 2611-2637.

Blanchflower D. and Levin A. (2015) Labour market slack and monetary policy, *Working Paper, University of Dartmouth, Dartmouth, USA*.

Blanchflower D. G. & Oswald A. J. (2005). *The Wage Curve Reloaded,* NBER Working Paper No. 11338.

Blöchliger H. and King, D. (2006) Less than you thought: the fiscal autonomy of sub-central governments. *OECD Economic Studies* **43**(2), pp. 155-188.

Blundell R. Crawford C. and Jin W. (2014) What can wages and employment tell us about the UK's productivity puzzle?*, The Economic Journal* **124**, 377-407.

Bolton P. and Roland G. (1997) The breakup of nations: a political economy analysis*, The Quarterly Journal of Economics* **112**, 1057-90.

Bosch N. (2009) The reform of regional government finances in Spain, in Institut d'Economia de Barcelona (Ed) *World report on fiscal federalism*, pp. 58. Institut d'Economia de Barcelona, Barcelona.

Brülhart M. and Jametti M. (2006) Vertical versus horizontal tax externalities: an empirical test*, Journal of Public Economics* **90**, 2027-62.

Brülhart M. and Parchet R. (2014) Alleged Tax Competition: The Mysterious Death of Bequest Taxes in Switzerland*, Journal of Public Economics* **111**, 63-78.

Buettner T. (2008) Fiscal equalisation in Germany, in N. Bosch and J. Duran (Eds) *Fiscal federalism and political decentralisation: lessons from Germany, Spain and Canada*, pp. 137-147. Edward Elgar, Cheltenham.

Carr-Hill R., Martin S., Peacock S., Sheldon T. and Smith P. (1997) A new formula for distributing hospital funds in England. *Interfaces* **27**(1), 53-70.

Chi W. & Li B. (2008). Glass ceiling or sticky floor? Examining the gender earnings differential across the earnings distribution in urban china, 1987–2004. *Journal of Comparative Economics, 36*(2), 243-263.

Christie A. and Swales J. K. (2010) The Barnett allocation mechanism: formula plus influence?*, Regional Studies* **44**, 761-75.

Clark, K. and Drinkwater, S. The labour-market performance of recent migrants. *Oxford Review of Economic Policy*, **24**(3) 495-516.

Comerford D. & Eiser D. (2014) Constitutional change and inequality in Scotland, *Oxford Review of Economic Policy*, **30** (2), pp. 346-373.

Commission on devolution in Wales (2012) *Empowerment and responsibility: financial powers to strengthen Wales.* Commission on Devolution in Wales, Cardiff.

Commission on Scottish Devolution (2009) *Serving Scotland Better: Scotland and the United Kingdom in the 21st Century: Final Report, June 2009.* Commission on Scottish Devolution, Edinburgh.

Commonwealth Grants Commission (2010) *Report on GST Revenue Sharing Relativities - 2010 Review.* Commonwealth of Australia, Canberra.

Connolly S., Bevan G. and Mays N. (2010) *Funding and performance of healthcare systems in the four countries of the UK before and after devolution.* Nuffield Trust, London.

Dahlby B. and Wilson L. S. (2003) Vertical fiscal externalities in a federation*, Journal of Public Economics* **87**, 917-30.

Day K. M. and Winer S. L. (2001) Interregional migration and public policy in Canada: an empirical study*, Human Resources Development Canada and Statistics Canada, Ottawa*.

Deaner, B. and Phillips, D. (2014) Government spending on public services in Scotland: current patterns and future issues. Briefing Note 140, Institute for Fiscal Studies, London.

Department of Health (2009a) *Resource allocation: weighted capitation formula, sixth addition.* Department of Health, London.

Department of Health (2009b) *2009-10 and 2010-11 PCT Revenue Allocations.* Retrieved August 2, 2011, from http://www.dh.gov.uk/prod_consum_dh/groups/dh_digitalassets/@dh/@en/documents/digitalasset/dh_091447.pdf

Department of Health (2010) *Equity and Excellence: Liberating the NHS.* Department of Health, London.

Department for Work and Pensions (2012) *Impact assessment for the benefit cap* https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/220178/benefit-cap-wr2011-ia.pdf.

Dickey, H. (2014) The impact of migration of regional wage inequality: a semiparametric approach. *Journal of Regional Science*, **54**(5), 893 – 915.

Dustmann, C., A. Glitz, and T. Frattini. The Labour Market Impact of Immigration. *Oxford Review of Economic Policy* **24** (3),.

