Agglomeration and Labour Markets: The impact of transport investments on labour market outcomes[1]

November 2010

Sabine D'Costa (LSE and SERC)
Steve Gibbons (LSE and SERC)
Henry Overman (LSE and SERC)
Panu Pelkonen (University of Sussex and SERC)

Later version published in: R. Crescenzi and M. Percoco (Eds.), Geography, Institutions and Regional Economic Performance, Advances in Spatial Science, DOI 10.1007/978-3-642-33395-8_13, © Springer-Verlag Berlin Heidelberg 2013.

Abstract

One of the possible causes of poorer labour market outcomes for workers in peripheral regions is the small size of cities in these regions. Given this possibility, and the difficulty of affecting city size directly, a frequent policy response has been to invest in transport in order to increase access to markets. In this chapter we investigate how local labour markets respond to these potential transport improvements. We use data on individual workers in the UK to assess how area wages respond to better market access and examine whether this variation is due to a changing composition of the labour market or to higher wages for existing workers. Our results indicate that the increase in wages associated with reductions in transport times stems from changes in the composition of the workforce and that wage increases for local workers with unchanged characteristics are minimal.

Key words: transport investment, agglomeration, regional labour markets.

JEL codes: R42, R23.

[1] This chapter is based on work for the Northern Way on Strengthening Economic Linkages Between Leeds and Manchester: Feasibility and Implications (full report: http://www.thenorthernway.co.uk/page.asp?id=817). We are grateful to the Northern Way for funding and to a steering group for advice. We thank Lizze Diss at the Department for Transport and Dan Graham at Imperial College London for their help with the ward to ward GTC for driving. We thank Peter Wiener at Steer Davies Gleeve and John Jarvis at Yorkshire Forward for advice on constructing the train counterfactuals. This work contains statistical data from ONS which is Crown copyright and reproduced with the permission of the controller of HMSO and Queen's Printer for Scotland. The use of the ONS statistical data in this work does not imply the endorsement of the ONS in relation to the interpretation or analysis of the statistical data. This work uses research datasets which may not exactly reproduce National Statistics aggregates. Copyright of the statistical results may not be assigned, and publishers of this data must have or obtain a licence from HMSO. The ONS data in these results are covered by the terms of the standard HMSO "click-use" licence.

1 Introduction

There is increasing interest in the role of cities in driving recovery from recession and economic growth more generally. In the UK this partly reflects the fact that, after a long period of relative decline a number of cities have, until recently, experienced improved economic performance (ODPM, 2006). At the same time, evidence suggests there may be continued growth in these cities, particularly if the UK economy continues to move from manufacturing to services. Producers of services benefit in a variety of ways when located in cities and research suggests these benefits may be larger than for manufacturers (Rosenthal and Strange, 2004). If services do benefit more from cities, a continued shift towards services points towards a future in which economic activity could be concentrated in a (small) number of larger cities.

Amongst UK policy makers this raises a number of questions. Will this growth be concentrated mostly in the South East? If so, is there anything that policy can, or should, do to counteract this? What role might growth in Northern cities play in increasing growth in the wider Northern economy? Which cities in the North might drive this growth and what, if anything, is the appropriate role for policy? This chapter is concerned with the last of these questions.

In this chapter, we consider the role of transport as a way of increasing the size of the local economy. A larger local economy may help firms be more productive. Such agglomeration economies – the beneficial effects of a larger local economy – may arise for a variety of reasons (Duranton and Puga, 2004). A large local economy may facilitate sharing of resources (for example of infrastructure such as airports), matching of capacity (for example of workers to firms) or learning (for example transfer of knowledge between firms). Can we say anything about the likely impact of these effects if we improve transport between locations? In this chapter we use labour market data for individual workers to provide a partial answer to this question.

Our research is related to several strands of work that attempt to answer this question. One strand, following Auscher (1989) treats transport infrastructure as public capital and considers the effects using country or regional level aggregate data. Studies that build on this approach are, at best, inconclusive about the impact of transport on productivity (Gramlich, 1994; Boarnet 1997). More recently, attention has shifted from the impact of infrastructure as capital to capturing the effects of the transport network. Rice, Patacchini and Venables (2006) provide a nice example using data for UK (NUTS 3) regions. More recently, a second strand of research considers the relationship between transport and economic performance by looking at the link between agglomeration (or accessibility) and productivity using micro-data for firms or workers (Holl 2010, Melo and Graham 2010). Melo and Graham (2010) is the closest research to that described here. They estimate the effects of agglomeration economies on wages of workers in the UK finding that a 1% increase in market potential leads to a 0.1% increase in wages, an effect halved when taking workers and firms' unobserved heterogeneity into account. These findings are in line with those reported in this chapter although we use more detailed data on actual transport costs in our analysis.

