2023

AGGLOMERATION AND TRANSPORT APPRAISAL: NEW DEVELOPMENTS AND RESEARCH DIRECTIONS

James Laird and Eivind Tveter

Peak Economics and Møreforsking

31st March 2023

Blank.

AGGLOMERATION AND TRANSPORT APPRAISAL: NEW DEVELOPMENTS AND RESEARCH DIRECTIONS

Final report to Department for Transport

31st March 2023

James Laird

Peak Economics

Eivind Tveter

Møreforsking

Blank.

Contents

ABSTRACT

The theoretical and empirical literature on agglomeration and its role in transport appraisal is now about two decades old. The seminal theoretical contribution by Venables (2007) using a monocentric urban model has been extended to situations where there are multiple workplace locations (Eliasson and Fosgerau, 2019), to situations of imperfect competition (Kidokoro, 2015, Kanemoto, 2013) and to situations where transport projects trigger land use change and specialisation (Venables, 2017).

The evidence base on agglomeration elasticities has continued to grow. The Donovan et al. (2021) meta-analysis identified 294 studies. This large evidence base permits much firmer conclusions to be drawn on the likely range of well-estimated urbanisation agglomeration elasticities, with Donovan et al. citing a range of 0.015 to 0.039. Whilst difficult to estimate there is small but growing evidence that the public sector experiences benefits from agglomeration. There is a greater understanding that clustering can happen at very local levels. But it can also span several hundred miles (related to freight haulage distances), with innovative, creative activities also being sparked by connections to high speed rail or air networks. It is no longer thought to be a phenomenon that operates purely at the city level.

Over the last twenty years there has also been a growing interest in the city as a place for consumption, and the benefits cities offer to consumers. The amenity benefits of cities, including the gains from variety, or consumption externalities are non-market benefits, in that they do not affect firm productivity. Consumption externalities are additional to transport user benefits, but other amenity benefits double count benefits already included in the transport appraisal. Turning to the agglomeration calculations it is now known that the spatial zoning system can influence the agglomeration benefits associated with transport projects, and that there is likely to be a long ramp-up period before agglomeration economies appear.

Despite the improved theoretical treatments and the much larger empirical evidence base there remain gaps. In terms of impact on appraisal, embryonic research indicates that localisation economies and consumption externalities may add up to two thirds of productivity related agglomeration benefits, but the evidence base supporting these elasticities remains thin. The role of the Access To Economic Mass parameter in determining the scale of the agglomeration elasticity and in fully capturing interactions that are both very local, at a city level and over a regional level has not been explored. Within the UK the empirics of agglomeration in transport appraisal have remained broadly unchanged aside from periodic updates of the parameters, primarily the GDP and employment data. The underlying equations remain as they were twenty years ago, although the parameters have been updated more recently. Recommendations have been made as to how guidance can be updated to reflect the improved evidence base, and areas where future research may be conducted to address policy needs.

1 INTRODUCTION

1.1 Background

Cities and agglomeration benefits

Cities exist partly because firms, workers and households obtain benefits from co-locating with each other. These are agglomeration benefits. Agglomeration benefits in the production sphere have received the most attention. These are the productivity benefits that firms experience from co-locating (Marshall, 1890, Jacobs, 1984). More recently the consumption benefits of households have been identified as an important reason households co-locate in cities (Glaeser et al., 2001). Households benefit from more diversity being available in cities, affecting the way households interact, socialise and consume services (e.g. education, healthcare, retail, leisure services). There are however also economic costs to co-locating in terms of congested networks, crime and high rents.

Better transport effectively brings people and businesses closer together – not in distance but in a temporal sense. This effectively makes co-location more intense. With more intense co-location, household welfare and business productivity improves. These are the agglomeration benefits. One way to measure 'better transport' is through changes in generalised transport cost (GTC) – an amalgamation of both the financial and cognitive costs of travel. Co-location, that is how close businesses are to each other can be measured through Access To Economic Mass (ATEM)[1]. Linking changes in agglomeration to changes in productivity through an agglomeration elasticity then provides a measure of the agglomeration benefits.

The TAG approach

Agglomeration benefits in TAG (Department for Transport, 2020) are calculated in what is easiest to think of as a four step process.

1. Calculate a trip weighted average $\overline{GTC}_{ij}^{m}$ by mode for each origin-destination pairs across all time periods and journey purposes.

$$\overline{GTC}_{ij}^{m} = \frac{\sum_{p,t} GTC_{ij}^{m,p,t}.T_{ij}^{m,p,t}}{\sum_{m,p,t} T_{ij}^{m,p,t}} \qquad (1)$$

$GTC_{ij}^{p,m,t}$ is the generalised transport cost between zones *i* and *j* for journey purpose *p* by mode *m* and in time period *t*.

$T_{ij}^{p,m,t}$ are the number of trips between zones *i* and *j* for journey purpose *p* by mode *m* and in time period *t*

Notes: The origins *i* and destinations *j* are required to be at a Local Authority District level of aggregation, so $\overline{GTC}_{ij}$ may need aggregating to such a zonal system. WITA uses a trip

[1] Alternative terms with similar function forms, Economic Mass and Economic Density, have also been used. Access To Economic Mass was first used by VENABLES, A. J., LAIRD, J. & OVERMAN, H. 2014. Transport investment and economic performance: Implications for project appraisal.

weighting averaging process for this. Only commuting and business journey purposes are considered. Freight can be included as a sensitivity test, and non-work purposes are excluded.

2. Calculate ATEM by industry (I). TAG uses a power decay functional form. The decay parameter α^I varies by industry. ATEM is calculated for each mode and then summed to get the ATEM for the zone by industry.

$$ATEM_i^I = \sum_j \sum_m \frac{E_j}{\left(GTC_{ij}^m\right)^{\alpha^I}} \quad (2)$$

Where: $ATEM_i^I$ is the ATEM of zone *i* for industry *I*

E_j is employment in zone j and is the sum of employment across all industries in zone *j*.
α^I is the decay parameter in the power decay functional form, and varies by industry.

3. Calculate the percentage change in productivity due to the change in GTC by industry.

$$\theta_i^I = \left(\frac{ATEM_i^{I,DS}}{ATEM_i^{I,DM}}\right)^{\rho^I} - 1 \quad (3)$$

Where: θ_i^I is the percentage change in productivity in zone *i* for industry *I*
ρ^I is the agglomeration elasticity by industry *I*
DM is the Do Minimum scenario, and
DS is the Do Something scenario

4. Calculate the economic benefits arising from changes in agglomeration ($WI1_i$) for each zone

$$WI1_i = \sum_I \theta_i^I . GDPW_i^I . E_i^{I,DS} \quad (4)$$

Where: $WI1_i$is the GDP impact in zone i (known as Wider Impact 1 (WI1) in TAG terminology)
$GDPW_i^I$ is the GDP per worker in zone *i*, industry *I*
$E_i^{I,DS}$ is the employment in zone *i* for industry *I* in the Do Something

1.2 Study objectives and report structure

The TAG methodology described above was first set out in 2005 (Department for Transport, 2005) and has changed very little since, partly because the theoretical work that underpins it (Venables, 2007) is considered robust (Laird and Venables, 2017). Along with other Wider Impact guidance it was reviewed in 2014 (Venables et al., 2014). The review process led to the suite of Wider Impact TAG Units to be developed by the Department (Units A2-1 to A2-4), with the agglomeration aspects appearing in Unit A2-4. There was no substantive change to the agglomeration guidance formulas as part of this review. The main changes that affect the benefit estimates have been through updates to the data which accompany the guidance: the agglomeration elasticities and the GDP and employment data. There have been two sets of agglomeration elasticities, both of which were estimated by Professor Dan Graham (Graham, 2007, Graham et al., 2009). Both capture urbanisation economies, the economies from co-locating next to others in general. The latest set were innovative at the time in that they not only disaggregated by industrial sector, but also explicitly allowed agglomeration impacts to decay at different rates (by industrial sector). The GDP and employment data is re-visited intermittently by the Department to ensure it is current. These data are contained in the TAG Wider Impacts Dataset last updated in July 2021 (Department for Transport, 2021).

The 2014 review identified some weaknesses in the methodology (Venables et al., 2014). There was no treatment of localisation economies (the external economies to firms from co-locating near to similar firms), the GTC averaging method across modes and time periods (Steps 1 and 2 above) was based on intuition and not evidence[2], and the treatment of journey purposes likewise (business and commuting were included, but not freight or non-work) (Step 1 above). An obvious weakness is the lack of a public sector agglomeration elasticity in the guidance. This arises as Graham et al. (2009) estimated the agglomeration elasticities to firm level data, which excludes the public sector. Another obvious weakness is that the Graham et al. (2009) elasticities were estimated using an ATEM function with distance as the denominator, but are applied with GTC as the denominator (Step 2 above). Empirically some of these evidence gaps are difficult to address due to collinearity and endogeneity in the data (Combes and Gobillon, 2015, Combes et al., 2011), and the Department's recent empirical efforts have struggled for these reasons. One tangible output of the Department's efforts has been a review piece that sets out the theoretical and empirical foundations to the inclusion of agglomeration benefits in a transport appraisal (Graham and Gibbons, 2019).

Since the Venables et al. (2014) review, some experimentation with the functional form of ATEM has occurred in a small section of the transport economic literature (KPMG, 2013, Gibbons et al., 2019, Knudsen et al., 2022), the discussion around the consumption (amenity) benefits to households in cities (Ahlfeldt and Pietrostefani, 2019, AitBihiOuali, 2022) has again raised the issue of double counting with other benefits in the transport appraisal, and the policy agenda on economic growth and regional growth (levelling up/re-balancing) leads to an interest in who the beneficiaries of further agglomeration are.

The objective of this think piece therefore is to re-examine the agglomeration literature with the aim of identifying contributions that can be made around these evidence gaps, and future research directions. There is a particular interest in the beneficiaries of agglomeration, double counting with other benefit categories in a transport appraisal, benefits of localisation economies, evidence on the benefits of agglomeration to the public sector, and the structure of ATEM. The latter would most likely be informed by recent regional economic literature research which has continued to focus on the identification of the micro-mechanisms that lead to agglomeration economies (see e.g. Laird and Tveter (2021) for a review).

