The Importance of Graduates for the Scottish Economy: A “Micro-to-Macro” Approach

Kristinn Hermannsson
Katerina Lisenkova
Patrizio Lecca
Peter G McGregor
J Kim Swales

University of Strathclyde
The Importance of Graduates for the Scottish Economy:
A “Micro-to-Macro” Approach

Kristinn Hermannsson†
kristinn.hermannson@strath.ac.uk
Katerina Lisenkova†
katerina.lisenkova@strath.ac.uk
Patrizio Lecca†
patrizio.lecca@strath.ac.uk
Peter G McGregor†
p.mcgregor@strath.ac.uk
J Kim Swales†
j.k.swales@strath.ac.uk

† Fraser of Allander Institute, Department of Economics, University of Strathclyde

* The authors gratefully acknowledge the support of the ESRC and the Higher Education Funding Bodies of the UK for The Overall Impact of HEIs on Regional Economies (ESRC, RES-171-25-0032). This project is one of nine that are funded under the joint UK Higher Education Funding bodies and ESRC Initiative on The Impact of HEIs on Regional Economies. All UK Higher Education funding bodies are involved: the Scottish Funding Council, the Higher Education Funding Councils for England and Wales, and the Department for Employment and Learning Northern Ireland.
Abstract

There have been numerous attempts to assess the overall impact of Higher Education Institutions on regional economies in the UK and elsewhere. There are two disparate approaches focusing on: demand-side effects of HEIs, exerted through universities’ expenditures within the local economy; HEIs’ contribution to the “knowledge economy”. However, neither approach seeks to measure the impact on regional economies that HEIs exert through the enhanced productivity of their graduates.

We address this lacuna and explore the system-wide impact of the graduates on the regional economy. An extensive and sophisticated literature suggests that graduates enjoy a significant wage premium, often interpreted as reflecting their greater productivity relative to non-graduates. If this is so there is a clear and direct supply-side impact of HEI activities on regional economies through the employment of their graduates. However, there is some dispute over the extent to which the graduate wage premium reflects innate abilities rather than the impact of higher education per se.

We use an HEI-disaggregated computable general equilibrium model of Scotland to estimate the impact of the growing proportion of graduates in the Scottish labour force that is implied by the current participation rate and demographic change, taking the graduate wage premium in Scotland as an indicator of productivity enhancement. We conduct a range of sensitivity analyses to assess the robustness of our results.

While the detailed results do, of course, vary with alternative assumptions about future graduate retention rates and the size of the graduate wage premium, for example, they do suggest that the long-term supply-side impacts of HEIs provide a significant boost to regional GDP. Furthermore, the results suggest that the supply-side impacts of HEIs are likely to be more important than the expenditure impacts that are the focus of most “impact” studies.

Keywords: Supply side impact; higher education institutions; computable general equilibrium model. JEL Codes: I23, E17, D58, R13.
1. Introduction and background

The numerous past studies of the regional impacts of higher education institutions (HEIs) fall into two categories, focussing on either the demand- (expenditure) side or the supply-side, “knowledge economy”, effects of HEIs on regional economies (see e.g. Florax, 1992, for an early review.)

The demand-side literature explores the “expenditure impacts” of HEIs, typically including a part of their students’ expenditures. These all employ some kind of “multiplier” analysis, focussing on HEIs as a sector that is the source of indirect and induced demand in the home region, through their intermediate purchases and employment demands. A number of these studies have a Scottish focus (Blake & McDowell, 1967; Brownrigg 1973, Battu et al 1998; Kelly et al 2004; Hermannsson et al, 2010a,b).

In contrast analyses of the contribution of HEIs to the “knowledge economy”, relate to the impact of HEIs on the supply side of regional goods markets. Here the focus is often “interregional” in the sense of impacts being transmitted over spatial boundaries where distance matters. The approach began by incorporating spatial effects more effectively into a knowledge production function in which the impact of HEIs is separately identified (Jaffe (1989)).\(^1\) In a wider context, studies of the knowledge economy encompass a broad range of typically more descriptive, case-study-based approaches, though the generality of their results is questionable (see e.g. Goldstein (2009)).\(^2\) Furthermore, many of these analyses are microeconomic in orientation, and so do not fit in an obvious way with the system-wide focus of the expenditure impact studies of HEIs.

---

1 See e.g. Anselin et al (1997) and Varga (1998) for early examples. Acs (2009) provides a review of these and subsequent developments of this approach.

2 There is recent UK evidence that strongly suggests that the “bugs and drugs” conception of “knowledge transfer” that has often been the focus of this literature is unwarranted: active knowledge exchange occurs across a very wide range of subjects areas. See Kitson et al, (2010).
If we want to understand the system-wide impacts of HEIs on regional economies, which is presumably of interest to both regional and national governments, the existing literature has two significant limitations. Firstly, studies of expenditure impacts focus exclusively on the impacts of the host region and assume an entirely passive supply side that seems to preclude, for example, any meaningful consideration of the transmission mechanisms from HEIs to regional economies that are emphasised by the “knowledge economy” literature. Secondly, the knowledge transfer literature tends generally to focus on micro-(or meso)-economic aspects, with no means of assessing system-wide impacts. Moreover, the scope of this literature does not extend to a comprehensive account of the supply-side impacts of HEIs (though nor does it profess to do so).

Accordingly, we have two completely disparate literatures on the regional impacts of HEIs that are seemingly irreconcilable in terms of their underlying vision of regional economies. Furthermore, in terms of their coverage of the regional impacts of HEIs they are not comprehensive. The most striking and important omission is that there is little attempt to provide a quantitative estimate of the impact of graduates on the host regional economy (but see Bradley and Taylor, 1996; Florax 1992), although there has indeed been recognition, and attempted measurement, of the potential role of graduate migration flows as an element of the knowledge transfer mechanism (Faggian and McCann, 2006; Anderson et al., 2009). This omission seems to be extremely serious given that the production of human capital is so fundamental to what HEIs actually do. The production of human capital embodied in their graduates is a crucial dimension of HEI activity, but is one that is currently neglected in studies of regional impacts. Of course, all contributors recognize this role and its potential importance, but neither of the main regional literatures makes any attempt to measure its impact at a system-wide level. In this paper we address

---

3 It may be objected that there is no reason why regional governments should be concerned about the impacts of their own region’s HEIs on other economies. However, if the scale of these is sufficient there maybe feedback effects to the own-region, which may be either positive or negative. Furthermore, if these spillover effects are positive they could possibly be used to negotiate an improved fiscal “deal” with the national government. They are also likely to be interested in the impact of other regions HEIs on their region.

4 Though Varga et al (2010) is an exception. This is a multi-level modelling approach that combines micro-econometric analysis of knowledge production functions with static spatial CGEs to explore medium-term tendencies to spatial concentration or dispersion, and a DSGE macroeconomic model to determine dynamics of adjustment. Our approach differs in: incorporating all of the impacts within a single framework; allowing for a fuller range of HEI impacts.
this gap using a “micro-to-macro” approach that exploits the micro-econometric evidence on the impact of HEIs to simulate the overall impact of graduates on the Scottish economy using an HEI-disaggregated computable general equilibrium model of the Scottish economy.