There has been considerable speculation that the small size of cities in peripheral regions may have negative effects on labour market outcomes and that this may help explain their relative under performance. Given that options to directly increase the city size are limited an alternative is to invest in transport to improve access to other markets. This raises two important questions. First, does improving accessibility *cause* changes in labour market outcomes? Second, if so, what kind of changes occur? There is a small literature that addresses the first question (Ahlfeldt and Feddersen 2010). In this chapter we focus on the second question and ask – if observed correlations between accessibility and labour market outcomes capture causal mechanisms, what does this imply about how local labour markets adjust to transport improvements? Following the long running debate on whether policy should focus on people or place (Kain and Perksy, 1969) we are particularly interested in whether changes in wages are likely to benefit existing residents. To examine these issues we use data on individual workers to see how wages vary with local labour market size. We assess the extent to which these differences arise from changing composition (e.g. large cities have more educated workers) as opposed to higher wages for existing workers. We then use our estimates, coupled with realistic assumptions about policy induced changes in transport costs, to assess the effect of increased integration on labour market outcomes. Specifically, we consider the effects of improving transport links between two of the North's largest cities: Manchester and Leeds. These cities are of interest because, while both have recently experienced strong growth, research finds little evidence of interaction in terms of business connections or commuting (IPEG/CUPS, 2008; Lucci & Hildreth, 2008) despite the fact that the cities are closely located. More generally, the case study provides important lessons on the magnitudes of possible labour market effects based on specific proposals for a real world transport improvement.

2 Agglomeration and labour markets

We are interested in how labour market size affects labour market outcomes, particularly wages. Our starting point is the observation that larger places tend to have higher productivity and wages. Economists refer to the productivity effects associated with increased levels of economic activity as *agglomeration economies*. This chapter focuses on agglomeration economies that arise in production, that is because of the productivity effects of physical proximity. Higher productivity, in turn, tends to lead to higher wages. We refer to this as the effect of better access to economic mass.

The literature emphasises three sources of agglomeration economies: linkages between intermediate and final goods suppliers, labour market interactions, and knowledge spillovers. Input-output linkages occur because firms save transaction costs by locating close to their suppliers and customers. Larger labour markets may allow for a finer division of labour or provide incentives to invest in skills. Finally, knowledge spillovers arise when spatially concentrated firms or workers are more easily able to learn from one another. See Duranton and Puga (2004). In this chapter, we only consider the overall effect of access to economic mass.

As discussed in the introduction, the small size of peripheral economies may have negative implications for labour market outcomes, partly explaining their relative under performance.[2] One argument for improving transport over and above cost savings is therefore that it increases the size of the local labour market. To consider this we use individual data to see how the level and growth of wages are affected by local labour market size. This provides another way of identifying the overall agglomeration effects studied in the research referred to above.

It is increasingly recognized, however, that the *composition* of the labour market may account for a large part of the relationship between wages, productivity and local market size. For example, large cities may attract more educated workers. Because more educated workers also earn more this leads to a positive relationship between city size and wages. When we measure agglomeration economies by looking at how wages change with city size we actually capture the changing *composition* of the labour force. Alternatively, larger cities may make workers more productive whatever their education level. That is, there is a *place-based* effect whereby larger cities pay higher wages. Our research assesses the extent to which the relationship between accessibility (our measure of local labour market

[2] See, for example, the Manchester Independent Economic Review (2009) which considers this issue.

size) and wages arises from changing composition as opposed to higher wages for existing workers. We then use our estimates, coupled with realistic assumptions about changes in transport costs, to assess the effect of increased integration on labour market outcomes. This allows us to paint a richer picture of the potential gains, the distribution of effects and the structural changes that might be needed to achieve them.

2.1 Methodology and Data

To assess the magnitude of agglomeration economies we need to see how wages differ with labour market size. We then want to break these overall effects down into those from changing composition versus those from place-based effects. To do this, we need to look at wages for individuals who are otherwise identical but who live in different sized labour markets. Ideally, we would do this by randomly allocating people across places. In reality, fortunately, the UK government does not decide where people live. People are therefore able to sort across places in non-random ways. If we observed everything about an individual (age, sex, education) that affected their wage, even in the absence of random allocation, we could still identify place effects by comparing wages for people with identical observable characteristics who live in different places. Unfortunately, even with detailed data, we cannot be certain that we observe everything that might affect wages. For example, in our data, we have no information on cognitive abilities or motivation. One way to get round this problem is to follow individuals as they move across places. Providing that ability is fixed over time, if we see the same individual earning more in larger labour markets we may be more confident in attributing this to a place-based effect. Although something may have changed that both affected their earnings potential and their location, in the absence of random allocation, (or a policy change that as good as randomly assigns people) tracking individuals over time is the best we can do to identify place-based effects of changing labour market size.

To do this we need data on where individuals work, their wages and the individual characteristics that might affect wages. We would also like to be able to follow individuals over time, particularly as they move across labour markets. In the UK, such data is available from the Annual Survey of Hours and Earnings (ASHE).[3] ASHE is constructed by the Office of National Statistics (ONS) based on an annual 1% sample of employees on the PAYE register with workers sampled in multiple years. It provides information on individuals including their work postcodes (which we map to other geographical units using the

[3] Previously the New Earnings Survey.