Consequently, following this introductory section, Chapter 2 discusses the economics of agglomeration including the production and consumption externalities of cities, current issues in

[2] The Graham et al. (2009) agglomeration elasticities are based on crow-fly distances between wards. The empirical analysis therefore cannot inform how to combine generalised cost between modes, time periods, etc. Judgements were therefore made in the original work (dating back to 2005), and these have remained unchanged.

their empirical measurement, and their inclusion in a transport cost benefit analysis with or without double counting. In Chapter 3 we turn to the issue of public sector productivity gains as a result of increased agglomeration, whilst in Chapter 4 we examine some of the more technical issues surrounding the ATEM function and how that impacts on the agglomeration elasticity and the benefits. Implications for policy and research are presented in Chapter 5.

2 THE ECONOMICS OF AGGLOMERATION

2.1 Productivity benefits to firms

Agglomeration economies are external economies of scale. They are business enhancing factors that are outside the firm, but affect the firm's performance. Urbanisation agglomeration economies, the benefits of co-locating in an urban area, are external to the firm, but internal to the whole urban area. If the urban area gets bigger, or denser, and urbanisation economies are positive, firms located in them get more productive. Likewise, localisation agglomeration economies, the benefits of co-locating with similar businesses, are external to the firm, but internal to the industry. Therefore, if an industry gets larger and localisation economies are positive, firms within that industry get more productive. Agglomeration economies are therefore a form of spatial productivity externality. They are spatial as they are driven by proximity. They are an externality, as the productivity of a firm or worker is affected by the actions of other firms and worker.

The productivity benefits of agglomeration are considered to stem from matching, sharing and learning behaviours of firms and workers, with the Marshallian sources of thicker labour markets, knowledge spillovers and better vertical linkages now considered aggregate effects (Rosenthal and Strange, 2001, Duranton and Puga, 2004, Rosenthal and Strange, 2004, Puga, 2010, Duranton and Puga, 2020). These microfoundations and the evidence for them have been recently reviewed for the DfT (Laird and Tveter, 2021). Readers are referred to that review for more information on these microfoundations.

In terms of evidence on the aggregate effects of agglomeration economies on productivity, Rosenthal and Strange (2004) in their review considered that agglomeration economies give rise to an increase of productivity of between 4 and 11% for a doubling of city size corresponding to an elasticity range of between 0.03 and 0.08 . In their meta-analysis Melo et al. (2009) found a mean value of 0.058 with a standard deviation of 0.115, a 5th centile of -0.09 and 95th centile of 0.292. This is quite a large range. This variation is also not random. Melo et al. (2009)'s analysis indicates service industries benefit the most from agglomeration, whereas manufacturing sits around the economy wide average. Study characteristics also matter as the results vary systematically with the data utilised and the empirical strategy adopted. It is generally considered that the more robust studies have lower elasticities. Ahlfeldt and Pietrostefani (2019) in their meta-analysis of 47 recent estimates find an average agglomeration elasticity for wages of 0.04 which doubles for developing countries. They do not try to explain any of the variation in the estimates. Donovan et al. (2021) conduct a very large meta-analysis of 6,684 estimates of agglomeration elasticities from 294 studies covering 54 countries and six decades. They find a global average of 0.072. However, for robust estimates estimated to wage data[3] they consider the average to be 0.026 (s.e. 0.006) with a 95% credibility interval of 0.015–0.039. They do not present variations by industrial sector, but find that there is little variation by country nor by national income. This contrasts with Melo et al. (2009) who found differences between countries, and Ahlfeldt and Pietrostefani (2019) who found differences between high income and middle/low income countries. They also found that elasticities for manufacturing sectors seem to have declined over time. A fourth meta-analysis by Grover et al. (2021) focusing on developing countries finds that agglomeration elasticities in developing countries

[3] statistically significant elasticity, published in a peer reviewed journal, but controlled for publication bias, derived from panel data, uses wages, measures agglomeration as population potential, controls for individual worker effects, own-skills and human capital and addresses endogeneity between wages and agglomeration.

are similar to those in developed countries, despite the apparently large estimates for productivity increases with city size in China, India and Africa.

The Graham et al. (2009) urbanisation agglomeration elasticities used by DfT in their TAG guidance have an average economy wide elasticity of 0.04, with higher observed elasticities for producer services sector 0.083 compared to manufacturing (0.021), construction (0.034) and consumer services (0.024). These elasticities seem broadly consistent with this broad evidence base. They are towards the lower end of the range observed by Rosenthal and Strange, based on older studies, but towards the upper end of the range considered credible by Donovan et al. for high quality empirical studies. They are however becoming dated, and there is the potential that current elasticities may now differ. Having said that the meta-analysis work by Donovan et al. (2021) does give some confidence to the continued use of the elasticities, as it suggests that agglomeration elasticities may be reasonably stable over time. The temporal stability of elasticities does however remain an ongoing field of research, so it is too early to make firm conclusions in this regard.

2.2 Estimating and choosing agglomeration elasticities for use in transport appraisal

The empirics of estimating agglomeration economies is a field to which a degree of consensus on appropriate methods has now developed. These are summarised by Combes and Gobillon (2015), and Graham and Gibbons (2018). The latter is a DfT sponsored paper. In summary there is a need to separately identify skills and local effects – with agglomeration being a local place based effect. There is also the need to treat endogeneity, though endogeneity is less of a problem if controls are used in the estimation around which sources of endogeneity may occur (e.g. skills). Successful estimations typically use micro-level panel data and use an identification strategy based on those who move location. A recent successful estimation with a transport appraisal agglomeration parameters in mind is by Knudsen et al. (2022), who estimated agglomeration elasticities for use in Danish transport appraisal guidance.

There now exists a substantial evidence base on agglomeration elasticities, with Donovan et al. (2021) identifying 294 studies and 6,694 values. If these values appear stable between countries, those which are based on weak methodologies can be excluded, as Donovan et al. appear to have done so then there is an argument that the meta-work should be used as a basis for the elasticities to be used in appraisal guidance. This would be instead of using the values from a single study. Unfortunately, the difficulty with this is that the majority of the literature on agglomeration elasticities, stems from the regional economics field, rather than transport appraisal. Donovan et al. identify that a third of the studies in their sample use a 'potential' functional form for ATEM, as used in TAG, with a distance decay (as in equation 2). The others use a mass variable determined by administrative boundaries. It would be necessary to focus on the subset of data that uses a potential form. The inter-relationship between the distance decay parameter and the elasticity would also make it difficult to use meta-analysis to select appropriate agglomeration elasticities and decay parameters for use in guidance. This is because the elasticities are paired with their decay parameters (see also Section 4.4). For the moment therefore it appears likely that we will continue to require individual studies to be our source of agglomeration elasticities in appraisal guidance.

Within the literature there is a debate as to whether to use wages or Total Factor Productivity (TFP) as the productivity measure. An example of TFP agglomeration elasticities are those derived by Graham et al. (2010) which TAG values are based on. Transport applications of wage derived agglomeration elasticities include those by D'Costa et al. (2013) for the UK and Knudsen et al. (2022) for Denmark. On one hand wage data is easily available, which makes estimation more tractable, but on the other hand wages are likely to be influenced by only a subset of agglomeration mechanisms, principally matching and learning. Sharing mechanisms between firms are more capital productivity related. From that perspective TFP would be the better measure of productivity, and arguably it has other advantages from an estimation perspective (Combes and Gobillon, 2015 pp282-284, Graham and Gibbons, 2019). Added to this debate is that TFP elasticities might be biased upwards by the manner that in dense locations journeys are shorter (see Ahlfeldt and Pietrostefani (2019) and section 2.8). Therefore, business travellers and freight spend less time travelling and more time engaged in productive activities. Thus, in dense locations, where businesses are very close, businesses will be more productive than those located in less dense location, even in the absence of agglomeration economies.[4] This will likely bias empirical estimates of agglomeration elasticities in an upwards direction, particularly TFP agglomeration elasticities.[5] Wage derived agglomeration elasticities are unlikely to be so affected, as in competitive markets the wage (i.e. the marginal product of labour) is not affected by business time savings – certainly at the margin. The debate as to the most appropriate productivity measure to use is not therefore as conclusive as one may wish it to be, with positives and negatives associated with both measures.

2.3 Additionality to transport user benefits

Venables (2007) uses a simple mono-centric urban model to demonstrate the additionality of the productivity related agglomeration benefits to transport user benefits. The essence of this argument is that transport user benefits only equal total economic value under conditions of perfect competition. The existence of agglomeration economies, as spatial externalities, means additional benefits to the user benefits will arise.

In Venables' model economic activity is located in the centre, households commute to the centre, and land rents completely offset commuting costs (*c*) (Figure 2-1 (A)). City size is determined by the difference between worker productivity (wages) in the city and in the city hinterland. City size (*N*) is stable when the wage gap (x-$\bar{x}$) between the city and the rural hinterland exactly equals the land rent/commuting cost premium paid by city residents.

[4] This is likely most relevant between rural and urban locations. Between urban areas, trip lengths are relatively stable. See for example National Travel Survey Table NTS9911: Average number of trips (trip rates) by trip length, region and Rural-Urban1 Classification: England, 20212. https://www.gov.uk/government/statistical-data-sets/nts99-travel-by-region-and-area-type-of-residence#trip-length

[5] Some commentators might view this upward bias as a form of double counting. The distinction between an upward bias and double counting is blurred at the boundary. We interpret double counting to be where benefit categories are interrelated and move together as demands and costs change. For example, value of travel time savings that vary with congestion levels will give time saving benefits that double count any reliability benefits. Such time saving benefits and reliability benefits will move proportionately, and both are capturing elements of the same impact. Similarly, surpluses in the land market will double count consumers' surplus in the transport market. Again, both are measuring the same thing and move proportionately, as one increases so will the other. Here agglomeration benefits are not related to travel demand flows. They are a function of potential accessibility. Furthermore, the error is likely most relevant between rural and urban locations, rather than between different urban locations. Thus the error is likely non-linear with respect to density and not proportional to increases in density. We therefore interpret this as an overestimate of the agglomeration benefits, rather than a double counting of transport user benefit.