It might be objected that, even if this gap is apparent in the regional literatures, it is surely not present at the national level. We contend that, in fact, it is. We do, of course, know a great deal about some aspects of the impact of higher education at the national level. In a microeconomic context, the impact of higher education on private market benefits has been studied extensively and reviewed e.g. by Blundell et al (1999, 2005), Psacharaopoulos and Patrinos (2004) and Machin and McNally (2007). While the regional dimension has not been central to this literature, it has been accommodated within it. However, this literature is micro-oriented and again seems to have little in common with the macroeconomic expenditure impact studies of HEIs.

There also exists a literature on the macroeconomic returns to education in general, and a smaller literature on the macroeconomic returns to higher education. (See Sienesi and Van Reenen (2004) for a general review of macroeconomic returns and Gemmell (1997) for a review of growth theory and its relation to higher education impacts.) These macroeconomic growth models (which incorporate a disaggregated labour input) are typically regarded as capturing the total (social plus private) returns to education, from which private returns may be subtracted to yield an estimate of the social returns. They cannot therefore capture the system-wide counterpart to the private market returns to higher education.\(^5\) Furthermore, this approach is not easily reconciled with the extant “knowledge economy” literature, although many of the fundamental ideas are common to both literatures, particularly through linkages to Romer’s (1986, 1990) R&D-oriented variant of endogenous growth theory. Finally, the approach does not identify the transmission mechanisms from higher education to aggregate economic activity. Here we

\(^5\) The precise interpretation of the returns varies, of course, with the precise specification of the higher education variables included in the growth equation.
seek to develop an integrative, “micro-to-macro” approach, which overcomes some of these limitations.

It is surely desirable to be able to explore the impacts of both demand and supply effects of HEIs in a single, unified framework. Furthermore, this framework should be capable, at least in principle, of accommodating many of the HEI impacts that have been identified through micro-econometric estimation. We believe that a regional computable general equilibrium (CGE) approach can be useful in the present context and illustrate this here with a model of Scotland. On the one hand this accommodates the multi-sectoral structure of IO systems, and can be used to identify the demand effects of an aggregate Scottish HEI sector on the economy of Scotland. Indeed, such a system emulates the behaviour of an augmented regional IO model of comparable aggregative structure for a demand disturbance under passive supply conditions (e.g. McGregor et al, 1996), but its applicability is not restricted to such conditions.

On the other hand, the model can also be used to simulate the supply side effects of HEIs, whether through the impact of its graduates on host regions, or through technological spillovers of the kind emphasized by the literature on the knowledge economy. Our approach is “micro-to-macro” in that we begin by seeking to identify the supply-side transmission mechanisms that operate at the micro/meso-level, use the available evidence to specify and calibrate them, and then simulate their system-wide impact through a regional CGE model. In terms of previous literature our analysis is closest to that of Giesecke and Madden (2006), who analyse the case of the University of Tasmania, whereas we consider Scottish HEIs as a whole, and focus on the impact of projected increases in the proportion of graduates in the Scottish labour force on the Scottish economy. Our approach therefore also serves to provide an integrated demographic-economic analysis of HEI impacts.

In this paper we focus exclusively on the system-wide impacts of the productivity enhancing effects of graduates. We simplify our approach to the extent that the available evidence seems to allow, and handle unavoidable uncertainties through sensitivity analysis. In Section 2 we briefly review the evidence on the graduate wage premium and
its usefulness as a measure of productivity differences between graduates and non-graduates. In Section 3, we motivate our approach, and our assumption of a constant wage premium and productivity differential despite a projected increase in the proportion of graduates in the Scottish labour force, while Scotland’s graduate “retention rate” is maintained. We calculate a productivity-adjusted Scottish labour force, and simulate the impact of the implied (labour-augmenting) productivity stimulus on a range of macroeconomic and sectoral outcomes using our HEI-disaggregated computable general equilibrium (CGE) model of Scotland. The results of our simulations, which we report in Section 4, illustrate the likely orders of magnitude of the impact of graduates on the Scottish economy if current higher education policy is maintained, and demand for graduates keeps pace with the supply, in line with recent trends. Any one of our assumptions may, of course, prove to be incorrect, but we seek to illustrate their likely importance through a sensitivity analysis in which we vary our assumptions about the evolution of the labour force, the scale of the wage premium and the strength of the signalling effect. We conclude in Section 5, where we discuss the implications of our analysis and identify possible extensions.

2. Graduate wage premia and productivity differentials

We briefly review the evidence on the graduate wage premium and then consider the evidence on the extent to which this reflects a genuine productivity differential. We use this review of the evidence to inform the simplifying assumptions that we adopt when we discuss the implementation of our micro-to-macro approach in Section 3.

2.1 Graduate wage premia

One of the most striking features of the graduate labour market over the last few decades is the apparent insensitivity of the graduate wage premium to the scale of the increase in the HEI participation rate. Scotland, as well as the rest of the UK, has recently
experienced a significant increase in higher education participation rates\(^6\) (see Figure 1). The participation rate for men has increased from 19.5\% in 1984 to 41.2\% in 2007. For women the change is even more marked, from 18.2\% in 1984 to 52.9\% in 2007. Recently, there has been a decline for both men and women. Other things being equal, we would expect such a major increase in the supply of graduates to result in a fall in their “price”, but the graduate wage premium has exhibited remarkable stability over the period.

**Figure 1. Higher Education Age Participation Index, Scotland, 1983/84-2007/08**

\(^6\) As a measure for participation rate we use age participation index. The Age Participation Index measures the number of new young (under 21) Scottish entrants to HE divided by the number of 17 year-olds in Scotland. For more details see [http://www.scotland.gov.uk/](http://www.scotland.gov.uk/)
The longest wage premium series for Scotland can be found in Walker and Zhu (2007). They report graduate wage premia separately for men and women and for different cohorts for 3-year groups starting from 1996 until 2005. They define the graduate wage premium as wage of graduates relative to A-levels. The aggregate graduate wage premium for the period 1996-2005 is mostly constant – it increased slightly for men from 28% to 35% and decreased slightly for women from 45% to 41%.

If we look for British-wide evidence, O’Leary and Sloan (2005) report graduate wage premia for Great Britain disaggregated by earnings quartile, subject and cohort. They find that between 1993 and 2003 the wage premium for men was largely stable, while that for women has declined. The breakdown reveals that the decline for women is more pronounced at the bottom of skill/earnings distribution, is more concentrated in Arts than in other disciplines and the effect is much stronger among recent cohorts of graduates. O’Leary and Sloan (2005) explain these by differences in supply of different types of graduates.

Walker and Zhu (2008), using a somewhat different methodology, but essentially the same data set (LFS for 1994-2006), compare wage premia for pre-expansion and post-expansion cohorts, but find no statistically significant decline for men and, remarkably, weakly significant 10% increase for women. They explain this by possible upward ability bias of OLS estimates. One possible way of controlling for it would be IV estimates. However, numerous studies using institutional supply constraints as instruments find that estimated in this way returns to education are typically as big or bigger than the corresponding OLS estimates. For a review see Card (2001).

Recent evidence for the UK is therefore a little mixed. However, given the dramatic increase in the relative supply of graduates observed in recent decades, the graduate wage premium seems remarkably insensitive to this. Furthermore, this evidence is not restricted to the UK’s experience (e.g. Machin and McNally (2007)). However, there is, of course, no “law” in operation here. For example, Goldin and Katz (2007) in their excellent
analysis of century-long history of returns to education in the US show that over the past century the college wage premium fluctuated between 30% and 60%.