National Statistics Postcode Directory (NSPD)), earnings including base pay, overtime pay, basic and overtime hours. We use basic hourly wage in our analysis. Individual occupation codes and data on the characteristics of an individual's job (public sector, part time, collective agreement, industry coded by SIC2003) also come from ASHE. ASHE does not provide years of education so we classify workers as belonging to one of four skill groups constructed using a mapping of SOC (Standard Occupational Classification) codes to skill groups.[4]

Information on employment and the industrial composition of areas comes from the Business Structure Database (BSD) which provides an annual snapshot of the Inter-Departmental Business Register (IDBR) accounting for approximately 99% of economic activity in the UK.[5] For the occupation structures of areas we aggregate the individual data in ASHE. Finally, area proportions of workers belonging to high-skill and intermediate-skill groups are based on LFS (Labour Force Survey) data. More detail on ASHE is provided in Gibbons, Overman and Pelkonen (2010), while information on data cleaning is provided in Northern Way (2009).

We follow existing research by focusing on the relationship between wages and 'access to economic mass' rather than city size. Imposing city boundaries is essentially arbitrary whereas measures of access to economic mass treat space as continuous by taking into account access to all other areas discounted by distance or transport cost. We construct two measures of access to economic mass.

The first is based on Generalised Transport Costs (GTC) when driving. Ward to ward generalized costs (driving) were provided by the UK Department for Transport (DfT). They comprise fuel and non-fuel vehicle operating costs and the value of time multiplied by travel time. The data have been averaged for peak and off peak. The exact formulae for these calculations can be found in the DfT's Transport Economics Note (DETR 2001) and more details on our use of GTCs are in Northern Way (2009). The access to economic mass for ward j is calculated by adding up employment in all other wards using inverse-GTC (driving) weighting. That is. $A_{it} = \sum emp_{jt} \times GTC_{ij}^{-1}$ where emp_{jt} is employment in ward j at time t, and GTC_{ij} is the ward-to-ward GTC for driving. To allow employment in ward *i* to contribute to its own access to economic mass, we set $GTC_{ii} = 0.5 \times GTC_{i1}$ where GTC_{i1} is the minimum ward-to-ward GTC for ward i (i.e. GTC for the "closest" ward). Therefore, a ward is assigned an aggregate of employment in other wards, with employment in more distant

[4] Northern Way (2009) provides more details.
[5] The 99% coverage was last verified in 2004/05, although there is no reason to think that this is not still the case.

places contributing less than employment close by. Each worker is assigned the access to economic mass value for the ward in which their employer is located.[6] Note that this index is identical to the effective density index used by Graham (2006) (although we prefer to refer to it as a measure of *access to economic mass* or *accessibility*).

Our second access to economic mass measure is calculated using train GTC. The train GTC are calculated using data provided to us by the DfT.[7] The final GTC matrix by train is a weighted sum of in-vehicle, wait and walk times (multiplied by the respective time value) and fare matrices. The index is constructed in an identical manner to the index based on driving GTC (although based on ward, not LA data). We refer to the employment accessibility measure based on driving GTC as “Car Accessibility” and that based on train GTC as “Train Accessibility”.

2.2 Results

We start by considering a model that captures agglomeration economies, ignoring the distinction between composition and place-based effects. Specifically, we run regressions that explain individual wages as a function of accessibility:

$$\ln(w_{it}) = \alpha_t + \theta \ln(A_{it}) + \varepsilon_{it}$$

where w_{it} is individual *i*'s wage at time t, A_{it} is one, or both, of the accessibility variables, ε_{it} is an error term representing unobservable factors, α_t is a time varying parameter and θ a time invariant parameter (both to be estimated). The alphas capture the increases in wages over time, while theta captures the impact of accessibility (assumed constant over our relatively short time period). Results are reported in Table 1.

We report results using only Car Accessibility (column 1), only Train Accessibility (column 2) then both together (column 3). When entered separately both are positively and statistically significantly associated with wages. When including both together we find the coefficient on Car Accessibility is positive but insignificant while that on Train Accessibility is both positive and significant. These effects of accessibility remain essentially unchanged if we drop individuals that work in London (column 4). Finally, for comparison we present results based on Travel to Work Area (TTWA) employment rather than accessibility (column 5).

[6] While ASHE contains information on both home and work postcode, NES only provides the latter so we need to base our measure of access to economic mass on work rather than home location.
[7] They stem from Base Year (2004) Rail 'Level of Service' skims based on UK Midman rail data

In terms of magnitudes, the coefficient on TTWA employment in column 5 is the easiest to interpret. It tells us that a 10% increase in TTWA employment is associated with a 0.7% increase in wages. This is consistent with the existing literature on the effect of city size on productivity which reports the effect of a 10% increase in city size varying from around 0.2% to 2% with most estimates under 1%. The coefficients on the accessibility measures are harder to interpret because they are calculated using GTC weighting of employment across all areas. Taken at face value, the coefficient of 0.344 for Train Accessibility implies that a 10% increase in employment in all Local Authorities, or a 10% reduction in the GTC between all Local Authorities, would increase wages by around 3.4%. For the moment, we focus on how these coefficients change as we introduce individual characteristics. Later, however, we calculate changes in accessibility consistent with proposed transport interventions which gives a feeling for the magnitude of the wage effects.