A transport improvement shifts commuting costs downwards. This destroys the equilibrium between the rural and city populations, and the city expands to (*N + dN*) in Figure 2-1(B). At this point the wage gap equals commuting costs for the longest commuting journey once again. The change in resources used for commuting is $\eta - \alpha$, which combined with the change in output ($\beta + \eta$), yields a net benefit from the transport improvement of $\alpha + \beta$. These are the transport user benefits as measured in a normal transport appraisal.

If we now take it that agglomeration economies exist, that they are positive and diminishing in city size.[6] This then leads to the concave (through the origin) wage gap curve depicted in Figure 2-1(C), rather than the fixed (horizontal) wage gap in Figure 2-1(B). The reduction in commuting costs, by leading to an expansion in city employment now also leads to an increase in wages in the city. The result is more workers are attracted to live in the city than in Figure 2-1(B). The benefit from the transport cost reduction is $\alpha + \beta + \delta$, where δ is the increase in wages due to increased agglomeration. That is the agglomeration benefits.

With a robust measure of the agglomeration elasticity and projected changes in employment, we can measure the full benefit from the transport improvement by summing transport user benefits with the projected increase in wages[7]. In TAG terminology this would be "*dynamic agglomeration*". This is because land uses change, it's dynamic. With employment shifting to the city dis-agglomeration maybe occurring elsewhere, and such costs would also need to be included in the appraisal (Kanemoto, 2013).

Of course, employment is not entirely concentrated in the centre of a city and is dispersed over a city or even a region. Therefore a transport improvement can bring workplaces closer together, as well as changing the number of people who work in a location. Eliasson and Fosgerau (2019) consider this in a model where households and workplaces are distributed over any spatial area. Workers can choose between workplace locations, their choice dependent on the net of tax wage and transport costs. Wages are dependent on the productivity of a location, which in turn is a function of proximity to workers in other locations. In this situation productivity of a location is a function of transport costs between workplaces (firms), or as Eliason and Fosgerau term it job-to-job accessibility. In their model household residence is fixed. They show that in this situation agglomeration benefits (spillovers) are driven changes in the numbers in workers at workplaces (as per Venables) and by job-to-job accessibility changes.[8] The agglomeration benefits stemming from the job-to-job accessibility changes are also additional to transport user benefits.

[6] If agglomeration economies (in aggregate) did not exhibit diminishing returns to city size, then cities would expand indefinitely. Negative costs to agglomeration include congestion on infrastructure both transport and other infrastructure, higher rents and higher wages.

[7] Only changes in wages occur in the Venables model, but it would be the change in GDP in a broader economic model.

[8] Eliason and Fosgerau (2019) consider that the matching effect is internalised through higher wages received by workers, and that changes in user benefits therefore capture matching benefits. However, some of the matching effects are external to the worker. Firstly, the worker vacates a job that would be a better match for others, and secondly the higher productive of the match is shared between workers and firms (firms see higher profits without any increase in their transport demand) (Graham and Gibbons, 2019).

The requirement to model changes in land uses is challenging. A simplifying assumption therefore is to assume fixed land use change. In TAG this is known as "Static agglomeration". This could be seen as a restricted version of the Eliasson and Fosgerau model.

FIGURE 2-1 (A): URBAN EQUILIBRIUM

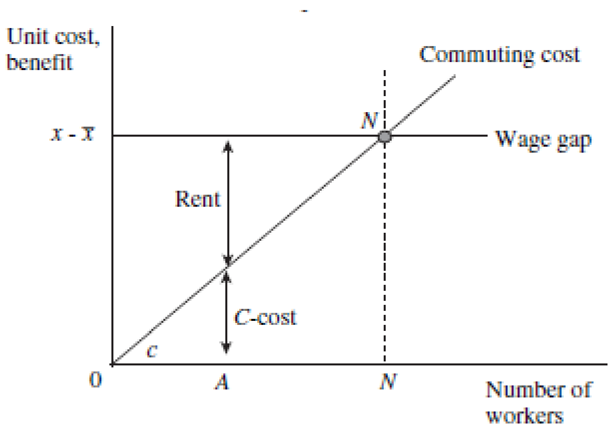


FIGURE 2 1 (B): NET GAINS FROM TRANSPORT IMPROVEMENT

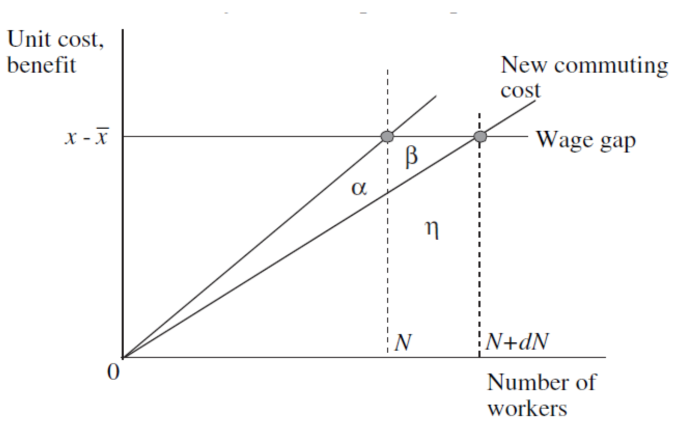
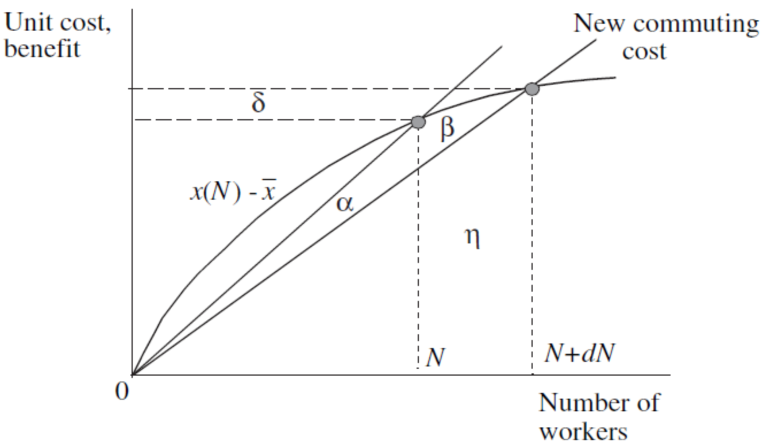


FIGURE 2 1 (C): NET GAINS FROM TRANSPORT IMPROVEMENT WITH ENDOGENOUS PRODUCTIVITY



Source: Venables (2007)

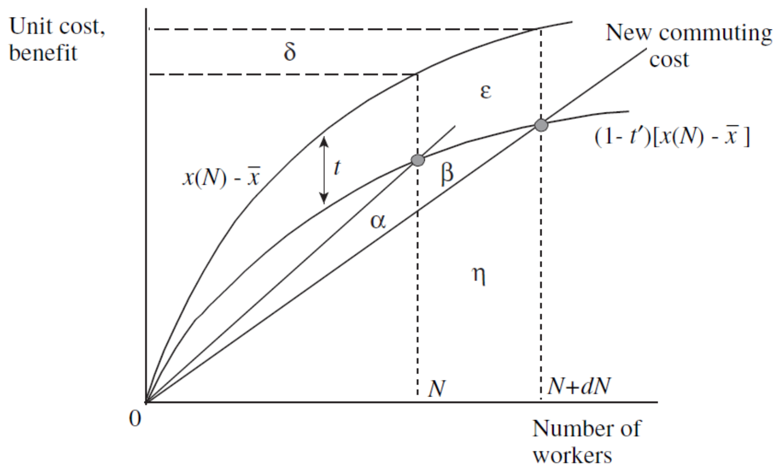
2.4 Interaction between agglomeration benefits and other wider economic impacts

Agglomeration economies will not occur in isolation from other forms of market failure. Potentially they may interact with them. Two particular ones that have received attention are that of labour taxes and coordination failure. The inter relationships between the market failures mean that treating them independently in a form of partial equilibrium analysis will only ever be a simplification and approximation to full general equilibrium analysis.

Agglomeration and labour taxes

Venables (2007) introduces labour taxes into the framework described in Figure 2-1. This is reproduced in Figure 2-2. The productivity gain from the city is the curve $x(N)-\bar{x}$, but workers' decisions are taken on the basis of the curve $(1-t')[x(N)-\bar{x}]$. Consequently, the output increment produced by marginal city workers now exceeds their commuting costs. Commuters only receive the net of tax share of the extra output they produce, but pay the whole of commuting costs. The difference, area ε, accrues to government as tax revenue. Total benefits therefore equal $\alpha + \beta + \delta + \varepsilon$. This additional tax revenue is the Move to More/Less Productive Jobs wider impact that is included in TAG.

FIGURE 2-2: NET GAINS FROM TRANSPORT IMPROVEMENT WITH ENDOGENOUS PRODUCTIVITY AND TAX WEDGE (T)



There are several points to note. Firstly, the relevant tax rate for the calculation of ε is the marginal tax rate t', not the average tax rate. Secondly, the increase in labour employment that drives the increase in city size could come about from both an increase in labour supply (due to higher real wages – a combination of lower commuting costs and higher labour productivity) or a displacement of labour. This has important implications for the consistency in the parameters being used for the estimation of changes in labour supply, Move to More/Less Productive Jobs (M2MLPJ) and agglomeration. Laird et al. (2020) considered these parameters in TAG to be inconsistent with each other and in need of revision.

Related to these underlying mechanisms there could be social costs elsewhere in the economy if employment has been displaced. Thirdly, and as illustrated in Figure 2-2 the rising and non-linear boundary between areas δ and ε will result in some overlap (i.e. overestimation of total benefits) if

δ and ε are estimated using fixed parameters over all affected workers. For example, if the areas are estimated using 'rectangles' rather than 'trapeziums', which is what happens in TAG. Only a general equilibrium analysis would avoid this potential (partial) double counting of benefits, though whether this is necessary in practice would very much depend on the scale of the in employment and agglomeration changes.