Dustmann, C., T.Frattini, and I. P. Preston. The Effect of Immigration along the Distribution of Wages *Review of Economic Studies* **80** (1), 145-173.

Dustmann, C., F. Fabbri, and I. Preston The Impact of Immigration on the British Labour Market *Economic Journal* **115** (507), F324-F341.

Eiser, D. (2016) Does Nicola Sturgeon's stance on the 50p tax rate undermine the case for tax devolution? Holyrood Magazine, March 2016 https://www.holyrood.com/articles/inside-politics/does-nicola-sturgeon%E2%80%99s-stance-50p-tax-rate-undermine-case-tax-devolution

Eiser, D. (2017) Fiscal decentralisation, fiscal incentives and accountability: the implications of Scotland's new fiscal settlement. Working Paper, University of Strathclyde.

Elliot R., Bell D., Scott A., Ma A. and Roberts E. (2005) Devolved government and public sector pay reform: Considerations of equity and efficiency*, Regional Studies* **39**, 519-39.

Esteller-Moré Á. and Solé-Ollé A. (2001) Vertical income tax externalities and fiscal interdependence: evidence from the US*, Regional Science and Urban Economics* **31**, 247-72.

Faggian A., Rajbhandari I. and Dotzel, K. (2017) The interregional migration of human capital and its regional consequences: a review. *Regional Studies* **51**(1), 128 – 143.

Feld L. (2011) Written evidence to the Scottish Parliament Scotland Bill Committee. Available at: http://archive.scottish.parliament.uk/s3/committees/scotBill/documents/49.SBProfessorLarsPFeld.pdf

Feld L. P. and Kirchgässner G. (2003) The impact of corporate and personal income taxes on the location of firms and on employment: some panel evidence for the Swiss cantons*, Journal of Public Economics* **87**, 129-55.

Fingleton B., Garretsen H. and Martin R. (2012) Recessionary shocks and regional employment: evidence on the resilience of UK regions*, Journal of Regional Science* **52**, 109-33.

Firpo S., Fortin N. M. & Lemieux, T. (2009). Unconditional quantile regressions. *Econometrica, 77*(3), 953-973.

Firpo, S., Fortin N. & Lemieux T. (2007). Decomposing wage distributions using recentered influence function regressions. *University of British Columbia (June 2007)*

Forbes T, Evans D. and Scott N. (2010) Implementing health and social care policy: England and Scotland compared. *Policy Studies* **31**(6), 591-611.

Friedman, S. and MacMillan, L. (2017) Is London really the engine room? Migration, opportunity hoarding and regional social mobility in the UK. *National Institute Economic Review*, **240**, R58 – R72.

Fujita M., Krugman P. R. & Venables, A. (2001). *The spatial economy: Cities, regions, and international trade* MIT press.

Galbraith J. and Stone M. (2011) The abuse of regression in the National Health Service allocation formulae: response to the Department of Health's 2007 'resource allocation research paper'. *Journal of the Royal Statistical Society A* **174**(3), 547-567.

Galego A. and Pereira J. (2013). Decomposition of regional wage differences along the wage distribution in Portugal: the importance of covariates. CEFAGE-UE Working Paper 2013/16.

Garrett G. and Rodden J. (2003) Globalization and fiscal decentralization*, In Kahler M. and Lake D. (Eds.) Governance in a global economy: Political authority in transition*, 87-109, Princeton University Press.

Gibbons S., Overman H. G. & Pelkonen P. (2010). Wage disparities in Britain: People or place? Spatial Econometrics Research Centre (SERC) Discussion Paper No.60, London School of Economics.

Glaeser, E. L. (1999). Learning in cities. *Journal of Urban Economics, 46*(2), 254-277.

Gordon I. (2016) Quantitative Easing of an International Financial Centre: How Central London Came So Well Out of the Post-2007 Crisis. Spatial Economics Research Centre Discussion Paper SERCDP0193, London School of Economics.

Gordon I. and Kaplanis I. (2014) Accounting for big city growth in low paid occupations: immigration and/or service class consumption. *Journal of Economic Geography* 90(1) 67-90.

Gregg P., Machin S. and Fernández-Salgado M. (2014a) The squeeze on real wages–and what it might take to end it, *National Institute Economic Review* **228**, R3-R16.