The graduate wage premium might also change over time in response to the quality of graduates. There is an argument that relative “quality” of graduates is going to decrease as participation in higher education increases. This argument is based on the assumption that potential entrants into HEIs are ordered according to their abilities and thus, as participation rate increases less able individuals are able to get into the higher education. This, however, will not necessarily be the case. Depending on what is the main reason for non-participation – low returns to education caused by low ability or high cost and supply constraints. If the first reason predominates the relative “quality” of graduates will decrease as participation increase. However, if the latter reason dominates, the relative “quality” of graduates can actually increase as participation widens. Both theoretical and empirical studies showed that it can change ether way (Card, 2001; D’Amato and Mookherjee, 2008; Freeman, 1996; Galor and Zeira, 1993; Mookherjee and Ray, 2003). This factor, however, should not affect our scenarios because we are not projecting large increases in the participation rate of the relevant age cohort.

We use the evidence of the comparative constancy of the graduate wage premium in recent UK history to motivate an important simplifying assumption, in which we treat human capital as homogenous. The difference between graduates and non-graduates is simply the quantity of human capital that these two groups possess on average. This approach allows us to treat the labour market as unified, and so avoid a number of complexities. Graduates and non-graduates are treated like perfect substitutes; it is “as if” it simply takes more non-graduates to perform the same task as graduates. Our sensitivity analysis provides some feel for the significance of this assumption.

2.2 From graduate wage premia to productivity differentials

In the absence of direct measures of productivity it is common to assume that productivity is closely correlated with observed wage rates. We follow this approach and assume that the graduate wage premium, at least to a significant degree, reflects the higher
productivity of graduates. For our purposes, however, it is important to understand how much of this wage differential can be attributed to the impact of higher education. The correlation between earnings and education is a well-established fact. The presence of correlation, however, is not sufficient to establish causality. There are two main strands of literature on this matter.

The Human capital school has its origins in the works of Mincer (1958), Schultz (1960) and Becker (1964, 1975). This tradition maintains that education directly increases human capital, which in turn increases the productivity of workers. Education spending should be viewed in this case as an investment in human capital. According to the human capital school the graduate wage premium represents the payment for the higher productivity of more educated workers that is directly attributable to their university education. This approach is very straightforward and appealing. Education is viewed as being a productive sector creating skills. However, not all researchers agree.

An alternative perspective is that of the sorting school. This stems from the works of Spence (1973) and Stiglitz (1975). The most extreme version of this theory maintains that education does not enhance human capital (and as a consequence productivity), but simply serves the purpose of revealing innate ability to employers. The sorting is required because of the presence of information asymmetry, in particular, the inability of the employer to observe the “true” productivity of the worker. The impact of asymmetric market information was first described by Akerlof (1970), using the example of the used cars market. In this model buyers are only willing to pay the price of the average quality goods presented on the market. This average is going to adjust until only worst quality “lemons” are traded. Both parties are interested in avoiding this situation, which is where sorting comes into play. Depending on which party makes the first move – informed or uninformed – sorting takes the form of either signalling – informed party acquires a signal – or screening – uninformed party sets a screen. In the context of labour market and education, it is more common to talk about “signalling” and we follow this convention.

Returning to the labour market example, workers with higher ability will try to use a signal to differentiate themselves from the workers with lower ability. In our case this
signal is the level and quality of education. For this model to work, the cost of acquiring a signal should be negatively correlated with ability. Those with higher ability would choose to acquire a “stronger” signal – more and better quality education – because it is less costly for them. In equilibrium ability would be perfectly revealed by the strength of a signal.

The literature differentiates two types of signalling hypothesis: strong signalling – education has no impact on productivity and serves only as sorting device, and weak signalling – education both improves productivity and provides a signal for the market. Most of the empirical studies fail to accept the strong signalling hypothesis. However, weak signalling effects are commonly observed in empirical data. In this context signalling is the part of the productivity differential, and therefore of the wage premium, that is not caused by education per se.

Signalling theory is often thought to imply that graduates are overpaid, i.e. they are less productive than their wage premium would suggest. This is not true. Both theories predict the same equilibrium outcome – more productive individuals have higher level of schooling and are paid more. However, the role of education in each case is different. In human capital theory, education is responsible for the increased productivity, while in the limiting case of signalling theory productivity is causally independent of education. Hence, one theory predicts a causal link between education and productivity, while the other one (at least in its limiting form) does not. Because we are trying to measure an effect of HEIs on the economy through their impact on labour force productivity, it is important to determine which part of the wage premium can be directly attributed to the impact of HEIs.

The main difficulty in testing signalling versus human capital theory is that, as was mentioned above, they predict observationally equivalent equilibrium outcomes. For a formal discussion see Lange and Topel (2006). Over the past four decades researchers have investigated a number of different methods to distinguish between these two effects. These include studies of returns to a year of schooling that leads to a qualification relative to one that does not; studies of relative returns to education in screened and unscreened
sectors; twin studies; employer learning studies etc. For a good recent review see Brown and Sessions (2004). Most scholars find that the effect of signalling on the wage premium is very modest. As our baseline level we use the 10% estimate reported by Lange and Topel (2006) based on their model with employer learning.

3. A “micro-to-macro” approach

We noted in the introduction that system-wide analyses of national impacts of HEIs have tended to use growth models to identify total (social plus private returns), and then use microeconomic estimates of private returns to deduce the scale of social returns, which is in effect a “macro-less-micro” approach. Here we explain and motivate our proposed “micro-to-macro” approach; discuss how we project the changes in the proportion of graduates in the Scottish labour force; identify the micro-econometric evidence of the private market returns to higher education and the evidence on signalling that we employ here; apply the implied productivity differential to our labour force projections to yield the overall stimulus to labour efficiency; outline the HEI-disaggregated CGE model of Scotland, which we then employ, in Section 4, to simulate the system-wide impacts of a growing proportion of graduates in the Scottish labour force.

3.1 The motivation for our approach

We propose to explore the system-wide or macroeconomic impacts of HEIs by adopting a “micro-to-macro” approach. The essence of this approach is to use the evidence on micro-econometric impacts of HEIs to inform both the specification of a regional, HEI-disaggregated CGE model and the nature of the shocks that HEIs transmit to their host regional economies. The idea is to exploit the often sophisticated and extensive micro-econometric evidence on the effects of HEIs to infer their likely macroeconomic impacts.

Our micro-to-macro approach has a number of strengths. Firstly, we can, in principle, isolate the system-wide ramifications of any particular demand or supply side impact associated with HEI activity. Presently our concern is with the system-wide impact of the
productivity stimulus associated with graduates, but other impacts can also be accommodated provided relevant empirical evidence exists. In contrast, the macroeconomic (growth) approach can at best identify the aggregate impact of HE activity, including the effects of any externalities. Secondly, in a broader context, the micro-to-macro approach can be used to measure the system-wide impacts of the social and the non-market private benefits of higher education, such as those that arise through enhanced health (but are not reflected in earnings). McMahon (2009, chpt. 4) reviews this literature and suggests that these wider impacts of HEIs may be substantial.

Thirdly, the transmission mechanism from any particular supply side or demand side stimulus to the wider economy from HEIs can, in principle, be captured by the model, at least in broad-brush terms, and the causal sequence is clear in any subsequent simulation of impacts. Fourthly, the micro-to-macro approach is intuitive, coherent and transparent, since it is not bedevilled to the same extent by the unavoidable problems of the interpretation of aggregate growth models based on varying theories, methods, assumptions and databases. Fifthly, while, in principle, the macro growth approach can be implemented at the regional level that is our present focus, in practice, this is often not straightforward given the quality and availability of regional data generally, constraints which have limited the application of economic growth models in a UK regional context. However, the modelling framework that makes the micro-to-macro approach feasible can readily be implemented for regions provided an appropriate input-output table exists. Overall, we believe that the micro-to-macro approach provides a useful additional means of exploring both demand and supply-side regional impacts of HEIs in a system-wide context.