Eliasson and Fosgerau (2019) also address the issue of agglomeration and taxation in their model where there are many workplaces instead of just a single workplace in the CBD. They also show that the changes in income tax revenue are additional to agglomeration benefits and user benefits. They indicate that all changes in income tax revenue are additional when households and workplaces are spatially distributed. There is the change in income tax revenue arising from the Move To More/Less Productive jobs, as per the Venables model, but also they disaggregate the agglomeration benefits to existing workers (who do not move) – stemming from what they term increased job-to-job accessibility. This is disaggregated between workers (as increases in net wages) and government (as increased income tax). Whilst the gross change in GDP is included in TAG, from 'static' agglomeration, this is not disaggregated between workers and government (as an exchequer impact).

Agglomeration and imperfect competition

Kidokoro (2015) show that there is an interaction between the imperfect competition wider impact and agglomeration. This is because capital is subject to productivity increases through agglomeration, and capital affects the monopolists' profits. They consider this to be an error in TAG. However, given the typical size of the imperfect competition WI[9], this error is likely to be small, and could probably be ignored.

Agglomeration, land use change and coordination failure

Some commercial developments may potentially be large. High density developments in city centres, and urban regeneration around new rail HSR stations are good examples. If the profitability of the development project for one decision taker depends on investment by others, then there is potential for a coordination failure. It is not in the interest of any single investor to invest, but each would invest if they knew that others were. The positive interdependence between firms could stem from agglomeration (Laird and Venables, 2017). Coordination failures thus lead to low level traps, the private benefits of development are lower than the social benefits. They require some policy mechanism to coordinate individual actions and break out of the trap. Transport investment can be such a mechanism. Transport infrastructure, including possibly the coordinated actions by government to deliver the infrastructure, may act as a catalyst in a growing city as signal that a location will develop. This signal may for example happen by increasing property prices.

Whilst Venables et al. (2014 p76) give a simple two firm example of the problem, as far as we are there have not been any attempts to quantify the coordination benefits that transport investment may bring. Qualitative arguments are often made that the transport investment is a trigger for further investment. How much of this is due to the increased productivity of the location due to

[9] In a TAG appraisal the imperfect competition WI is 10% of business user benefits. Agglomeration impacts for urban schemes typically lie around 50% to 75% of business user benefits. Thus even a large 25% error in the imperfect competition WI is not likely to have a large impact on the sum of business benefits and WIs (1 to 2%). The error in the PVB would be even smaller as non-work user benefits form a substantial component of the PVB (often around 50%).

transport user benefits, and how much is a function of overcoming a market failure to investment remains an unresearched topic.

2.5 Long distance agglomeration benefits

How far do agglomeration economies reach?

In TAG agglomeration benefits are considered to mainly be relevant within functional urban regions (Department for Transport, 2020). The rule of thumb is that they are small relative to the other benefit categories in the appraisal once distances between places of economic mass exceed 45 minutes. Effectively this stems from the functional form and parameters of the ATEM function, which has a power decay form. Once a certain 'distance' has been reached the power decay function means changes in ATEM are small. This has been contested, most notably by KPMG in their HS2 work (KPMG, 2013) with alternative ATEM functions being used. National Highways, with their interest in the strategic long distance road network has also explored this through several consultancy studies. The latest of which being the 2022 study by Jacobs and Atkins. This tranche of work effectively posits that agglomeration impacts can span long distances. Inter-city transport projects or upgrades to long distance transport links reduce the impedance of travelling between these places thereby effectively bringing places closer together. This then generates agglomeration productivity gains for places that are far apart. Within the UK there have been several consultancy studies that have attempted to estimate 'long distance' agglomeration elasticities. None of this work has been published nor, as far as we are aware, has it been accepted by peers. Given this apparent empirical failure, it is worth rehearsing the reasons why we might consider that agglomeration benefits can operate at long distances, and then think whether the scale of these benefits will be substantive or not, before considering if the field warrants additional research.

Rosenthal and Strange (2020) in their paper *How Close Is Close? The Spatial Reach of Agglomeration Economies* identify that the different micro-mechanisms that underpin agglomeration economies operate at different spatial scales. Specifically, they identify that «*sharing of physical inputs, for example, is often associated with truck transport and can extend over regional distances. Labor market pooling is likely to have effects within commuting areas, which is to say at the metropolitan level. Knowledge spillovers as envisioned by Marshall (1890) are unplanned and are likely to be highly local.*» They go on to identify clustering effects at regional, metropolitan, neighbourhood and sub-neighbourhood levels. Clustering is evident at both longer distances and very, very close distances – so close that firms can even cluster between adjacent floors in high rise office blocks (Liu et al., 2020). These clustering effects also vary systematically by industry, with finance being seen as being very tightly clustered, and manufacturing more dispersed (Rosenthal and Strange, 2020). One interpretation of these different clustering patterns is that the different micro-mechanisms to agglomeration have different levels of importance in different industries. If knowledge spillovers are very important, then clustering will occur at local levels. If the primary benefits of agglomeration to some firms is the sharing of inputs between firms, then clustering can occur over much longer distances – potentially several hundred miles. Research and development is an interesting 'in-between'. In their analysis Rosenthal and Strange (2020) show that IT clusters, but unlike finance with a single main cluster in new York (in their case study region of north east USA) there are multiple clusters spread out over a region. They identify that the location of universities is key to this spatial pattern. Of course universities do not operate in isolation and they maintain connectivity with the broader academic network through conferences and other external interactions.

Agglomeration productivity benefits are external to the firm and arise as a consequence of interactions between firms, and between firms and workers. When interactions are most intense is when we would expect agglomeration impacts to be strongest. These interactions are between people and people need to travel to have face-to-face interactions. These interactions therefore require transport trips, thus we can see a link between transport connectivity, interactions and agglomeration. A transport project creates the opportunity to form new interactions, which are evident as new trips (or changes in trip origins and destinations) (Tveter et al., 2017, Tveter, 2018). As the number of trips between origins and destinations tend to diminish with distance then we would expect productivity impacts of agglomeration forces to also diminish, though of course the relationship between trips, interactions and agglomeration may be non-linear. In a long distance context transport projects that create a step change in interactions are therefore of interest: new high speed point to point services either HSR or air may fit into this category. Gibbons and Wu (2020) find that manufacturing firms close to airports in China are more productive than those elsewhere *ceteris paribus*. As part of their analysis they use a domestic market access variable as used in the agglomeration literature. They are however not able to disaggregate the productivity premium between use benefits of the airport and agglomeration, nor between channels of agglomeration. For high speed rail Sun et al. (2021) find that knowledge creation (as measured by patent numbers) is stimulated by new high speed rail lines and stations. They conjecture that the enhanced ability provided by high speed rail to travel for face-to-face meetings or conferences, provide the interactions necessary to stimulate this additional knowledge generation. To the extent that wage data tends to show that the agglomeration premium on wages dissipates almost completely over moderate distances, e.g. Rice et al. (2006) estimate this to be 80 mins for the UK, it would indicate that worker productivity is not affected over long distances. This would imply that it is TFP that is affected at longer distances.

Turning to the micro-foundations to agglomeration economies, we would expect that labour market matching effects will play out over the commutable distances. That is at the local and city levels, although the ability to work from home does extend the effective commute, it will likely remain these effects will remain at the local and city level. The sharing mechanisms will also play out at these local and city levels, but in addition they could play out over much longer distances as businesses are able to trade with other firms quite some distance away. The learning effects may also play out over local neighbourhood levels, the city level and longer distances. Knowledge is passed from one person to another through interactions, new knowledge is created from people interacting, etc. A lot of the impact of learning mechanisms is likely to play out locally, but the need for some long distance interaction, e.g. between research centres or design teams that accidentally meet at conferences will have a role to play. Thus, it is likely that only a subset of the sharing and learning mechanisms underpinning agglomeration economies are likely to play out over longer distances.

From this discussion, we would conclude that long distance agglomeration impacts are likely to apply to only a subset of firms and arise through only a subset of the micro-mechanisms that underpin agglomeration. For the firms/industries to which they apply the elasticities are therefore likely to be smaller than those for shorter distances – which would also include within-labour markets and neighbourhood impacts (the matching and learning mechanisms). Long distance agglomeration impacts appear to be real but are hard to identify. Whether they are significant component of benefits (vis a vis user benefits) in a transport appraisal context has not been investigated. Conceptually, we would expect long distance transport agglomeration benefits to be smaller than ‘conventional’ TAG type benefits due to the manner that they apply to a subset of the underlying mechanisms to agglomeration economies, and very likely a subset of firms.

Empirically, studies commissioned within the UK to identify long distance agglomeration elasticities appear to have failed. Alternative approaches are likely needed. This could include better data or alternative approaches to defining the ATEM function for long distance trips. We discuss the functional form of ATEM in Chapter 4. From a research perspective Rosenthal and Strange (2020) and Donovan et al. (2021) consider that further investigations in this area are of merit. The latter present averages from their meta-analysis on spatial scope, with significant elasticities associated with national and international scope. Having said that they define scope as the population size to which the elasticity appears to apply, which is not the same as distance. Certainly, further work is required in this area before any definitive statements on the scale and relevance to transport appraisal of long distance agglomeration elasticities can be made.

Urbanisation and localisation

An alternative mechanism by which transport projects can affect agglomeration over long distances is through land use change and increased city specialisation (Venables, 2017). The mechanism that creates the agglomeration is an indirect consequence of the transport project. Over long distances reductions in transport costs permit increased trade. This leads to increased specialisation – following principles of Ricardian comparative advantage. If the industries in which a city or region specialises in experience positive localisation agglomeration economies then there will be agglomeration benefits, even if the cities remain far apart in travel times and total employment levels do not change.

Two challenges exist in the application of this approach to transport appraisal. The first is the identification of a consistent set of urbanisation and localisation elasticities, of which there are only five examples in the literature (Pierce, 2020). Localisation elasticities are typically viewed as smaller than urbanisation elasticities and attenuate quite quickly (see e.g. Graham (2009) for a comparison of UK and localisation and urbanisation elasticities).[10] The second is how to predict land use change and increased specialisation that may arise as a result of the transport project (Pierce et al., 2019).