Table 1: Regressions of wages on accessibility

	[1]	[2]	[3]	[4]	[5]
	Car	Train	Both	Without London	TTWA Employment
ln Car Accessibility	0.230*		0.084	-0.040	
	(0.092)		(0.122)	(0.035)	
ln Train Accessibility		0.344***	0.258**	0.217***	
		(0.093)	(0.093)	(0.036)	
ln Employment					0.069***
					(0.008)
R^2	0.085	0.086	0.09	0.06	0.085
Observations	1102527	1119582	1102527	884953	1119582

Notes: Dependent variable is log hourly earnings and the explanatory variables are logarithms of car and train accessibility, or log TTWA employment. All estimations are based on panel data for 1998-2007, and include year effects. Standard errors (reported in brackets) are clustered at the TTWA level. ***, **, * denote significance at the 1%, 5% and 10% levels respectively.

As we explained above, the problem with these results is that they do not distinguish between the two different explanations of the positive correlation between accessibility and wage. To separate out these effects, we need to control for the fact that individual characteristics that affect wages may be correlated with accessibility. To do this, we include these individual characteristics in our wage regressions to give:

$$\ln(w_{it}) = \alpha_t + \beta' X_{it} + \theta \ln(A_{it}) + \varepsilon_{it}$$

where X_{it} are individual characteristics, beta is a parameter to be estimated and all other notation is as before. Beta captures the effect of individual characteristics on wages leaving theta to capture the effect of accessibility controlling for composition.

In order to separate composition from place-based effects we want to control for predetermined individual characteristics that are correlated with the accessibility of the places in which they live. These characteristics - e.g. gender - can become correlated with labour market size if individuals with different productivities sort into places of different sizes. Clearly the sex of a worker is not determined by accessibility even if males and females then choose to live in different places so that sex is correlated with accessibility. However, there are some individual characteristics that may at least partly be *determined* by accessibility. If for example good accessibility causes a person to choose a higher paid occupation (which is possible if agglomeration economies cause some occupations to be more prevalent in larger labour markets) then we may want to attribute the resulting effect on wages to accessibility *not* to occupation. Controlling for occupation in our wage regressions will yield estimates of the effect of accessibility that net-out any effects arising from occupational choice.

An additional challenge is that an association between composition and accessibility could arise because better transport connections have evolved between labour markets with more productive workers. This suggests caution because the direction of causality may not run from accessibility to labour market composition, but in the opposite direction. Improving transport would not then be effective in changing composition or raising productivity. Other than controlling for a limited number of other area characteristics we do not address this issue, so our estimates are upward biased and the effects that they imply will never be fully realised by improving transport or otherwise increasing accessibility.[8]

These issues complicate our analysis. In short, we want to control for individual characteristics that can be regarded as predetermined, not determined by accessibility in the place in which a person currently works. Unfortunately, there are some characteristics like occupation, education and industry which are partly predetermined, but may be partly determined by the place in which a person works. If we control for these factors, we control for composition effects arising both through sorting (which we want to eliminate), and

[8] Results in the academic literature suggest that the issue of reverse causality is likely to be much less important than that of composition. See Graham and Melo (2010) and Combes, Duranton and Gobillon (2011).

through changes in individual characteristics induced by accessibility (which we do not necessarily wish to eliminate).

The approach we employ is simply to estimate wage equations using various sets of individual characteristics, whilst recognising that controlling for characteristics that are partly determined by accessibility yields lower bounds to the overall effect of accessibility on wages, whereas failing to control for predetermined characteristics is likely to upward bias our estimates. We start by introducing characteristics that are most likely to be predetermined and then adding in characteristics where we are less certain. Results are reported in Table 2.[9] . Column 1 just replicates results from Table 1 where we do not control for any individual characteristics. Column 2 shows what happens when we control for sex, age and age squared which, as argued above, are certainly predetermined. The coefficient on Car Accessibility drops while that on Train Accessibility increases although neither change is statistically significant.

[9] We only report coefficients on the accessibility measures. Thee full results can be found in an appendix to Northern Way (2009)

Table 2: Regressions of wages on accessibility and other variables

	[1]	[2]	[3]	[4]	[5]	[6]	[7]
ln Car Accessibility	0.084	0.074	0.071	0.054	0.046	0.069***	0.070***
	(0.122)	(0.118)	(0.080)	(0.066)	(0.058)	(0.016)	(0.021)
ln Train Accessibility	0.258**	0.277***	0.173**	0.165***	0.170***	0.049***	0.030***
	(0.093)	(0.090)	(0.059)	(0.049)	(0.044)	(0.014)	(0.010)
R^2	0.090	0.218	0.513	0.622	0.638	0.918	0.918
Observations	1102527	1091551	1091551	1091551	1090528	1090528	1090528

Notes: Dependent variable is log hourly earnings and the explanatory variables of interest are logarithms of car and train accessibility. All estimations are based on panel data covering years 1998-2007, and include year effects. Column [1] has no controls; [2] adds age, age squared and gender; [3] adds years of education; [4] adds occupational characteristics (1-digit level) and dummies for part-time, public sector and collective wage agreement; [5] adds 1-digit industry controls; [6] adds individual fixed effects; [7] adds area level characteristics as described in the text. Standard errors (reported in brackets) are clustered at the travel-to-work area level. ***, **, * denote significance at the 1%, 5% and 10% levels respectively.