We formulate our scenarios in terms of the increases in labour productivity attributable to the growing proportion of graduates in the labour force. In order to calculate the size of the shock we have to project two things: the future size and skill composition of the Scottish labour force, and the future productivity difference between graduates and non-graduates attributable to the effect of higher education. We consider each in turn.
3.2 Future labour force.

Our baseline scenario for the future skill composition of the Scottish labour force reflects a direct projection of the future number of graduates. The central assumption is that the number of graduates from the Scottish universities after the 2006/07 academic year changes proportionately to the number of people aged 20-25 and that the retention rate of graduates within the regional labour force remains constant. This is a convenient combination of assumptions that captures the step change from the 1980s, but abstracts from the possible endogeneity of HEI participation and retention rates. The original skill composition is calculated from the NOMIS age-specific shares of graduates and the 2006 population structure. The new graduates are distributed within the 20-35 age group proportionately to the 2006 distribution of graduates\(^7\). The distribution is limited to the 20-35 age group because in 2006 it contained about 88% of all HEI graduates and each older cohort accounted for less than 1% of graduates. The proportion of graduates in older cohorts (36+) is assumed to stay constant (only people get older). The future size of the potential labour force (population age 20-64) and its groups is taken from the 2008-based ONS principal population projection.

Retention rates are calculated based on the Higher Education Statistics Agency’s (HESA) Destination of Leavers from Higher Education Survey (DLHE) data set for 2002-07. For the baseline scenario we use the “UK net retention rate”. This is calculated as the total number of UK graduates employed in Scotland 6 months after graduation divided by the total number of UK graduates that graduated from Scottish universities. The retention rate therefore takes into account the retention of students from Scottish universities as well as the net inflow of graduates from other UK regions. The UK net retention rate for Scotland in 2006/07 was 88%. It is very stable over the 5 years for which we have data, and

\(^7\) Calculated from the Higher Education Statistics Agency (HESA) data
fluctuates within one percentage point. Based on these calculations, about 30 thousand new graduates entered the Scottish labour force in 2006.

Figure 2 plots the projected future aggregate share of graduates in the Scottish labour force implied by our assumptions. By 2051 the share of graduates in the labour force will stabilize at 46% (starting from just above 34% at the beginning of the period). The change is remarkable given that we are not projecting an increase in the number of graduates from Scottish universities. Rather, it is the interaction of demography with an assumed constant number of graduates that generates these results. Older cohorts have a significantly lower proportion of graduates in them than more recent ones. Accordingly, through time “less skilled” older cohorts are replaced by “more skilled” younger cohorts, and the total share of graduates in the labour force increases.

---

8 Foreign graduates are underrepresented in the DLHE and we make a correction for this. In the 2006-07 DLHE of all Scottish graduates only 4.4% were of non-UK origin. While according to the general HESA database, which has comprehensive coverage, foreign students accounted for 16.6% of the total student population in that academic year. We excluded from the retention rates calculation those foreign students who are not covered by the DLHE survey. In 2007/08 academic year they constituted 12.2% (16.6\%-4.4\%) of the total number of graduates with a first degree. The implication is that 73% of foreign students (100\*12.2/16.6) leave. This is a large and growing share (in 2002 it was 10.4\%) and is potentially problematic. However, at the moment we know nothing about this group and to treat them as UK graduates would clearly be inappropriate because they are less likely to stay than domestic graduates. However, this group was the target of the Scottish Government’s Fresh Talent Initiative, which sought to encourage them to remain (Lisenkova et al, 2009)
3.3 Graduate wage premium and productivity

To illustrate the varied findings relating to the graduate wage premium, we take the 30% to 60% range for the wage premium identified by Goldin and Katz (2007), and adopt the 10% signalling estimate of Lange and Topel (2006). This combination encompasses the range of national and regional estimates of graduate wage premia for the UK.\(^9\)\(^,\)\(^10\)

---

\(^9\) Probably the most influential study is that by Blundell \textit{et al} (2005), who estimate a graduate wage premium of 26% relative to those who leave school with A-levels. Since this is an extraordinarily thorough study that exploits an unusually detailed database, our choice of the range of wage premia to explore may seem optimistic. However, our “minimum” estimate, of 30%, once adjusted for signalling gives a very similar estimate of the productivity stimulus. Since the econometric analysis presented in Blundell \textit{et al} (2005) is able to control for ability, it could be argued to be less susceptible to the signalling critique than other micro-econometric studies, which typically have fewer control variables available in their database. However, the Blundell \textit{et al} (2005) study applies exclusively to male graduates, whereas we are concerned here with the impact of all graduates. All of the evidence suggests that graduate wage premia are much higher for females than for males, and we have settled on a mean wage premium of 45%.
Naturally, the sensitivity analysis provides a feel for how the order of magnitude of the results would be affected by choosing different estimates of the wage premium, and it is comparatively straightforward to infer the impacts that are likely to be associated with smaller values of the wage premium.

Our projection of the proportion of graduates in the labour force is combined with our assumptions about the future graduate wage premium and the strength of the signalling effect to calculate a series of productivity-adjusted labour force estimates. The total productivity-adjusted labour force is calculated as the sum of non-graduates and graduates weighted by their productivity difference (measured by the graduate wage premium reduced by the effect of signalling)\(^\text{11}\). Because we are not adjusting the potential labour force for age-specific labour force participation and unemployment rates our implicit assumption here is that these remain constant.

The size of the long-run labour productivity shock is calculated as a growth rate in the productivity-adjusted labour force between 2006 and 2051. To eliminate the scale effect of the change in total population, the series is divided by the change in the size of the labour force\(^\text{12}\). This allows us to focus exclusively on the effect of the changing skill composition on the productivity of the labour force.

3.4 Simulation strategy

The purpose of our simulations is to determine the likely system-wide consequences of the improvement in productivity implied by our projections of the increasing share of graduates in the labour force. The stimulus is introduced as an increase in the productivity

---

10 A member of our HEI research team, Robert Wright, estimated the Scottish-specific wage premium to be 58%, relative to all non-graduates.

11 Productivity-adjusted labour force = non-graduates + graduates * \{(1 + (graduate wage premium x (1-signalling effect)))\} So, for example, a 30% wage premium in the presence of a 10% signalling effect implies graduates are 27% more productive.

12 The future productivity-adjusted labour force can change for two reasons: change in the size of the labour force and change in the skill composition of the labour force. Because we are only interested in the effect of the latter the gross change in the productivity-adjusted labour force is discounted by the change in the total labour force, thus, leaving only the net effect of the change in the skill composition.
of labour across all 25 sectors of the model; it takes the form of labour-augmenting, or Harrod-neutral, technical progress. In a partial equilibrium context, the determinants of the employment effect of such a change has been understood since Hicks’ (1932) identification of laws of derived demand. The present general equilibrium context complicates matters in that the key wage-elasticity of the demand for labour reflects a responsiveness to all of the effects of wage changes, including income and compositional effects. An increase in labour efficiency reduces the effective price of an efficiency unit of labour, and so stimulates the demand for labour in efficiency units. Employment rises, falls or remains the same depending on whether the general equilibrium wage elasticity of labour demand is greater, less or equal to unity. This, in turn, depends on all the key elasticities in the model, including of course, the elasticity of substitution between labour and capital in each sector; the sectoral shares of labour in value-added and the elasticity of supply of capital. In our model capital accumulation takes time and so the value of the latter increases through time, as does the wage elasticity of labour demand.\textsuperscript{13} However, if households and firms are forward-looking, they anticipate expansion, bring forward their investment and consumption plans and so increase the short-run wage elasticity of employment demand.