Regarding the impact on transport appraisal, in exploratory research Pierce (2020) case studied an improved rail link between Manchester and Leeds. He used scenarios, following a survey of the literature, to define the range of elasticities and land use change used in the analysis. He found that localisation agglomeration economies are relevant to the appraisal, but are smaller than urbanisation economies. It is worth noting that the case study he considered generates significant urbanisation economies as benefits. With fixed land uses localisation benefits were half the urbanisation benefits. When land uses changed with increased specialisation this gap narrowed. Land use change increased both urbanisation benefits and localisation benefits, but localisation increased by more (+35% versus +7%).

The implication is that localisation agglomeration benefits should be expected, both under static and dynamic clustering (land use change) scenarios. For projects where urbanisation agglomeration benefits are expected to be small, but land use change is expected (i.e. for transport projects that improve connectivity between places that are a long distance apart) then localisation benefits will be larger than urbanisation benefits. They could quite likely be relevant to project appraisal. The implication for TAG is the need to explore how best to bring localisation economies into the

[10] At the moment TAG only provides elasticities on urbanisation. As these elasticities were derived in a manner that did not control for localisation then they will be upwardly biased (i.e. capture more than just urbanisation effects) MELO, P. C., GRAHAM, D. J. & NOLAND, R. B. 2009. A meta-analysis of estimates of urban agglomeration economies. *Regional science and urban Economics,* 39**,** 332-342, DONOVAN, S., DE GRAAFF, T., DE GROOT, H. L. & KOOPMANS, C. C. 2021. Unraveling urban advantages—A meta-analysis of agglomeration economies. *Journal of Economic Surveys*..

appraisal framework. There will be a difficulty in successfully estimating urbanisation and localisation elasticities simultaneously due to collinearity (Graham and Gibbons, 2019).

2.6 Time lag before agglomeration benefits appear

The literature on agglomeration distinguishes between static and dynamic agglomeration mechanisms (see e.g. de la Roca and Puga (2017)). This should not be confused with the TAG terms of static and dynamic clustering. Static agglomeration mechanisms occur instantaneously. For example, matching of workers and firms. When a worker moves to a different (larger) city and is better matched to a job, their productivity increases at the time of their move. Dynamic mechanisms on the other hand take time to feed through. Learning mechanisms are an example. A worker will continue to pick up skills and improve their productivity for many years after moving between cities. The limited evidence that is available suggests that dynamic agglomeration mechanisms may be around 50% of the total agglomeration impact and take up to ten years to fully appear following a worker's move between cities (de la Roca and Puga, 2017, Carlsen et al., 2016b, D'Costa and Overman, 2014).

If we think that a worker's move (i.e. a land use change) will only be triggered following a transport investment, then an additional friction on the formation of agglomeration economies will occur if land use change is not instantaneous. In congested cities there is evidence from the induced traffic literature that land use changes continue for at least five years up until the point that the roads are fully congested again (Duranton and Turner, 2011, Hymel, 2019). However, in less congested locations land use change may continue for up to fifteen years (Tveter et al., 2017).

Considering the frictions caused by the dynamic agglomeration mechanisms and those caused by the frictions on land use change, may imply that agglomeration impacts may take between ten and fifteen years to fully appear in an economy following a transport investment. This has an implication on firstly the ability to detect agglomeration economies ex post, and secondly may impact on the cost benefit analysis. With respect to the cost benefit analysis Tveter and Laird (2018) found that the impact is very dependent on the discount rate and the appraisal period. For high growth countries with high discount rates (e.g. 7%) then the estimated benefits may be more than twice what they should be. For countries with low discount rates and long appraisal periods, such as the UK, then the error is more like a 10% overestimation. Within the context of an overall transport appraisal in the UK this is probably liveable – given agglomeration benefits are in the region of between 20 and 30% of total benefits (Tveter, 2020) which would imply an error of the order of 2–3% in total economic benefits.

2.7 Consumption externalities

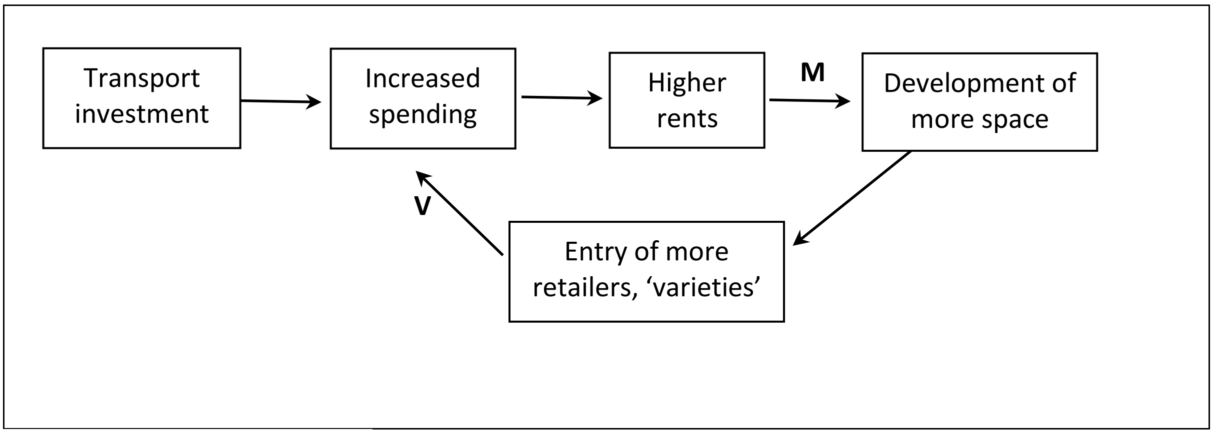
In addition to productive benefits cities are also considered to offer consumption benefits for households. In particular, cities offer benefits of variety in non-tradeable goods such as restaurants, theatres and even an attractive mix of social partners. Successful cities also offer households other benefits: high aesthetics and physical setting, good public services and easy access to a range of services and jobs (Glaeser et al., 2001).

We can think that transport investment can impact on two of these four attributes: the variety of non-tradeable goods, and increasing the range of services and jobs available (through greater speed giving greater accessibility). We therefore concentrate on the former here (the benefits of increased variety), and why they could be considered as consumption externalities and additional to transport user benefits. The second of these is closely related to productivity benefits of agglomeration through the labour market matching mechanisms.

When transport stimulates changes in land use and economic growth in a city, this will lead to new retail, office, housing, etc. developments. The benefits from these developments only exceed the user-benefits in certain circumstances. Laird and Venables (2017) offer a conceptualisation of this. The context they use is that of a retail development, but the arguments are more general. With reference to Figure 2-3, a transport improvement increases spending in a place, as visits respond to lower travel costs and footfall increases. Increased spending raises profitability, which incentivises the development of more space, redeveloping the site – by extension, or perhaps by building taller. Associated with this is the entry of more shops, in turn making the place a more attractive destination, thereby creating the feedback loop which leads to a second round of higher rents, etc.

User-benefits trigger this process, and wider-benefits arise if (and only if) there are interactions with market failures. There are, arguably, two sources of market failure in this process, labelled M and V. The first, M, arises as there may be barriers preventing the level of development reaching an efficient level (or conversely leading to excessive development above efficient levels) and hence creating gaps between marginal benefits and costs. The second is at point V, and captures the idea that places become more attractive as they attract more stores.

FIGURE 2-3: LAND USE CHANGE IN A COMMERCIAL DEVELOPMENT



Source: Laird and Venables (2017)

Whilst barriers to entry (M) are clearly a source of market failure, there is no particular reason to consider that they are systematically related to agglomerations other than indirectly via forms of governance or the coordination failure argument made earlier. However, the attractiveness (V) argument is clearly related to agglomeration size, as larger cities can support a larger non-tradeable sector and therefore offer more variety.

The attractiveness argument requires that entry of new stores creates some consumer benefit over and above the value of their spending. [11] This will arise if stores are differentiated from each other,

[11] This can also be re-stated as a consumer variety effects that trigger the second (and subsequent) rounds of land use change, which mean that the rule of half for transport user benefits 'fails' – in that it does not capture all the benefits. BATES, J. 2006. Economic evaluation and transport modelling: theory and practice. . *In:* AXHAUSEN, K. (ed.) *Moving Through Nets: The Physical and Social Dimensions of Travel. .* Oxford: Pergamon.

and is a variety effect. If there is no variety effect and new shops (products) are perfect substitutes then there is no additionality. Laird and Venables (2017) draw from international trade literature to argue that this variety effect could create a wider benefit mark up of 10–20% of expenditure in the development, though caution that this is likely to be context specific. Ahlfeldt et al. (2015) estimate an elasticity of residential externalities (amenities) of 0.15 for the case of Berlin, this is double the productivity elasticity estimated. In their synthesis Ahlfeldt and Pietrostefani (2019 p104) consider an elasticity of 0.12 for the pure gains from variety to be reasonable, though this is based on very limited studies. For a 1% increase in density they estimate this to give a welfare benefit that is a third of the productivity gain on wages.[12] An alternative approach for estimating this variety benefit can be through the trip distribution component of a multi-stage model, see for example Geurs et al. (2010) for an application. They call this a destination utility effect, where changes in population and employment lead to changes in the attractiveness of destinations in the model. The use of the trip distribution model would also likely have the added advantage of being able to capture the increased range to existing services, which is Glaeser's fourth attribute of successful cities mentioned above.

These arguments are generalisable from the retail sector to other non-tradeable sectors: theatres, cafes, hairdressers, etc. They are also generalisable to industrial sectors: manufacturing, business services (office development), etc. However, in this latter case the variety effect is a re-statement of the agglomeration (productivity) effect. To avoid double counting between wider economic impacts therefore only the consumer benefits from increased variety (Glaeser et al., 2001, Glaeser and Gottlieb, 2006) should be included in the transport appraisal in addition to agglomeration (productivity) benefits. Even then care would need to be undertaken to avoid double counting of other benefits included in the transport appraisal and ensuring only the consumer variety impacts are included.

In addition to the need of careful accounting to ensure there is no duplication with productivity related agglomeration economies, Laird and Venables (2017) add a cautionary note that when land uses change these benefits have to be placed in the context of displacement effects. Would the activity – retail, manufacturing, commercial or residential – take place somewhere else, absent the transport improvement? If so, is it subject to the same market failures? Effects across all geographical areas then have to be combined – some of them positive, and others negative. This of course is also true should productivity agglomeration benefits be partially driven by land use change.