The next individual characteristic that we include is education. Although there is some evidence linking educational outcomes to accessibility, the causal effect (if indeed it is causal) is not large. Given our aggregated skills classification we would argue education as largely predetermined. However, as column 3 makes clear, sorting means that education is quite strongly correlated with accessibility, at least for Train Accessibility. This suggests that higher educated workers get paid higher wages and tend to work in areas with higher accessibility by train. Once we control for education the association between wages and Train Accessibility is considerably weakened.

Next we control for occupation, whether the individual works in the public sector, works part time and is subject to a collective pay agreement. We could think of these characteristics as associated with either the individual or the job. If the latter, it is a little harder to be certain that these characteristics are predetermined. Fortunately this issue is moot as introducing these controls has little effect on the coefficients on the two accessibility variables (column 4). A similar story applies when adding industry controls (column 5).

To summarise, when we control for composition based on the observable characteristics of individuals (and jobs) the effect of accessibility is reduced by between a quarter and a third. So far, however, we have only controlled for the observable characteristics of individuals. Given that we observe individuals over time we can use panel data techniques to control for unobservable characteristics of individuals, such as ability, that might be positively associated with both wages and accessibility. Specifically, we include individual fixed effects to control for time-invariant unobservables. This implies that the effects of accessibility are estimated from individuals that move over time (for individuals that do not move, we cannot be sure whether higher wages are something to do with that individual or with the place in which they work). The specification is:

$$\ln(w_{it}) = \alpha_t + \beta' X_{it} + \theta \ln(A_{it}) + \lambda_i + \varepsilon_{ct}$$

where everything is as defined before, except for the inclusion of individual fixed effects λ_i.

As can be seen from Column 6 the effect on the coefficients on the accessibility measures are considerable. For Train Accessibility the coefficient is decreased by a factor of 3 and now smaller than the coefficient on Car Accessibility (although not significantly so). Car Accessibility is now significant for the first time. Controlling for composition using both the observed and unobserved characteristics of individuals provide the best estimate of the relationship between wages and accessibility. We view these coefficients as the upper bound of the likely effect on individuals who do not change sex, age, education *etc.* as a result of increasing accessibility.

In the results reported so far, we only allow for place-based effects to be explained by accessibility. It is possible that other area characteristics that are correlated with both accessibility and wages might actually be the source of place-based effects. To consider this we control for a number of additional area based characteristics. Following Wheeler's (2006) work on wage growth for the US these include measures of TTWA industrial and occupational diversity to allow for the possibility that diversity might be more important for wages than size. Industrial diversity of a TTWA j is calculated using a Herfindahl index: $\sum_j (E_{ijt} / E_{it})^2$ where E is employment, *j* is two-digit industry, *i* is TTWA and t is year. Occupational diversity is an analogous measure using employment by two-digit occupation instead of SIC. We also include the shares of high and intermediate skills in TTWA working age population (with low skills the omitted category). Finally, we include two digit TTWA industry shares to see if the industrial composition makes any difference.

Column 7 shows what happens when we include these additional area characteristics. The effect of Car Accessibility is essentially unchanged, while that of Train Accessibility falls somewhat further. Detailed results in Northern Way (2009) show that TTWA share of high skills and the share of activity in Other Services are the only two significant area characteristics. These are positively correlated with Train Accessibility which reduces the coefficient on that measure of accessibility. Of course, these results may partly reflect the fact that large places attract lots of skilled workers. Without more evidence on the channels, and given that the coefficients on the access variables do not change too markedly, we prefer to use the results in Column 6 (ignoring other area characteristics) when considering the counterfactuals described below.[10]

Finally, we consider whether effects differ depending on the skill level of workers by running regressions separately for each skill group. Table 3 shows the results for our preferred specification including individual fixed effects. It is interesting to note that the effects of improving both Car and Train Accessibility may be slightly stronger for those with intermediate level skills than higher skills, while the lower skilled may not benefit at all from increased Train Accessibility. We use the average effects in what follows, ignoring the fact that the effects might differ somewhat across individuals.

Table 3: Regressions of wages on accessibility split by skill group.

	Skill group 1	Skill groups 2 and 3	Skill Group 4
ln Car Accessibility	0.054***	0.074***	0.049***
	(0.011)	(0.017)	(0.010)
ln Train Accessibility	0.003	0.054***	0.019*
	(0.015)	(0.016)	(0.009)
R^2	0.826	0.894	0.863
Observations	46057	894873	149598

Note: Regressions dividing by skill groups. FE reports coefficients from a regression including a full set of individual controls and is equivalent to specification [6] in Table 2. ***, **, * denote significance at the 1%, 5% and 10% levels respectively.