In all of the simulations presented below the migration function is switched off. This means that there is no inflow or outflow of labour generated by the change in the returns on labour. Because our goal is to isolate the impact of the increased productivity of the labour force due to the increasing proportion of graduates within it, we preclude endogenous population adjustment. If the size of the labour force is allowed to adjust through migration the change in employment and GDP for a given increase in the labour productivity is larger.

The previous two sections explain how we generate a series of projected changes in the productivity of the Scottish labour force in response to the increasing proportion of

\textsuperscript{13} See e.g. Hanley et al (2009) and Turner (2009) for detailed discussion of the determinants of an efficiency change in production (in their case in the use of energy).
graduates in the labour force. The size of the long-run labour productivity shock, for the baseline scenario with an assumed 45% graduate wage premium, of which 10% reflects a signalling effect, is 4.1%. This is the long-run labour productivity impact of the proportion of graduates increasing from just above 34% to 46%. Of course, the stimulus to productivity in the early years of the simulation is very modest, since it takes time for the proportion of graduates, and therefore productivity, to increase significantly.

3.5 The HEI-disaggregated CGE Modelling Framework

To simulate the system-wide impact of increases in labour productivity on the Scottish economy we employ a computable general equilibrium (CGE) model, AMOS, which is explicitly disaggregated to accommodate a separate HEI sector. AMOS is a CGE modelling framework parameterised on data from Scotland. Essentially, it is a fully specified, empirical implementation of a regional, inter-temporal, general equilibrium variant of the Layard, Nickell and Jackman (1991, 2005) model. It has three domestic transactor groups, namely the personal sector, corporations and government; and four major components of final demand: consumption, investment, government expenditure and exports. The model has 25 sectors, of which the Scottish HEI sector is one. A good general description of CGE modelling is given in Greenaway et al (1994) and an extensive review of regional CGE models can be found in Partridge and Rickman (1998, 2010).

In this version of the model, consumption and investment decisions reflect intertemporal optimization with perfect foresight (Lecca et al, 2010a,b). However, for comparative purposes we also report the results of the myopic version of the model, which has a recursive dynamic structure, since this yields some interesting differences in terms of the short-run employment responses to productivity enhancements. Real government

---

14 AMOS is an acronym for A Macro-micro Model Of Scotland. The model is calibrated using a Social Accounting Matrix based around the 2004 Scottish Input-Output Tables, rolled forward to 2006. (Scottish Government 2007). See e.g. Hermannsson et al (2010c) for details.
expenditure is exogenous. The demand for Scottish Rest of the UK (RUK) and Rest of the World (ROW) exports is determined via conventional export demand functions where the price elasticity of demand is set at 2.0. Imports are obtained through an Armington link (Armington, 1969) and therefore relative-price sensitive with trade substitution elasticities of 2.0 (Gibson, 1990). We do not explicitly model financial flows, our assumption being that Scotland is a price-taker in financial markets.

It is assumed that production takes place in perfectly competitive industries using multi-level production functions. This means that in every time period all commodity markets are in equilibrium, with price equal to the marginal cost of production. Value-added is produced using capital and labour via standard production function formulations so that, in general, factor substitution occurs in response to changes in relative factor-prices. Constant elasticity of substitution (CES) technology is adopted here with elasticities of substitution of 0.3 (Harris, 1989). In each industry intermediate purchases are modelled as the demand for a composite commodity with fixed (Leontief) coefficients. These are substitutable for imported commodities via an Armington link, which is sensitive to relative prices. The composite input then combines with value-added (capital and labour) in the production of each sector's gross output. Cost minimisation drives the industry cost functions and the factor demand functions.

In the simulations reported in this paper, the labour market is characterised by a regional bargaining function, in which the bargained real wage is inversely related to the unemployment rate. The bargaining function is parameterised using the regional econometric work reported in Layard, Nickell and Jackman (1991, 2005). Detailed discussion of the model and underlying algebraic structure are available in Harrigan et al (1991) for the myopic variant and in Lecca et al (2010a,b) for the inter-temporal version of AMOS. The model is calibrated to a purpose-built, HEI-disaggregated IO table and Social Accounting Matrix (SAM) for 2006. The process of constructing the HEI-disaggregated IO table is described in Hermannsson et al (2010c).

It is important to recognise that, in the simulations reported below, the only exogenous change that is introduced into the model is the increased labour productivity due to the
growing share of graduates in the labour force. The results should therefore be interpreted as deviations from what would have occurred if labour force productivity had remained unchanged. For simplicity, we assume that we start from a steady-state equilibrium, although we have seen that there have been significant changes in the percentage of graduates in the labour force in recent years.

4. Results

4.1 Base case

As explained in earlier sections our base case assumes a constant graduate wage premium of 45%, a constant net graduate retention rate of 88% and a constant signalling effect equivalent to 10% of the graduate wage premium. When combined with our projections of the proportion of graduates in the labour force, these assumptions imply a long-run stimulus to labour productivity of 4.1%. This is the labour productivity increase that is implied by the gradual rise in the proportion of graduates in the labour force to 46%. Of course, this effect builds up through time, reflecting the gradual build-up in the proportion of graduates in the labour force depicted in Figure 2. When we simulate the impact of this using our HEI-disaggregated CGE model of the Scottish economy, we obtain the long-run results reported in Table 1. In the present context the long-run refers to a position where all capital stocks have fully adjusted, and all current cohorts have been replaced, so that the proportion of graduates in the Scottish labour force is 46%.

As we would expect for a beneficial supply side disturbance of this type there is a stimulus to gross regional product, and a downward pressure on prices. Furthermore, the stimulus is substantial, with an increase of 4.2% in GRP. Recall that this result is based on an assumption of unchanged HE policy: the total number of graduates is constant in this simulation. A key transmission mechanism is from improved regional competitiveness, through a stimulus to trade, with exports to RUK and ROW increasing by 4.3% and economic activity generally being stimulated. Importantly, we are assuming no changes in the economy of the rest-of-the UK.
Table 1. Long-run impacts of a 4.1% increase in labour productivity (% changes from base)

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>GRP</td>
<td>4.2</td>
</tr>
<tr>
<td>Consumption</td>
<td>1.0</td>
</tr>
<tr>
<td>Investment</td>
<td>3.7</td>
</tr>
<tr>
<td>Employment</td>
<td>0.4</td>
</tr>
<tr>
<td>Unemployment rate</td>
<td>-5.9</td>
</tr>
<tr>
<td>Nominal wage</td>
<td>-0.6</td>
</tr>
<tr>
<td>Real wage</td>
<td>0.7</td>
</tr>
<tr>
<td>CPI</td>
<td>-1.2</td>
</tr>
<tr>
<td>Exports to RUK</td>
<td>4.3</td>
</tr>
<tr>
<td>Exports to ROW</td>
<td>4.3</td>
</tr>
<tr>
<td>Capital Stock</td>
<td>3.7</td>
</tr>
</tbody>
</table>

Notice that in this simulation employment actually increases in the long run: ultimately the stimulus to employment from improved competitiveness, for example, dominates the fact that any given level of output can now be produced with less labour input. Of course, the fall in the price of an efficiency unit of labour stimulates the demand for labour in efficiency units, but in general employment can fall (and does in the short-run if transactors are myopic – see below).