Gregg P., Machin S. and Fernández-Salgado M. (2014b) Real wages and unemployment in the big squeeze, *The Economic Journal* **124**, 408-32.

Green A. and Livanos, I. (2015) Involuntary non-standard employment and the economic crisis: regional insights from the UK, *Regional Studies* **49**(7), 1223-1235.

Greer S. (2004) *Four way bet: how devolution has led to four different models for the NHS.* UCL Constitution Unit, London.

Haldane A. (2015) Drag and Drop, *Speech, 19 March 2015, Rutland.*
www.bankofengland.co.uk/publications/Pages/speeches/default.aspx

Hallwood P. and MacDonald R. (2009) *The political economy of financing Scottish government.* Edward Elgar, Cheltenham.

Hanlon, P. Lawder, R. Buchanan, D. Redpath, A. Walsh, D. Wood, R. . . . Chalmers, J. (2005) Why is mortality higher in Scotland than in England and Wales? Decreasing influence of socioeconomic deprivation between 1981 and 2001 supports the existence of a 'Scottish effect'. *Journal of Public Health* **27**(2), 199-204.

Heald, D. (2009) Memorandum by Professor David Heald, in House of Lords Select Committee on the Barnett Formula *The Barnett Formula: Report with evidence*, pp. 316-324. HL Paper 139. House of Lords, London.

Heald, D. (2014) Page 5 pf written evidence to the Scottish Parliament's Finance Committee, June 2014.
http://www.parliament.scot/S4_ReferendumScotlandBillCommittee/20141211Public_Papers(1).pdf

Hilber C. A. L., Lyytikäinen T. and Vermeulen W. (2011) Capitalization of central government grants into local house prices: Panel data evidence from England, *Regional science and urban economics* **41**, 394-406.

HM Government (2010) The Coalition Agreement: our programme for government.
https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/78977/coalition_programme_for_government.pdf

HM Government (2016) The Welsh Fiscal Framework.

http://gov.wales/docs/caecd/publications/161219-fiscal-agreement-en.pdf

HM Treasury (2010a) *Funding the Scottish Parliament, Welsh Assembly Government, and Northern Ireland Assembly: statement of funding policy.* HM Treasury, London.

HM Treasury (2010b) *Budget 2010.* The Stationary Office, London.

HM Treasury (2011) *Rebalancing the Northern Ireland economy*. The Stationary Office, London.

HM Treasury (2012) *Public Expenditure Statistical Analysis.* The Stationary Office, London.

HM Treasury (2016) Scotland's Fiscal Framework.

https://www.gov.uk/government/publications/the-agreement-between-the-scottish-government-and-the-united-kingdom-government-on-the-scottish-governments-fiscal-framework

Holmes C. & Mayhew K. (2012). *The changing shape of the UK job market and its implications for the bottom half of earners.* Resolution Foundation, London.

House of Commons Health Committee (2009) *Health Inequalities: Third report of session, 2008-9.* The Stationary Office, London.

Home Rule and Community Rule Commission (2014) *Cambell II.* Scottish Liberal Democrats, Edinburgh.

IMF (2017) Are countries losing control of domestic financial conditions? Chapter 3 of Global Financial Stability Report, April 2017. IMF, Washington.

Independent Commission on Funding and Finance for Wales (2010) *Fairness and accountability: a new funding settlement for Wales.* Welsh Government, Cardiff.

Industrial Communities Alliance (2015) *Whose recovery? How the upturn in economic growth is leaving older industrial Britain behind.* Industrial Communities Alliance, Barnsley

Jeffery C. (2012) Dis-united Kingdom? *Public Policy Research* **19**(1), 14-16.

Jeffery C. (2014) *Devolution Commission reflections: comparisons and contrasts* URL: http://www.futureukandscotland.ac.uk/blog/devolution-commission-reflections-comparisons-and-contrasts.

John P. and Ward H. (2001) Political manipulation in a majoritarian democracy: central government targeting of public funds to English subnational government, in space and across time. *British Journal of Politics and International Relations* **3**(3), 308-339.

Jones P. S. and Green A. E. (2009) The quantity and quality of jobs: changes in UK regions, 1997–2007*, Environment and Planning A* **41**, 2474.

Kay A., Bristow G., McGovern M. and Pickernell, D. (2005) Fair division or fair dinkum? Australian lessons for intergovernmental fiscal relations in the United Kingdom. *Environment and Planning C* **23**, 247-261.