2.3 The labour market impacts of closer integration

We can now use our results to assess the labour market impact of improving accessibility. To do this we construct a counterfactual Train Accessibility measures based on a 20 minute reduction in train travel time between Manchester and Leeds.[11] As already explained, this is an investment that has been the subject of considerable interest from UK agencies concerned with narrowing the gap between the North and South of England. It also provides a natural way of translating the abstract coefficients on our accessibility measures into more concrete estimates of the effects of a real world investment.

We calculate the impact on wages by multiplying the percentage changes in accessibility by the relevant coefficient on Train Accessibility that we reported in Table 2 (repeated in the

[10] Results in Northern Way (2009) show that excluding London does not make that much difference to the Train Accessibility coefficients that are the main focus of our counterfactual analysis below.

[11] We also allow for the second round (or knock on) effects on journeys between LAs not directly affected (e.g. Liverpool to Hull) that may see improved journey times as a result of the improved network. Northern Way (2009) provides more details on the construction of counterfactuals.

last row of the table). Results are reported in Table 4. The column marked L-M -20m gives the percentage change in Train Accessibility produced by the 20 minute reduction in journey time. The first column reports the total effects of this change (including any compositional changes). These range from a 2.7% increase in wages in Wakefield to a 1.06% increase in Tameside. Column 2 shows what happens as we control for age and sex. The estimate of the percentage wage effect increases slightly because the coefficient on Train Accessibility is slightly higher. Column 3 controls for education which leads to the first big reduction in the estimated size of the effect. Columns 4 and 5 show smaller changes as we first introduce occupation and then industrial controls. Finally column 6 shows the large reduction when we allow for unobservable individual characteristics. As a reminder column 6 is our preferred estimate of the effect of increased accessibility controlling for the effects of composition. We see the results range from a high of 0.5 of a percent for Wakefield to a low of 0.2 of a percent for Tameside. As is clear, compositional changes account for the vast majority of the overall effect on wages.

We view this as a fundamental policy message: *if* transport investment has a causal effect on wages that is captured in the correlation between accessibility and wages, then most of the overall wage gains of improving accessibility come from the changing composition of labour markets *not* from improved wages for existing workers (i.e. those that do not change education, occupation, industry or ability in response to increased accessibility). As the composition of the Manchester-Leeds economies shifts towards higher educated, higher ability workers average wages rise by between 1.06% (Tameside) and 2.65% (Wakefield). But the gains to existing workers who do not change their characteristics in response to increased integration are considerably smaller. We return to the implications of this below.

Table 4: Percentage change in wages for a 20 minute reduction in Manchester-Leeds train time

LAD NAME	CR	L-M -20m	[1]	[2]	[3]	[4]	[5]	[6]
Bradford	L	6.59	1.70	1.83	1.14	1.09	1.12	0.32
Calderdale	L	6.05	1.56	1.68	1.05	1.00	1.03	0.30
Craven	L	6.3	1.63	1.75	1.09	1.04	1.07	0.31
Harrogate	L	6.98	1.80	1.93	1.21	1.15	1.19	0.34
Kirklees	L	6	1.55	1.66	1.04	0.99	1.02	0.29
Leeds	L	9.75	2.52	2.70	1.69	1.61	1.66	0.48
Selby	L	6.51	1.68	1.80	1.13	1.07	1.11	0.32
Wakefield	L	10.26	2.65	2.84	1.77	1.69	1.74	0.50
Bolton	M	6.17	1.59	1.71	1.07	1.02	1.05	0.30
Bury	M	6.24	1.61	1.73	1.08	1.03	1.06	0.31
Congleton	M	6.29	1.62	1.74	1.09	1.04	1.07	0.31
High Peak	M	5.22	1.35	1.45	0.90	0.86	0.89	0.26
Macclesfield	M	7.84	2.02	2.17	1.36	1.29	1.33	0.38
Manchester	M	10.07	2.60	2.79	1.74	1.66	1.71	0.49
Oldham	M	4.56	1.18	1.26	0.79	0.75	0.78	0.22
Rochdale	M	4.34	1.12	1.20	0.75	0.72	0.74	0.21
Salford	M	4.42	1.14	1.22	0.76	0.73	0.75	0.22
Stockport	M	7.62	1.97	2.11	1.32	1.26	1.30	0.37
Tameside	M	4.12	1.06	1.14	0.71	0.68	0.70	0.20
Trafford	M	6.4	1.65	1.77	1.11	1.06	1.09	0.31
Vale Royal	M	6.21	1.60	1.72	1.07	1.02	1.06	0.30
Warrington	M	6.86	1.77	1.90	1.19	1.13	1.17	0.34
Wigan	M	6.47	1.67	1.79	1.12	1.07	1.10	0.32
Multiply percentage change by			0.25800	0.27700	0.17300	0.16500	0.17000	0.04900

Notes: Table 4 shows percentage change in accessibility for a 20 minute reduction in train journey times between Manchester and Leeds (L-M-20m). Column [1] shows total effects including any compositional changes; [2] controls for age, age squared and gender; [3] controls for years of education; [4] controls for occupational characteristics (1-digit level) and dummies for part-time, public sector and collective wage agreement); [5] controls for 1-digit industry; [6] controls for individual fixed effects. The final row corresponds to the coefficients in columns [1] to [6] reported in Table 2.