In the long-run the increase in the employment of efficiency units of labour (which exceeds the change in actual employment by the size of the labour productivity shock) is greater than the change in value-added, which in turn is greater than the change in the capital stock. The increase in GDP exceeds the labour productivity increase because both employment and capital stock are increasing.

The reduction in the wage per efficiency unit of labour stimulates the demand for value-added through its impact on prices, via a competitiveness and real income effect, and this in turn stimulates the demand for both labour and capital services. However, the reduction in the relative price of an efficiency unit of labour stimulates the demand for it relative to capital, through a substitution effect, and the ratio of efficiency units of labour to capital increases. Nonetheless, the change in employment is less than that in capital. The
capital/worker ratio increases, reflecting the grater efficiency of workers.

The increase in the demand for labour and capital pushes up the real wage and the real rental rate. However, the overall level of domestic prices is falling because of the competitiveness effect, and the nominal wage and rental rates decline too. While the real wage rises, it does so by less (0.7%) than the stimulus to productivity (4.1%), so that the wage in efficiency units falls, so that the unskilled do get squeezed as a consequence.

Notice that the competitiveness effect is conditional on our assumption that labour efficiency is improving in Scotland relative to the rest of the UK (RUK) and the rest of the World (ROW). If other regions are experiencing similar increases in productivity, the competitiveness advantages would, of course be muted (but would be offsetting what would otherwise be a decline in Scottish competitiveness).

It is instructive to examine the time path of the simulated response of the Scottish economy to the projected increase in the proportion of graduates in the labour force. Figure 3 plots the GRP response to this increase. The middle 2 lines of the graph relate to the base case in which, as we have seen, GRP ultimately rises by 4.2%. In both the myopic and forward-looking cases, GRP approaches its long-run equilibrium level gradually, reflecting the projected build-up in the proportion of graduates in the labour force. In the forward-looking base case in which the wage premium is 45% (depicted by the dashed line WP 45% (FL)), however, adjustment is, as we would expect, more rapid than in the myopic case (WP45% (MYP)) as consumers and investors correctly anticipate the expansion and bring forward expenditures. The long-run equilibrium impact is, however, identical in each case (Lecca et al (2010a)).

The other cases depicted in Figure 4 differ from the base case only in respect of the wage premium that they assume (and which continues to be treated as invariant to the proportion of graduates in the working population). For a wage premium of 30% GRP eventually increases by 2.9%, and with a premium of 60%, the long run impact on GRP is 5.4%. As we would expect the long-run stimulus to GRP is directly related to the size of
the wage premium.

Figure 3. The impact of the increasing graduate composition of the labour force on Scottish GRP

The adjustment paths for employment are shown in Figure 4. The base case is shown as the two lines that meet in the middle of the right-hand-side of Figure 4 (at a 0.38% increase in the long-run equilibrium employment level). In the myopic case there is an extended period during which employment actually falls, reflecting the various factors that make the general equilibrium wage elasticity of employment demand lower in the short-run, including the fixed sectoral capital stocks in the first period. In the myopic case investment responds partially to rental rate changes and very gradually impacts on the capital stock, and consumption is income-constrained. In the forward-looking case investors anticipate yet higher profitability in the future and consumers anticipate higher wealth, leading both to bring spending forward relative to the outcome under myopia. In
effect, the short-run general equilibrium elasticity of employment demand with respect to the real wage is raised by the presence of forward-looking transactors.

**Figure 4. The impact of the increasing graduate composition of the labour force on Scottish employment**

In practice, neither the purely myopic, nor the perfect foresight case is likely to be realistic, but the two paths give an indication of the likely range of possible outcomes.

**4.2 Sensitivity analysis**

While we motivated our base case scenario on what we believe are the most plausible assumptions given the available micro-econometric evidence, clearly there is considerable uncertainty concerning our assumptions about various issues. In this section we conduct a sensitivity analysis around the most important factors influencing the results. Firstly, we explore the impact of alternative assumptions about the future size of the graduate wage premium and the strength of the signalling effect. Secondly, we vary our assumptions about graduate retention rates. Thirdly, we analyse the consequences of different
participation rates in higher education.

The size of the wage premium and the signalling effect

We consider three potential levels of the long-run graduate wage premium: 30%, 45% and 60%. They lie within the boundaries of the college wage premium observed in the US over the past century (Goldin and Katz, 2007), though recent UK estimates, as we have noted, would be towards the middle and lower end of this spectrum. We also consider the impact of three different signalling effects: 0%, 10% and 30%.

Table 2 presents the size of the long-run productivity stimulus implied by each of the nine possible combinations of graduate wage premium and signalling effects. Naturally, the size of the stimulus is directly related to the graduate wage premium, but inversely related to the strength of the signalling effect. The base case scenario corresponds to the combination of the 45% wage premium and 10% signalling, which as we have seen implies a long-run stimulus to productivity of 4.1%. The size of the labour productivity stimulus varies significantly depending on the combination of these two factors, implying long-run productivity stimuli between 2.2% and 5.7%.

Table 2. The size of the labour productivity shocks for simulations

<table>
<thead>
<tr>
<th>Graduate wage premium</th>
<th>Signalling</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0%</td>
</tr>
<tr>
<td>30%</td>
<td>3.1%</td>
</tr>
<tr>
<td>45%</td>
<td>4.5%</td>
</tr>
<tr>
<td>60%</td>
<td>5.7%</td>
</tr>
</tbody>
</table>

Summary results for these simulations are presented in Table 3. The long-run increase in GRP attributable to the changing skill composition of the labour force varies between 2.3% (for the case of the 30% wage premium and 30% signalling effect) and 5.9% (for a
60% premium with no signalling effect).

Table 3. The long-run increase in GRP in response to the corresponding productivity stimulus.

<table>
<thead>
<tr>
<th>Graduate wage premium</th>
<th>Signalling</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0%</td>
</tr>
<tr>
<td>30%</td>
<td>3.2%</td>
</tr>
<tr>
<td>45%</td>
<td>4.6%</td>
</tr>
<tr>
<td>60%</td>
<td>5.9%</td>
</tr>
</tbody>
</table>

The adjustment paths for GRP for the three scenarios with a 10% signalling effect can be seen in Figure 3, and the corresponding paths for employment are plotted in Figure 4. The adjustment paths are similar in all cases although, of course, the long-run equilibrium impacts differ as we would expect given the different scales of the productivity stimulus. The main differences are the more rapid adjustments apparent under perfect foresight, and the more positive short-run employment experience in that case.