2.8 Beneficiaries of agglomeration

Ultimately the beneficiaries of agglomeration are households, as households own businesses and government also acts on their behalf. However, this aggregate persepctive disguises that the incidence of benefits (and costs) from agglomeration may vary by household, by business and can also take different forms: income related, cost of living, transport related, environment related, etc. In their synthesis of the literature Ahlfeldt and Pietrostefani (2019) categorised the impacts of density (agglomeration) into fifteen categories. We reproduce these below in Table 2-1. They go on to provide elasticities to density for each of these categories, based on either a meta-analysis, their own analysis or a single high quality paper (also in Table 2-1).

[12] They subsequently adjust this downwards slightly to reflect potential double counting with vehicle miles travelled.

The categories in Table 2-1 are phrased in terms of perceived benefits. Thus, vehicle miles travelled reduction is a benefit. Reducing pollution is a benefit, as is reducing crime. Positive elasticities would therefore be seen as benefits from agglomeration. As can be seen most metrics improve with density. Not all do though. Social equity worsens, as does pollution, traffic speeds, mortality and reported well being. It's important to recognise the embryonic nature of some of these quantitative estimates. Not all are based on a range of studies. Ahlfeldt and Pietrostefani consider that (category 1) wage and productivity, (2) rent, (4) vehicle miles travelled (10) pollution and (12) average speed to be sufficiently well estimated to have a true causal interpretation. They consider that the other elasticities are best interpreted as representing associations in the data, and that these are areas for further research.

Now turning to productivity and wages. Whilst the average wage elasticity is 0.04 in Table 2-1, we know that the impact of density on wages varies by industry type (Melo et al., 2009, Graham et al., 2010), with business services having one of the largest elasticities associated with it. Therefore this average disguises inter-industry variation. Wages also systematically vary with industry with higher income workers in business services. Thus increasing density creates more benefits for high income workers than lower income workers. This is due to the fact that the worker productivity premium from agglomeration increases with education level (Behrens and Robert-Nicoud, 2015). By implication this means that increasing density increases wage inequality, which is why social equity (category 7) diminishes as density and wages on average increase.

Rents increase with density. This has distributional impacts between homeowners and renters. Homeowners gain from increases in density. Renters have to pay higher rents, but gain higher amenities from the higher density. The impact on them is therefore ambiguous. If we think that on average homeowners tend to be wealthier than renters or from a different demographic profile (young adults versus families), then we can see the impact of increasing density to have a distributional impact between these groups.

TABLE 2-1: IMPACTS OF AGGLOMERATION

Category		Outcome	Elasticity	Currently included in TAG
1	Productivity	Wage	0.04	Yes: Agglomeration
2	Innovation	Patent intensity	0.21	Yes: Agglomeration
3	Value of space	Rent	0.15	Yes: User benefits and agglomeration
4	Household car travel	Vehicle miles travelled (VMT) reduction	0.06	Yes: User benefits
5	Services access	Variety value (price index reduction)	0.12	No: consumption externality
6	Effect of public services	Local public spending	0.17	Yes: if productivity agglomeration includes the public sector
7	Social equity	Inter-quintile wage gap reduction	-0.035	No
8	Safety	Crime rate reduction	0.085	Partially, in a non-monetised way
9	Open space	Green density	0.28	Partially - some landscape values in TAG.
10	Pollution reduction	Pollution reduction	-0.13	Yes. Environmental impacts.
11	Energy efficiency	Energy use reduction	0.07	Yes. Reduced car VMT
12	Traffic flow	Average speed	-0.12	Yes. User benefits.
13	Mode choice	Car use reduction	0.05	Yes. User benefits
14	Health	Mortality rate reduction	-0.09	Partially. Road safety.
15	Well being	Self reported well being	-0.004	No. In the sense that wellbeing using a different paradigm, though utility is incapsulated by the above impacts.

Source: Ahlfeldt and Pietrostefani (2019) (columns 1 to 4) and authors' interpretation (column 5)

It is also worth drawing out that if a transport project increases density then a number of these density benefits are likely to double count those benefits already included in a transport appraisal, if not fully then partially. Very likely the only truly additional impacts of agglomeration that are not currently captured in a TAG transport appraisal including wider impacts would be (category 5) the value of variety. To a greater or lesser extent TAG captures the other benefits. Ahlfeldt and Pietrostefani estimate that the monetary value of gains in variety (the consumption externalities) is $115 per capita off a total benefit of $627 per capita, so about 18%, for a 1% increase in density.

There was a recent attempt to give a net value for the amenity benefits of cities in the north of England for the National Infrastructure Commission (AitBihiOuali, 2022). Putting aside some of the empirical challenges of the model estimation, the amenity values found include the impacts in Table 2-1 aside from wages and rents. As illustrated in that table there will be some double counting between this amenity value and the benefits in the transport appraisal, but the Ahlfeldt and Pietrostefani analysis would suggest that the majority of the benefit comes from the variety effect (consumption externalities). Consumption externalities would be additional to other benefits in a transport appraisal.

Of course in a real world with taxation and imperfect competition it will not just be households who gain. Eliasson and Fosgerau (2019) identify that the benefits of agglomeration will be split between workers (as higher wages) and government as increased taxation. There will also be an increase in returns on capital which in competitive markets will be bid away as reduced prices to consumers. Consumers (i.e. households) would therefore gain. As richer households typically consume more, then arguably richer households would gain the most from these price reductions.[13] If there is some imperfect competition then there will be some additional 'excess' profits experienced by businesses. In the UK profits are taxed, so these additional profits will be split between businesses and government. These additional business returns on capital would ultimately fall to shareholders as increased dividends. In a closed economy these would be all paid out to UK residents (either via direct ownership of small businesses or shareholdings in larger businesses or indirectly via investment or pension funds). In an open economy with overseas investors, there will however be some leakage of these benefits overseas.

To summarise, we can see that the beneficiaries from agglomeration are principally households (via increased wages, reduced consumer prices and other amenity benefits), and government through increased taxation. In imperfectly competitive markets business profits will also increase. These would be passed on to households as increased dividends, some of this may incur leakage overseas. Households will not all benefit equally. It is likely the richer households (with workers who are more likely to benefit from agglomeration economies, own their own home and consume more) gaining the most in nominal terms. Wage, amenity and rental impacts will be experienced locally in the city, but reductions in prices in tradeable goods could benefit households far beyond a city's boundaries.

[13] Such arguments also apply to price reductions driven by user benefits that households also gain from.

3 PUBLIC SECTOR ELASTICITY

3.1 Agglomeration economies in the public sector

Productivity benefits in the public sector are not included in the current TAG guidelines. The main reason for this is that the set of parameters is taken from the study of Graham et al. (2010), which are based on firm data and therefore exclude the public sector. However, the micro-mechanisms usually used as the explanation of agglomeration economies of transport improvements are conceptually just as valid for the public sector. The problem is that public sector production usually is not market based, i.e., has no observable prices, making the standard approach of examining agglomeration benefits unapplicable.

Our examination of the agglomeration economies of the public sector is informed by indirect evidence only. One reason for the lack of direct evidence is that wages to a lesser degree reflect the marginal labour productivity. The reason could be political interfering, a missing market for the output, or both. Another reason is that different productivity indicators at the firm level to a lesser degree reflect productivity differences. The standard approach in the literature of estimating the relationship between productivity and concepts of agglomeration, such as ATEM, is therefore to a lesser degree applicable. Still, we believe that the literature that examines how the cost of providing public services at the local level, and a closer look at a selection of estimated agglomeration elasticities will provide a second-best approach to understand how agglomeration affects the public sector.

In this chapter, we examine the literature that sheds some light on possible agglomeration benefits in the public sector. First, we review the empirical literature that examines agglomeration economies in the public sector. Second, we take a fresh look at recent estimates of agglomeration economies that include or exclude the public sector , and try to back out the likely size of the public sector agglomeration elasticity.

3.2 Empirical evidence of agglomeration economies in the public sector

Table 3-1 displays an overview of the papers we have found when searching the literature. We have only included papers that, in our opinion, say something about scale or agglomeration economies in the public sector. The table distinguishes between outcome (cost or quality), sector (fire and rescue, education, hospitals, or total local services), independent variable (mainly size or density), estimation technique, and finding.

A conceptual challenge when interpreting these findings is whether the possible benefits of agglomeration result in cost savings or better services (quality). Both possibilities are possible. One may examine this issue by examining the properties of the production function in the public sector. Another possibility is to look at the managerial incentives in public sector when there is a productivity uplift together with price elasticity of the public good. If the result mainly gives a cost reduction of public services the effects is in principle monetized, while if the benefits are harvested as better quality they are close to the amenities benefits of agglomeration discussed in Glaeser et al. (2001) and discussed earlier in Section 2.8. However, in both cases the benefits are additional to the standard user benefits.