2.4 Results: wage growth

To reiterate, our results so far suggest that any substantive impact on wage levels from greater integration of labour markets come mostly from changing the composition of individuals and partly from changing the composition of work via effects on industrial structure and occupation. The effects on workers who do not change individual characteristics (education, ability) are quite small. In this sub-section we briefly consider the related question of whether accessibility plays a role in driving individual wage growth rather than levels. That is, we consider the possibility that accessibility is more important for understanding the dynamics of the labour market.

The sample of individuals used to study wage growth is essentially the same as that used for wage levels (some additional trimming eliminates very large growth rates). The dependent variable is annualised percentage wage growth over the period of observation of the individual: $\ln(w_T - w_{t0})/(T - t_0)$ where w_{t0} is the individuals' wage in the first year they are observed and w_T is the wage in the final year. Wage growth is normalized by the number of years $T - t_0$ over which the individual is observed to allow for the fact that we observe individuals for different lengths of time.

We work through the same set of specifications as for wages. Because we are looking at wage growth over a period of years we need to decide which characteristics we measure at the start of the period and which we allow to vary over time. Sex is fixed and we measure age and experience at the start of the period. For the remaining individual and job characteristics we simply take the average over the period for which we observe the individual. We also time-average accessibility and area characteristics for each individual (thus allowing for the fact that individuals may move across TTWAs).

We start by regressing growth in wages on both accessibility measures. Results are reported in column 1 of Table 5. The effects are an order of magnitude smaller than those for wage levels. This is reassuring as large differences in growth rates quickly translate into very large differences in the levels of wages (because of the "compound interest" nature of wage growth). The meaning of the coefficient of 0.067 on Train Accessibility is that a 10% improvement in Train Accessibility increases annual wage growth by roughly 0.7 percentage points. We now start to introduce individual characteristics in the same order as for the wage regressions. As before, we only report the coefficients on accessibility. Adding sex, age and age squared (column 2) makes the negative effect on Car Accessibility insignificant and substantially reduces the coefficient on Train Accessibility. Adding education (column 3) has a similar effect. Adding occupational controls (1 digit occupation dummies plus part time, public sector and collective agreement) turns Car Accessibility positive and Train Accessibility negative (column 4). Once we include industry dummies (column 5) we are left with a very small effect of Car Accessibility on wage growth, but no effect from Train Accessibility. When we add in area controls, industrial diversity etc, even the effect of Car Accessibility disappears. Note that the fact that we consider average wage growth over the period means we cannot control for individual unobserved characteristics. These made a large difference for wage levels, but their omission here is of less concern because we do not find particularly strong evidence of an effect of accessibility on wage growth when we control for observed individual and area characteristics.

Table 5: Regressions of wage growth on accessibility and other variables

	[1]	[2]	[3]	[4]	[5]	[6]
ln Car Accessibility	-0.018**	-0.003	-0.003	0.006*	0.007**	0.005
	(0.006)	(0.005)	(0.006)	(0.003)	(0.003)	(0.004)
ln Train Accessibility	0.067**	0.0152**	0.010*	-0.012*	-0.011	-0.001
	(0.008)	(0.005)	(0.005)	(0.006)	(0.006)	(0.005)
Observations	248068	246125	246125	246125	246125	246125
R^2	0.00	0.08	0.08	0.10	0.10	0.10

Notes: All models have annualised percentage wage growth over the period of observation of the individual as dependent variables and the explanatory variables of interest are logarithms of car and train accessibility variables. Column [1] has no controls; [2] adds age, age squared and gender; [3] adds years of education; [4] adds occupational characteristics (1-digit level) and dummies for part-time, public sector and collective wage agreement); [5] adds 1-digit industry controls; [6] adds area level characteristics as described in the text. ***, **, * denote significance at the 1%, 5% and 10% levels respectively.

2.5 Labour markets and agglomeration: conclusions

Our results suggest that closer integration between labour markets may deliver additional benefits in terms of increased area wages. Whether these benefits are actually delivered by transport improvements depends on the extent to which the observed correlation between accessibility and wages is actually capturing the causal impact of transport. We have not done much to address this question in this chapter and the limited literature that does try to assess this causality (see Ahlfeldt and Feddersen, 2010 and Gibbons, Lyytikainen, Overman and Sanchis, 2010) urges considerable caution in attributing all (or indeed any) of this correlation to the causal effects of transport. Regardless of the causal relationship, the results in this chapter urge further caution for policy makers. While our estimates for a 20 minute reduction in train journey times between Manchester and Leeds have wages increasing by between 1.06% and 2.7%, nearly all of these wage effects come through the changing *composition* of the workforce (arising through sorting, and/or because people change their characteristics in response to changes in accessibility). The effects for any given individual who does not increase their education or skill levels (the place-based effects) are small at somewhere between 0.20 and 0.50 of a percent. Consistent with this, individual wage growth is faster in places with accessibility, but this effect appears to be driven by the fact that these places tend to have more educated workers. Once we control for this there is essentially no relationship between labour market size and wage growth.