Retention rates

The base line scenario assumes a UK net retention rate that, in addition to Scottish graduates, includes the net flow of graduates from other UK regions. This essentially means that our simulations are providing a measure of the impact of UK HEIs on the Scottish economy. Here we explore the impact of HEIs using the Scottish gross retention rate that only takes into account the retention of graduates from the Scottish HEIs that were working in Scotland 6 months after graduation. (So it excludes the net inflow of graduates from RUK that is included in the simulations reported above.) This can be regarded as a measure of the impact of Scottish HEIs on the Scottish economy. The size of the labour productivity shock and corresponding GDP increase for both types of retention rates are presented in Table 4.
Table 4. The long-run GDP increase for alternative retention rate assumptions

<table>
<thead>
<tr>
<th>Graduate wage premia</th>
<th>Retention rates</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>UK net retention rate</td>
<td>Scottish local retention rate</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Shock</td>
<td>GDP increase</td>
<td>Shock</td>
</tr>
<tr>
<td>30%</td>
<td>2.8%</td>
<td>2.9%</td>
<td>2.6%</td>
</tr>
<tr>
<td>45%</td>
<td>4.1%</td>
<td>4.2%</td>
<td>3.8%</td>
</tr>
<tr>
<td>60%</td>
<td>5.2%</td>
<td>5.4%</td>
<td>4.8%</td>
</tr>
</tbody>
</table>

Focussing on the Scottish gross retention rate implies a slightly lower stimulus to productivity of between 0.2 and 0.4 percentage points. This, of course, implies that the stimulus to GDP is lower, by between 0.2 and 0.4 percentage points. The differences in these GDP estimates provides a measure of the contribution of HEIs in the rest of the UK to the Scottish economy. A symmetrical study of the impact of Scottish HEIs’ graduates on other regions would require us to model the other regions of the UK explicitly.

Future skill mix of the labour force

We now explore the impact of alternative approaches to projecting the future skill composition of the labour force. The central assumption in this approach is that all future cohorts will reach the same share of graduates as the highest age-specific share attained in recent years. The age-specific shares of graduates in 2006 were obtained from NOMIS. For future years it is assumed that cohorts that were 25 or older in 2006 have already achieved the highest level of qualification by this year and in the future their skill composition will not change. We used this cut-off point because, in 2006, people aged 25 achieved the highest proportion of graduates, namely 46%. For cohorts that were younger than 25 in 2006 and for new cohorts that enter the labour force in the future, it is assumed that all of them will achieve the 46% share of graduates by the age of 25. For those aged 20-24 it is assumed that they will have the same age-specific shares of graduates as
cohorts that were in this age group in 2006. Thus, by 2046 all age groups have 46% of
graduates, except for those aged 20-24, who are assumed to be still in the process of
acquiring their qualification. The fragment of the projected skill composition is provided
in the Appendix for illustration.

The projected future skill mix is multiplied by the projected Scottish potential labour
force – population aged 20-64 – to arrive at the total future number of graduates. The rest
are assumed to be non-graduates. The total productivity-adjusted labour force is
calculated in the same way as in the baseline scenario.

As one alternative to this scenario we calculated the effect of increase in the maximum
age-specific graduate share from the current 46% to 50%. This level was chosen because
it had recently been a Scottish Government target for HEI participation. In our scenario
participation increases by 1 percentage point a year starting from 2011 and reaches 50%
by 2014. Table 5 presents the size of the shocks and the corresponding increases in GDP
associated with them (on the baseline assumption of 10% signalling).

Table 5. The impact of alternative participation rate assumptions

<table>
<thead>
<tr>
<th>Graduate wage premia</th>
<th>Retention rates</th>
<th>Current participation rate</th>
<th>50% participation rate</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Shock</td>
<td>GDP increase</td>
<td>Shock</td>
</tr>
<tr>
<td>30%</td>
<td>2.5%</td>
<td>2.6%</td>
<td>3.4%</td>
</tr>
<tr>
<td>45%</td>
<td>3.6%</td>
<td>3.7%</td>
<td>4.8%</td>
</tr>
<tr>
<td>60%</td>
<td>4.6%</td>
<td>4.8%</td>
<td>6.2%</td>
</tr>
</tbody>
</table>

These results cover a wide range of possible outcomes, reflecting a number of “what if”
simulations. The goal of our exercise is to provide broad-brush quantitative estimates of
the effect of graduates on GRP, since the available evidence does not permit precise
estimation of these effects. However, our simulations suggest that, across a wide range of
possible assumption, there is a substantial GDP impact: ranging from 2.5% at one end of the spectrum and up 6.4% at the other.

It would, of course, be useful to conduct further sensitivity analysis, for example, with respect to some of the key parameters and macroeconomic closures of the computable general equilibrium model, but we shall report this in future research given the length of the present paper.

5. Conclusions

In this paper we argue that there exists a major lacuna in the existing literature on the regional impacts of HEIs: the absence of any systematic attempt to assess the scale of their impact on regional economies that they exert through the enhanced productivity of their graduates. Of course, this mechanism is widely recognised, and its potential importance often emphasised, but there have been no systematic attempts to measure the scale of the impact at least in a UK context. This paper attempts to address this gap by developing a “micro-to-macro” approach that uses existing micro-econometric evidence on the scale of the graduate wage premium and the strength of any signalling effect to identify the differential productivity stimulus of graduates relative to non-graduates. We then project the share of graduates in the labour force, compute the implied productivity stimulus and simulate the system-wide impact of this using an HEI-disaggregated CGE model of Scotland.

In projecting the share of graduates we assume, in our base case, an unchanged total number of graduates and retention rates, but demographic change implies that through time there is a major increase in the proportion of graduates in the Scottish labour force, and therefore in productivity. As old cohorts with a lower share of graduates leave the labour force they are replaced by the young cohorts with higher levels of HEI participation, so that the aggregate share of graduates in the labour force increases.

The impact of the implied stimulus to labour productivity is then simulated within an
HEI-disaggregated CGE model of Scotland that is calibrated to a purpose-built database in which HEIs are treated as a separate sector of the economy. While the precise results inevitably depend upon assumptions, overall they strongly suggest that HEIs exert a significant impact on regional economies through the skills with which they imbue their graduates. These effects typically imply significantly larger impacts than the demand-side or expenditure effects of HEIs (Hermannsson et al., 2010a,b), when considered on as comparable a basis as possible. In those studies the focus is on HEIs as a sector that demands intermediate goods from other Scottish firms, and whose employees consume Scottish goods out of their incomes. The highest impact of combined HEI and student expenditures is 2.63% of GDP (under conventional input-output assumptions, with the regional public sector budget constraint ignored), which is significantly below the estimate of our base case in this paper (4.2%), although this is predicated upon a wage premium of 45% that is constant in the face of the increased proportion of graduates. However the relative scale of supply side effects is much more impressive once it is recognised that the estimated expenditure impact reflects the maximum possible impact of HEIs’ - and their students’ - expenditures, which typically would require inclusion of migration effects not included here. Furthermore, the expenditure analysis does not measure a marginal impact, but rather relates to a “hypothetical extraction” of the entire Scottish HEIs sector. Crucially, and in stark contrast, the supply-side impacts of graduates reported here reflect an assumption of a constant number of graduates interacting with ageing: they reflect the incremental effects that would arise with no change in HE policy.

Of course, the precise numerical results are dependent on our assumptions, and through our sensitivity results we have tried to give a flavour of the impact of relaxing these. However, what is clear is that, even if our most conservative assumptions are close to reality, the impacts of graduates on the supply-side of regional economies are substantial. Naturally, the research reported here can be extended in a number of directions, since we have adopted the simplest set of assumptions that the available evidence seems to allow. Firstly, we could further relax the assumed constancy of the wage premium and the graduate retention rate, and explore the possible endogeneity of both. Secondly, we can assess the importance of graduates for other regions and for the UK as a whole. The latter
would identify the extent to which regional effects depend on regional-specific competitiveness effects.