TABLE 3-1 EMPIRICAL EVIDENCE OF AGGLOMERATION ECONOMIES IN THE PUBLIC SECTOR

Study	Sector	Outcome	Independent variable	Unit	Estimation	Agglomeration economies?
Chakraborty et al. (2000)	Education	Cost	Students in district	Expenditure per student in school districts	Panel data	Yes
Kirjavainen and Loikkanen (1998)		Quality	Categorical (urban, rural, dens)	Performance indicator of Finnish senior secondary schools	2nd stage DEA regression	None
Gibbons and Silva (2008)			Urban density	Pupil test score in UK	Panel fixed effects regression	Yes
Holmgren and Weinholt (2016)	Fire and rescue services	Cost	Population	Expenditure per firefighter in Swedish municipalities	SFA (descriptive)	Yes
Duncombe and Yinger (1993)		Cost and Quality	Labour share	Translog production function of US fire departments	2SLS	IRC to quality/ CRS to population
Friedson and Li (2015)	Hospitals	Cost	Population density	Price of intermediate medical services in US hospitals	2SLS with panel data	Yes
Bates and Santerre (2005)		Quality	Hospitals per capita	Inpatient days per bed in US metropolitan area	Cross-sectional regression on differenced data	Yes
Cohen and Paul (2008)			Proximity to other hospitals	Inpatients and outpatients US hospitals	Flexible cost function	Yes
Byrnes and Dollery (2002)	Local public services	Cost	NA	NA	Literature review	Mixed evidence
Carruthers and Ulfarsson (2003)			Population density	Expenditures per capita in US metropolitan areas	Cross-sectional regression	Yes (most services)
Bönisch et al. (2011)			Population density	Expenditure of German municipalities	2nd stage DEA regression	Yes
Büttner et al. (2004)			Population density	Expenditures per capita in German states	Cross-sectional regression	No
Hortas-Rico and Solé-Ollé (2010)			Urbanized land per capita	Expenditures per capita in Spanish municipalities	Cross-sectional regression	Yes (most services)
Soukopová et al. (2014)			Population size	Expenditure per capita in Czech Rep. municipalities	Cross-sectional regression	No (only education)
Ladd (1994)			Change in population	Change in expenditure per capita in large US counties	Cross-sectional regression	Yes
Matějová et al. (2017)			Population size	Expenditure per capita in Czech municipalities	Cross-sectional regression	Mixed
Prieto et al. (2015)			Population size and number of dwellings	Expenditures per capita in Spain municipalities of basic infrastructure*	Translog production function equation and SURE estimation	Yes
Hauner (2008)		Quality	Population density	Three measures of public sector performance in Russian regions	2nd stage Cross-sectional regression	Only for health services

DEA – Data Envelopment analysis, SFA – Stochastic Frontier Analysis
* Water supply, sewerage and cleansing of residual waters, and paving and lighting.

The table includes 18 different studies. 12 consider cost output, 5 quality, and 1 both. Regarding sectors, 10 studies consider all or several local public services, three focus on education, three on hospitals, and two on fire and rescue services. Most studies use rather aggregated data at the municipality level or similar. However, a few studies use more detailed data, such as hospitals, schools, fire departments, or pupils. 9 of the studies use accessibility type measures in their estimation strategy, which in our opinion best captures the concept of agglomeration. We also regard the estimation strategies using panel data estimation, two stage least square (2SLS), or both as preferable. Below we review these studies grouped by sector.

Education

Three studies consider agglomeration economies in the education sector (Kirjavainen and Loikkanen, 1998, Gibbons and Silva, 2008, Chakraborty et al., 2000). Kirjavainen and Loikkanen (1998) examines efficiency of Finnish senior secondary schools using date envelopment analysis (DEA). The study reports a positive association between productivity and school size, although the association is rather weak. However, no statistical effects are found between urban, density, or rural areas. This implies no evidence of agglomeration economies. Chakraborty et al. (2000) investigates scale economies of public education using panel data from Utah school districts and a cost function. This study reports scale economies from district size using fixed effects panel data estimation. Gibbons and Silva (2008) explore the association between urban density and pupil attainment using pupils in England. They find small, but significant benefits from education in schools in densely urbanised settings. We put most weight on the two latter studies because they apply both the preferable indicator of agglomeration economies and a good estimation strategy (panel data estimation is regarded as better than cross-sectional analysis). Although there is only one example of each, this shows evidence of agglomeration economies in the education sector both for costs and quality.

Fire and rescue services

Two studies examine the efficiency of fire and rescue services (FRS) in relation to agglomeration. In an early contribution, Duncombe and Yinger (1993) uses a translog production function to estimate productivity in fire protection. The study reports constant returns to population size. In the second study, Holmgren and Weinholt (2016) study the efficiency of the Swedish fire protection service using DEA. They report that the cost of spending of FRS is clearly lower in a large sized municipality. But at the same time, they show that the largest cities have more workers in FRS and a shorter response time. Although the evidence is mixed and thin, there is some support of scale economies from city size in fire protection and rescue services.

Hospitals

Three studies examine agglomeration economies of hospitals. Bates and Santerre (2005) examine agglomeration economies in the hospital service industry by examining how inpatient days per bed at the US metropolitan level is affected by hospital clustering. They find that inpatient days per bed are lower when the hospital density increases. Interpreting inpatient days per bed as a quality indicator, this result implies that hospital clustering results in a better health service. The authors hypothesise that the mechanism from hospital clustering to better health services runs through a faster spread of knowledge in denser areas.[14] They also discuss a possible utilization of the sharing mechanisms through capacity utilization. Cohen and Paul (2008) uses more detailed data with hospital-level data and a cost function. Agglomeration is included as the proximity to other hospitals. This study also reports agglomeration economies, which is explained by benefits from knowledge sharing through proximity to other hospitals. Friedson and Li (2015) examine the extent agglomeration in the hospital industry enhances productivity. They focus on a particular mechanism by looking at how concentrated hospital services increase the specialization of intermediate medical labs. They find that this specialization reduces the price of these services. They interpret this effects as benefits working through competition and specialization.

Local public services in general

The largest group of evidence are the studies that examine several or all local services together using aggregated data at the level of municipality, metropolitan area, state, or region. All studies except one consider monetized effects. This strand of the literature is older and was initiated in the late 1950s. The typical approach is to consider per-capita expenditure of different services as the dependent variable. The early literature (from the 1950s until 2001) is reviewed in Byrnes and Dollery (2002) and they find mixed results of scale economies. This result holds both when examining the total cost of providing public sector services and when looking at different types of services, such as water, public libraries, public transportation, and education. Below, we review the more recent literature. We start with the studies using accessibility measures to identify the effect, and next the studies that consider indicators of city size.

The studies that consider density measures report mixed evidences. Carruthers and Ulfarsson (2003) examines how the expenditures per capita in 283 US metropolitan counties vary with density considering 12 different measures of public expenditure. They find that the per capita cost of services declines with density. Bönisch et al. (2011) examine efficiency in German municipalities using DEA. They also run a second stage regression after their efficiency analysis and include population density as an explanatory factor. This factor is estimated to slightly improve efficiency, which implies agglomeration economies. These findings are, however, inconsistent with the results of an earlier analysis of German municipalities (Büttner et al., 2004) which finds no effect on the cost of public goods from differences in population density, when estimating a range of different sectors separately. This study finds, however, the there is a cost-disadvantage for the smallest states.

The five studies that use population size also report mixed evidence. Hortas-Rico and Solé-Ollé (2010) examine the relationship between urban sprawl and the costs of providing local public services. They use data of 2,500 Spanish municipalities and estimate a cross-sectional analysis. They find that low-density areas have a higher provision cost in most spending categories. The exception is housing , basic infrastructure, and transport. Prieto et al. (2015) investigates economies of scale in association to population and housing in urban areas of providing the basic infrastructure water, sewage, cleansing of water, paving and lightning. The find scale economies in the cost of providing basic infrastructure, which is reinforced when the density increases. Ladd (1994) examines how population growth is associated with per capita spending in large US counties. In a cross-sectional regression analysis, higher density is associated with reduced per capita spending, but only in sparsely populated counties. Soukopová et al. (2014) examine returns to scale in the Czech Republic,

[14] The authors control for a range of factors in the analysis. The controls consist of population size, staff per bed, salary, physicians per bed, nurses per bed, health insurance, trainees per bed, per capita income, and average hospital size.

considering the provision of the services: sports and leisure activities; culture, church, and media; environmental protection, housing, municipal services and development; and education. They use a log-linear regression model and trend comparison for 205 municipalities. The study finds little evidence of increasing return to scale in the production of local services. Similar results are found in Matějová et al. (2017). They also find scale economies in collecting local fees, pre-school, and elementary education, but not for local administration using a cross-sectional regression analysis. Also, this study argues that there are inefficiencies for the smallest municipalities.

Only Hauner (2008) examines overall productivity effect of the public sector productivity using quality outcomes. This study considers how public sector performance in Russia varies with population density. Two of the quality indicators are based on indicators such as infant mortality, education, and poverty, while the last is based on efficiency performance from an DEA analysis. The findings suggest that higher population density improves public sector performance in health, but no robust effects are identified on education, social protection, and social sectors.

Summary

Although the evidence is mixed and our review gives no complete picture, some implications can be drawn. First, there seems to be a varying degree of scale economies when providing municipal services. The reason for this could be that only some types of public sector services have scale economies (education, fire and rescue services and hospitals), while other do not. Another reason could be that scale economies are only relevant up a point, that is it is only small municipalities that benefit of becoming larger. Given this interpretation one should not assume that there is a general return to scale in all forms of public sector production. Such an effect appears to be sector specific, but also quite relevant in remote areas with a small population size.

3.3 Private and public sector agglomeration elasticities

Another approach to explore the public sector agglomeration elasticity is to review the literature and try to extract some indirect evidence from it. We follow this approach by focusing on studies that have used state-of-the-art methods in recent years, and where public sector activities are in their sample. Based on this strategy we identified six studies (see Table 3-2).

TABLE 3-2 INDIRECT EVIDENCE OF PUBLIC SECTOR ELASTICITY

	Outcome	Model	Accessibility indicator	Including public sector	Country	Elasticity (subsample)
Börjesson et al. (2019)	Wages	Panel data estimation	ATEM with exponential decay	Yes	Sweden	0.028 (movers)
Carlsen et al. (2016a)	Wages	Panel data estimation	City size	No	Norway	0.03–0.04 (movers)
D'Costa and Overman (2014)	Wages	Panel data estimation	City size	Yes	UK	0.023 (full sample) 0.021 (only private)
Knudsen et al. (2022)	Wages	Panel data estimation	ATEM with exponential decay	No	Denmark	0.027 (movers)
Roca and Puga (2017)	Wages	Instrumental variable panel data estimation	City size	No	Spain	0.020 (short-term) 0.053 (medium term)

Maré and Graham (2013)	Gross output	Panel data and a translog production function	ATEM with power decay	Yes (but how?)	New Zealand	0.049 (full sample) 0.051 (education and health & community services)*

*Weighted average of elasticities in the sectors education; and health and community services.