King D. and Eiser D. (2016) Reform of the Barnett Formula with needs assessment: can the challenges be overcome?*, Regional Studies*, **50**(5), 790-804.

King, D. and Eiser, D. (2017) Allocating Grant to the UK's Devolved Territories by Needs Assessment: Lessons from School Funding Formulae in England and Scotland *Fiscal Studies* 38(1), 81-109*.Fiscal*

King D., Pashley M., and Ball R. (2007) Scotland's social services spending needs: an English view. *Environment and Planning C: Government and Policy* **25**, 918-940.

Lago-Peñas I., Lago-Peñas S. and Martinez-Vazquez, J. (2011) Guest Editorial, The political and economic consequences of decentralisation. *Environment and Planning C* **29**(2), 197-203.

Laws,J. (2008) Constitutional Guarantees, *Statute Law Review* **29** (1): 1-10

Lecours A. and Béland D. (2010) Federalism and fiscal policy: The politics of equalization in Canada*, Publius: The Journal of Federalism* **40**, 569-96.

Lluch J. (2012) Autonomism and federalism*, Publius: The Journal of Federalism* **42**, 134-61.

Local Government Association. (2013) Press Release, 2 July 2013: *Ten point plan to revive England's economy and public services*

Lodge G., Wincott D., Henderson A., and Jones R. W. (2012) *The dog that finally barked: England as an emerging political community.* Institute for Public Policy Research, London.

Mackay R. and Williams J. (2005) Thinking about need: public spending on the regions*, Regional Studies* **39**, 815-28.

Madson, J.B. and Andric, S. (2017) The immigration-unemployment nexus: do education and
Protestantism matter? *Oxford Economic Papers*, **69**(1), 165 – 188.

Manacorda, M., A. Manning, and J. Wadsworth. The Impact of Immigration on the Structure of Male
Wages: Theory and Evidence from Britain, *Journal of the European Economic Association* **10**
(1), 120-151.

Martin R. (2012) Regional economic resilience, hysteresis and recessionary shocks*, Journal of
Economic Geography* **12**, 1-32.

Martin R., Gardiner B. and Tyler P. (2014) *The evolving economic performance of UK cities: city
growth patterns 1981-2011.* Government Office for Science, London.

Martin R., Pike A., Tyler P. and Gardiner, B. (2015) Spatially rebalancing the UK economy: the need
for a new policy model. Regional Studies Association.

Migration Advisory Committee. Analysis of the Impacts of Migration. Home Office, London, 2012.

McAfferty, I. (2014) *The UK productivity puzzle: a sectoral perspective. Speech,19 June 2014*
www.bankofengland.co.uk/publications/Pages/speeches/default.aspx

McLaren J. and Armstrong J. (2014) Scotland's Economic Performance and the Fiscal Implications of
Moving to Independence*, National Institute Economic Review* **227**, R3-R13.

McLaren J., Armstrong J. and Harris R. (2010) *Spending on health.* Centre for Public Policy for
Regions, Glasgow.

McLean I., Lodge G. and Schmuecker K. (2008) *Fair Shares: Barnett and the Politics of Public
Expenditure.* IPPR, London.

McLean I. and McMillan A. (2005) The distribution of public expenditure across the UK regions*, Fiscal
Studies* **24**, 45-71.

McVicar D. (2006) Why do disability benefit rolls vary between regions? A review of the evidence
from the USA and the UK*, Regional Studies* **40**, 519-33.

Midwinter A. (2002) Territorial resource allocation in the UK: a rejoinder on needs assessment*,
Regional Studies* **36**, 563-77.

Miller H. and Pope T. (2016). *The changing composition of UK tax revenues.* London: Institute for Fiscal Studies (BN182).

Milligan K. and Smart M. (2013) The Devolution of the Revolution: Taxation of High Incomes in a Federation. University of Carleton Working Paper http://carleton.ca/economics/wp-content/uploads/seminar-paper-131108.pdf

Morgan K. (2001) The new territorial politics: rivalry and justice in post-devolution Britain. *Regional Studies* **34**, 343-48.

National Audit Office (2011) *Landscape Review: Formula Funding of Local Public Services.* Audit Office, London.

National Audit Office (2012) *Healthcare across the UK: a comparison of the NHS in England, Scotland, Wales and Northern Ireland.* The Stationary Office, London.

NHS Scotland Resource Allocation Committee (2007) *Delivering Fair Shares for Health in Scotland.* NHS Scotland, Edinburgh.