Thirdly, and relatedly, HEIs are an integral part of a UK system of HE, and Scotland is inextricably linked to the economy of the rest of the UK, through migration and bargaining mechanisms, it may useful to investigate the inter-regional effects of the productivity shocks. Fourthly, we could explore the system-wide impact of other supply-side transmission mechanisms, notably those coming through innovation and knowledge spillovers (e.g. Harris et al (2010a,b)). Fifthly, to the extent that the micro-econometric evidence exists we can also investigate the wider impacts of HEIs, notably social returns and non-market private returns (McMahon (2009); Hermannsson et al (2010d)). This is potentially crucial in evaluating the recommendations of the Browne (2010) report, which proposes a radical reduction in the extent of public subsidy of undergraduate education in the UK.

Sixthly, future analyses could seek to further explore the impact of the origin of graduates as well as their employment destination. This would allow, in principle, a comprehensive analysis disaggregating graduates’ impacts by location of HEIs for each region of the UK. Given devolution within the UK this analysis is likely to be of interest to both regional and national governments. Seventhly, the previous extension would be even more revealing if we allow for a degree of heterogeneity, for example, in respect of variations in productivity shocks across industrial sectors and by graduates’ subject areas, and even by HEI of graduation. Furthermore, we could extend the analysis to allow for distinct labour markets for graduates and non-graduates.

Finally, if the HE sector in the UK, or at least in England, is going to become much more market-driven following the Browne (2010) report, it would be instructive to attempt to incorporate aspects of this within the modelling system. This is especially important if the degree of divergence in HE policy across the countries of the UK increases as a consequence of devolved governments seeking to pursue differentiated policies. Approaches of the type developed in this paper hold the promise of allowing us to explore
the potentially significant impact on the regional distribution of HEI impacts.
References


the returns from education and training to the individual, the firm and the economy”, *Fiscal Studies*, vol. 20(1), pp 1-24


Freeman, S. (1996) “Equilibrium Income Inequality among Identical Agents”, *Journal of


Harris, R., Li, Q. C. and Moffat, J. (2010) The impact of higher education institution-firm
knowledge links on establishment-level productivity in British regions, Strathclyde Discussion Papers in Economics, 10-18.


Education, IZA Discussion Paper No. 1549


Appendix. Fragment of the projected skill composition of the Scottish labour force.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>20</td>
<td>11%</td>
<td>11%</td>
<td>11%</td>
<td>11%</td>
<td>11%</td>
<td>11%</td>
<td>11%</td>
<td>11%</td>
<td>11%</td>
<td>11%</td>
</tr>
<tr>
<td>21</td>
<td>25%</td>
<td>25%</td>
<td>25%</td>
<td>25%</td>
<td>25%</td>
<td>25%</td>
<td>25%</td>
<td>25%</td>
<td>25%</td>
<td>25%</td>
</tr>
<tr>
<td>22</td>
<td>34%</td>
<td>34%</td>
<td>34%</td>
<td>34%</td>
<td>34%</td>
<td>34%</td>
<td>34%</td>
<td>34%</td>
<td>34%</td>
<td>34%</td>
</tr>
<tr>
<td>23</td>
<td>37%</td>
<td>37%</td>
<td>37%</td>
<td>37%</td>
<td>37%</td>
<td>37%</td>
<td>37%</td>
<td>37%</td>
<td>37%</td>
<td>37%</td>
</tr>
<tr>
<td>24</td>
<td>45%</td>
<td>45%</td>
<td>45%</td>
<td>45%</td>
<td>45%</td>
<td>45%</td>
<td>45%</td>
<td>45%</td>
<td>45%</td>
<td>45%</td>
</tr>
<tr>
<td>25</td>
<td>46%</td>
<td>46%</td>
<td>46%</td>
<td>46%</td>
<td>46%</td>
<td>46%</td>
<td>46%</td>
<td>46%</td>
<td>46%</td>
<td>46%</td>
</tr>
<tr>
<td>26</td>
<td>46%</td>
<td>46%</td>
<td>46%</td>
<td>46%</td>
<td>46%</td>
<td>46%</td>
<td>46%</td>
<td>46%</td>
<td>46%</td>
<td>46%</td>
</tr>
<tr>
<td>27</td>
<td>44%</td>
<td>46%</td>
<td>46%</td>
<td>46%</td>
<td>46%</td>
<td>46%</td>
<td>46%</td>
<td>46%</td>
<td>46%</td>
<td>46%</td>
</tr>
<tr>
<td>28</td>
<td>45%</td>
<td>44%</td>
<td>46%</td>
<td>46%</td>
<td>46%</td>
<td>46%</td>
<td>46%</td>
<td>46%</td>
<td>46%</td>
<td>46%</td>
</tr>
<tr>
<td>29</td>
<td>45%</td>
<td>45%</td>
<td>44%</td>
<td>46%</td>
<td>46%</td>
<td>46%</td>
<td>46%</td>
<td>46%</td>
<td>46%</td>
<td>46%</td>
</tr>
<tr>
<td>30</td>
<td>44%</td>
<td>45%</td>
<td>45%</td>
<td>44%</td>
<td>46%</td>
<td>46%</td>
<td>46%</td>
<td>46%</td>
<td>46%</td>
<td>46%</td>
</tr>
<tr>
<td>31</td>
<td>45%</td>
<td>44%</td>
<td>45%</td>
<td>45%</td>
<td>44%</td>
<td>46%</td>
<td>46%</td>
<td>46%</td>
<td>46%</td>
<td>46%</td>
</tr>
<tr>
<td>32</td>
<td>45%</td>
<td>45%</td>
<td>44%</td>
<td>45%</td>
<td>45%</td>
<td>44%</td>
<td>46%</td>
<td>46%</td>
<td>46%</td>
<td>46%</td>
</tr>
<tr>
<td>33</td>
<td>43%</td>
<td>45%</td>
<td>45%</td>
<td>44%</td>
<td>45%</td>
<td>45%</td>
<td>44%</td>
<td>46%</td>
<td>46%</td>
<td>46%</td>
</tr>
<tr>
<td>34</td>
<td>40%</td>
<td>43%</td>
<td>45%</td>
<td>45%</td>
<td>44%</td>
<td>45%</td>
<td>45%</td>
<td>44%</td>
<td>46%</td>
<td>46%</td>
</tr>
<tr>
<td>35</td>
<td>39%</td>
<td>40%</td>
<td>43%</td>
<td>45%</td>
<td>45%</td>
<td>44%</td>
<td>45%</td>
<td>45%</td>
<td>44%</td>
<td>46%</td>
</tr>
<tr>
<td>36</td>
<td>38%</td>
<td>39%</td>
<td>40%</td>
<td>43%</td>
<td>45%</td>
<td>45%</td>
<td>44%</td>
<td>45%</td>
<td>45%</td>
<td>44%</td>
</tr>
<tr>
<td>37</td>
<td>38%</td>
<td>38%</td>
<td>39%</td>
<td>40%</td>
<td>43%</td>
<td>45%</td>
<td>45%</td>
<td>44%</td>
<td>45%</td>
<td>45%</td>
</tr>
<tr>
<td>38</td>
<td>38%</td>
<td>38%</td>
<td>39%</td>
<td>40%</td>
<td>43%</td>
<td>45%</td>
<td>45%</td>
<td>44%</td>
<td>45%</td>
<td>45%</td>
</tr>
</tbody>
</table>