Overall, the findings suggest that the aggregate effects of closer integration may be larger than the individual effects. This aggregated effect relies on structural changes moving the composition of better integrated labour markets towards higher skilled jobs. From a traditional cost-benefit perspective, these effects would *not* be counted as additional for individual projects if, as is likely, they come about because of greater attraction or retention of existing skilled workers. If they occur because existing workers increase their education or skills in response to changing economic opportunities some part of these higher gains may be additional (to the extent that the individual benefits of increasing, say, education, outweigh the costs). Regardless of the mechanism, if increased integration does lead to structural change (and again, we emphasise that this chapter has done little to address the crucial issue of causality) these compositional changes will increase aggregate output in better connected labour markets, and this will be of interest to policy-makers interested in the performance of the better connected places. In our case study, the estimated impact of closer integration between Manchester and Leeds is dependent on induced changes in the composition of the population. It represents an upper bound of the possible effects as we cannot rule out the possibility that some of this effect runs *from* the composition of the labour market *to* lower transport costs (rather than vice versa). We find evidence that the effect on wages for individuals who do not change their personal or job characteristics are small (between 0.2% - 0.5%). This modest impact on the wages of workers whose characteristics remain unchanged is likely to be offset or even reversed by induced increases in the cost of living.

3 References

Ahlfeldt, G. and A. Feddersen (2010) 'From periphery to core: economic adjustments to high speed rail'. Processed, LSE.

Aschauer, D.A. (1989), 'Is Public Expenditure Productive?' *Journal of Monetary Economics*, Vol. 23, pp 177-200.

Boarnet, M. G. (1997), 'Highways and economic productivity: Interpreting recent evidence', *Journal of Planning Literature* 11(4), 476-486.

Combes, P.-Ph., Duranton, G. and Gobillon, L. (2011), 'The identification of agglomeration economies', *Journal of Economic Geography* 11(2), 253-266.

DETR (2001) Transport Economics Note, DETR, London.

Duranton, G. and D. Puga (2004), 'Micro-foundations of urban agglomeration economies' in J. V. Henderson and J. F. Thisse (ed.), *Handbook of Urban and Regional Economics* (Vol 4) p.2063-2117 Elsevier.

Gibbons, S., Lyytikainen, T., Overman, H. G. and Sanchis, R. (2010), 'Productivity and employment impacts of agglomeration: evidence from transport improvements', SERC/LSE/UKTRC working paper.

Gibbons, S., Overman, H. G. and Pelkonen, P. (2010), 'Wage disparities in Britain: People or place?', SERC/LSE discussion paper 0060.

Graham, D. (2006), 'The localization of service industries in Britain', working paper, Centre for Transport Studies, Imperial College London.

Graham, D. and Melo, P. (2010), 'Agglomeration economies and labour productivity: Evidence from longitudinal worker data for GB's travel-to-work areas', Spatial Economics Research Centre (SERC) discussion paper 0031.

Gramlich, E. M. (1994), 'Infrastructure investment: A review essay', *Journal of Economic Literature,* American Economic Association, 32(3), 1176-96.

Holl, A., (2010) Market potential and firm-level productivity in Spain. Processed, FEDEA, Madrid.

IPEG/CUPS (2008), 'The Northern Connection', report for the Northern Way.

Kain, J and J. Persky (1969) Alternatives to the Gilded Ghetto, *The Public Interest* 14: 74-83.

Levinson, D. (1996) 'Accessibility and the journey to work', *Journal of Transport Geography*, Volume 6, Issue 1, March 1998, Pages 11-21, ISSN.

Lucci, P., and P. Hildreth (2008), 'City Links', Centre for Cities Report for the Northern Way.

Manski, C. F. (1993). Identification of endogenous social effects: The reflection problem. *Review of Economic Studies*, 60: 531-42.

Melo, P. C. and D. J. Graham (2010), 'Agglomeration economies and labour productivity: Evidence from longitudinal worker data for GB's travel to work areas', SERC Discussion Paper No. 31.

MIER (2009) 'The Case for Agglomeration Economies', Manchester Independent Economic Review.

Overman, Henry G., Redding, Stephen J. and Venables, Anthony J. "The economic geography of trade, production and income: A survey of empirics". In *Handbook of International Trade*, edited by E. Kwan-Choi and J. Harrigan. Basil Blackwell, 2003.

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

Rosenthal, S. S, and W. C. Strange (2004), 'Evidence on the Nature and Sources of Agglomeration Economies' J. V. Henderson and J. F. Thisse (ed.), *Handbook of Urban and Regional Economics* (Vol 4) . 2119-2171 Elsevier.

Wheeler, C. H. (2006) 'Cities and the growth of wages among young workers: Evidence from the NLSY', *Journal of Urban Economics* 60, 162-184.

Northern Way (2009), 'Strengthening Economic Linkages between Leeds and Manchester: Feasibility and Implications Strengthening Economic Linkages between Leeds and Manchester: Feasibility and Implications'.