Nuffield Trust (2009) *Developing a person-based resource allocation formula for allocations to general practices in England.* Nuffield Trust, London.

Oates, W. E. (1972). Fiscal federalism.

Oates, W. E. (2005). Toward a second-generation theory of fiscal federalism. International tax and public finance, 12(4), 349-373.

OBR (2016) Forecast evaluation report. http://budgetresponsibility.org.uk/docs/dlm_uploads/Forecast-evaluation-report-October-2016-1.pdf

OECD (2013) *Fiscal federalism 2014: making federalism work.* OECD, Paris.

Pessoa J. P. and Van Reenen J. (2014) The UK productivity and jobs puzzle: does the answer lie in wage flexibility? *The Economic Journal* **124**, 433-52.

Phillips, D. (2014), *Business as usual? The Barnett formula, business rates, and further tax devolution*, IFS Briefing Note No. 155, Institute for Fiscal Studies (available at http://www.ifs.org.uk/publications/7442)

Poghosyan, T., Senhadij, A., and Cottarelli, C. (2014) The role of fiscal transfers in smoothing economic shocks. In Cottarelli, C. and Guerguil, M. (Eds) Designing a European Fiscal Union, Routledge.

Pope T., Roantree B. & Grace C. (2015). *A survey of the UK tax system.* (Briefing Note No. BN09). London: Institute for Fiscal Studies.

Propper C. and Van Reenan J. (2010) Can pay regulation kill? Panel data evidence on the effect of labor markets on hospital performance*, Journal of Political Economy* **118**, 222-73.

Reed, H. and M. Latorre, The Economic Impacts of Migration on the UK Labour Market, *IPPR Economics of Migration Working Paper 3*, Institute for Public Policy Research, London, 2009.

Reform Scotland (2012) *DevoPlus: A stronger Scotland within the UK.* Devo Plus Group, Edinburgh.

Revelli F. (2005) On spatial public finance empirics*, International Tax and Public Finance* **12**, 475-92.

Rice P., Venables A. J. & Patacchini E. (2006). Spatial determinants of productivity: Analysis for the regions of Great Britain. *Regional Science and Urban Economics, 36*(6), 727-752.

Riley R., Rosazza-Bondibene C. and Young G. (2014) The Financial Crisis, Bank Lending and UK Productivity: Sectoral and Firm-Level Evidence*, National Institute Economic Review* **228**, R17-34.

Sala G. (2014) Federalism without adjectives in Spain*, Publius: The Journal of Federalism* **44**, 109-34.

Scottish Labour Devolution Commission (2014) *Powers for a purpose: strengthening accountability and empowering people.* Scottish Labour.

Scottish Government (2016) Government Expenditure and Revenue Scotland (GERS) 2015/16. Scottish Government, Edinburgh.

Seabright P. (1996) Accountability and decentralisation in government: An incomplete contracts model*, European Economic Review* **40**, 61-89.

Select Committee on the Barnett Formula (2009) *The Barnett Formula.* The Stationary Office, London.

Smith Commission (2014), *Report of the Smith Commission for further devolution of powers to the Scottish Parliament*, (available at https://www.smith-commission.scot/smith-commission-report/)

Smith P. C. (2006) *Formula funding of public services.* Routledge, Abingdon.

Smith P. C. (2003) Formula funding of public services: an economic analysis. *Oxford Review of Economic Policy* **19**(2), 301-322.

Sorens J. (2011) The institutions of fiscal federalism*, Publius: The Journal of Federalism* **41**, 207-31.

Stone M. and Galbraith J. (2006) How not to fund hospital and community health services in England. *Journal of the Royal Statistical Society A* **169**(1), 143-164.

Travers, T. (2015) Devolving funding and taxation in the UK: a unique challenge, *National Institute Economic Review* **233**, R5-R13

Trench A. (2012) *Devo more, devo plus and so on: extending devolution in the UK, and financing it* Presentation for PSA Territorial Politics Group Conference.

Trench A. (2013) *Funding Devo More: fiscal options for strengthening the Union.* IPPR, London.

UK Parliament (2012) *The Scotland Act*. The Stationary Office, London.

Tommasi, M., & Weinschelbaum, F. (2007). Centralization vs. Decentralization: A Principal-Agent Analysis. Journal of public economic theory, 9(2), 369-389.

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

Young R. (2002) *Economic factors in the secession calculus: A survey* The Conference on Economics of Political Integration and Disintegration, Catholic University of Louvain.