Carlsen et al. (2016a) estimates the city wage premium using a panel of Norwegian workers. The primary focus in this paper is to examine how the return to education affects the urban wage premium. In their instrumental variable (IV) estimation they find an initial impact on wages between 0.03 and 0.04 in the short run, and a medium premium between 0.04 to 0.05. This study excludes the public sector sectoral workers arguing that "wages are determined by national regulation, public sector workers" (p. 41). The study could, however, serve as a control study for the other studies that includes workers in the public sector.

The D'Costa and Overman (2014) uses a similar strategy as Carlsen et al. (2016a). Including workers from all sectors, they report an effect of city size of 0.023 and cannot find any difference between the initial and medium run effect. This study includes workers in the public sector. When they restrict their sample to only private sector workers the elasticity falls slightly to 0.021. Regarding differences between private and public sector workers this difference is consistent with a *higher* elasticity for public sector workers than the workers in the private sector. The difference is however too small to conclude that there is any significant difference. It does, however, at least indicate that the benefits of agglomeration are not any lower in the public sector than the private sector. It is however possible to argue that wages are set in the private sector and that the local level of wages in the public sector is set in a second stage of the wage bargaining process.

Both Knudsen et al. (2022) and Börjesson et al. (2019) estimate the effect on wages using ATEM functions with exponential decay using data at the individual level. The first study uses data from Denmark, and the latter from Sweden. Although the Swedish study includes all workers in their sample and the Danish study excludes public sector workers (and agricultural workers) their estimated elasticity is rather similar at between 0.026 and 0.028.

Roca and Puga (2017) consider the wage premium in Spain and attempt to identify a dynamic impact on the wage premium. They find a short run effect of wages pf 2 pct. and a medium run effect of 5 pct. This study excludes the workers in the public sector using the same arguments as Knudsen et al. (2022) and Carlsen et al. (2016a) that these wages are not market based. But the size of the estimates are comparable to D'Costa and Overman (2014) that includes the public sector.

Our final paper Maré and Graham (2013), examines sector heterogeneity using firm data from New Zealand and a standard ATEM formulation of accessibility. We do not understand how they manage to include public sector firms, but from their estimation results we observe that the estimated agglomeration elasticity for sectors that produce similar services to in the public sector (education and health and community services) have elasticities close to the economy weighted average. Hence, also in this study the results indicates that the observable benefit of agglomeration is similar in the public sector and other part of the economy.

Our review of elasticities of the public sector is unfortunately only based on indirect evidence. Based on the limited available indirect evidence we find no evidence to support the assumption that the effect in the public sector is zero, as implicitly assumed in the DfT framework. The available evidence is too thin to make any assertion of a distinct agglomeration elasticity for the public sector. If anything, our review supports the hypothesis that the effect in the public sector is closer to the economy weighted average, than it is to zero. The inclusion of a public sector elasticity in the productivity agglomeration calculations has an implication for the treatment of some amenity benefits (see Section 2.8), and there will be a need to avoid potential double counting.

4 ATEM

4.1 B2B and B2C, generalised cost and time periods

Agglomeration economies are dependent on interactions, between businesses (B2B) and between businesses and consumers (B2C). Between business interactions (B2B) are clearly important for creating new contracts in the supply chain (to help share the gains from variety) and for the generation and diffusion of knowledge. These sorts of activities might be associated with employers' business person-trips. The sharing of physical inputs is often associated with truck transport, particularly where inventory levels are important (Rosenthal and Strange, 2020). These might therefore be associated with freight trips, though the contracts that created the supply chain linkages might be dependent on person-trips as just mentioned.

Interactions between businesses and workers (B2C) are clearly relevant for labour market matching effects on quality, chances and mitigating hold-up problems. Here commuter trips would be relevant, though in this era of work-from-home, it would be necessary to factor in the number of work from home days in any commuting cost analysis (Laird and Tveter, 2021). However, B2B transport costs are also relevant for matching effects, as businesses need to share labour pools. Businesses can only do so if they are proximate to each other. Commuting costs between homes and workplaces are not a good indicator for whether businesses are sharing a labour pool, some other measure e.g. business to business transport costs, would be necessary for that.

If we think specifically about transport cost data then we might map the different mechanisms on to likely different trip types. Labour market matching mechanisms would likely map onto peak hour costs, person-business travel on to inter-peak transport costs and freight onto some mix of inter-peak and off-peak[15]. This is of course imperfect, as commuting can occur at any time of day, although it peaks in the peak, with person business trips and freight trips likely being the opposite – more intense between the peaks than in the peaks.

Turning to consumption externalities, the relevant trips associated with these would be non-work-other (in the TAG definition). These would be unlikely to occur in peak hours, but would be relevant in the inter-peak and the off-peak. Such trips would not seem overly relevant to the interactions that drive productivity related agglomeration economies.

The relevance of both B2B and B2C would suggest both should appear in ATEM with associated different elasticities and decay parameters. However, B2B and B2C measures are so highly correlated that it would be difficult to empirically separate out the two measures in an estimation (Graham and Gibbons, 2019).[16] See KPMG (2013) for an example of the empirical challenges of mixing ATEM functions. Thus, in practice, whilst not ideal, a choice between the two mass measures

---

[15] Inter-peak is during the working day between say 9am and 4pm, off-peak is evenings, night time and weekends.

[16] Graham and Gibbons (2019) compare ATEM functions calculated using an employment (B2B) and population (C2C) functions for Great Britain. They find them highly correlated (correlation coefficient of 0.95). They do not directly compare a B2B with a B2C function, but given employment density in an employment zone (B2B) will be a function of the number of workers that can access that zone (i.e. B2C) we would expect them to be very highly correlated.

invariably has to be made in empirical work, with a preference for B2B as that is related to more of the microfoundations to agglomeration economies than B2C.

We can conclude that there is no single time period or trip purpose that captures agglomeration economies. It is a blend. From a theoretical perspective, our view is that there is no obvious guide as to how transport costs across time periods and journey purposes should be combined. It is clear though that business to business person trips are likely to be very important. Non-work other trips are likely to not be relevant, and freight and commuting trips sit somewhere in-between. In TAG generalised cost formulations exclude freight and use a trip weighted combination of business and commuting trips across time periods. This has been based on judgement. Whilst there is no evidence to disprove, there is also no empirical evidence to support it. Ideally this should be tested empirically. As it is it remains an aspect for further research.

4.2 The role of modes in ATEM

As discussed previously the majority of the empirical agglomeration literature uses ATEM variables that do not have a travel cost dimension, e.g. city size or administrative areas. A subset of the studies have looked at how travel time affects agglomeration (e.g. Rice et al. (2006) in the UK) and how distance affects agglomeration (e.g. Graham (2007) and Graham et al. (2010) in the UK). As far as we are aware only D'Costa et al. (2013) has successfully explored the different contributions of different transport modes to agglomeration. They estimated separate modal elasticities for agglomeration economies. Alternatively, one might think that the different modes combine to form the ATEM. This is the DfT approach. They sum road and public transport economic mass to get a combined ATEM.

Looking at the evidence on clustering we see that clustering occurs at many different scales: very local (even between floors in high rise buildings), at the city level and at the regional level. Different modes are likely to be relevant at these different scales. At the neighbourhood level walking and cycling are likely to be very important, walking in particular. Meetings at coffee shops mid-way between office buildings (Atkin et al., 2022) might be best achieved on foot, rather than by car or public transport. The role of walking in the ATEM function has hardly received any attention in the literature with Rohani and Lawrence (2017) making the only attempt we aware of. Their exploratory empirical work shows relationships between economic density as measured using pedestrian networks and productivity. Their work does suffer from a difficulty in obtaining disaggregate productivity data.

A complete ATEM function that would aim to explain these close forms of clustering should therefore take account of proximity by all modes. This would of course require that motorised transport should also include access/egress costs by active modes. Whether different ATEM functions by mode should be used (as per D'Costa et al. (2013)) or some weighted average or a logsum combination of modal generalised costs would seem to be an empirical matter.

From the TAG perspective the current approach, which ignores active modes, seems appropriate in its ability to capture clustering at a very close level. Having said that the TAG approach advocates applying ATEM at a local authority district level, so it is not really capable of picking up these small clustering impacts anyway, so in that sense it is mutually self-consistent.[17] We turn to the spatial scale of the ATEM function and the agglomeration benefit calculations next.

4.3 Zonal aggregation

The current guidance in TAG requires that the agglomeration benefit calculations are undertaken at a Local Authority District (LAD) level. Average employment in a LAD is 86,710 (in 2021)[18], with the largest at 683,535. In many ways this spatial treatment of economic mass is consistent with a lot of the regional economics approach to measuring the agglomeration premium. For example, de la Roca and Puga (2017) use city size as a variable in their analysis of the impact of learning mechanisms in agglomeration economies. It is also consistent with the availability of some economic data (e.g. average GDP/worker).

However, this 'aggregate' treatment of economic mass causes a number of conceptual and analytical challenges with respect to transport appraisal. Conceptually, many of the agglomeration forces act at a neighbourhood level. We see clustering at a much finer level than the LAD spatial level (e.g. in a CBD). Clearly some micro-mechanisms act at a metropolitan level (labour market type matching effects) and as argued earlier, some will act over longer distances between LADs, but we expect a substantial component of the micro-mechanisms to be acting at a much smaller spatial scale. Empirically this is supported by the decay parameters, particularly those associated with producer services. The latter tends to dominate the agglomeration benefits from transport projects. Transport projects, due to their size, also tend to act primarily at a sub-LAD level, particularly urban investments. Even National Highways, the custodian of the strategic road network in England typically delivers projects that are 'spatially small'.[19]

Analytically calculating agglomeration benefits at an LAD level also poses some challenges. Invariably the transport cost data will need aggregating to an LAD level. This creates the difficulty of calculating intra-zonal travel costs for large zones, but importantly leads to the analysis being subject to the Modifiable Area Unit Problem (MAUP). The latter is a known phenomenon associated with spatial data. It stems from changes introduced in the correlation structure of the data through aggregation. It is considered to be persistent, unpredictable and cannot be eliminated (see Gotway and Young (2002) for a review). It can however be minimised. Tveter et al. (2022) argue that to minimise MAUP in agglomeration benefits the benefit calculation should be undertaken at as small a spatial scale as possible. If aggregation must occur then travel cost data should be aggregated using the harmonic mean. They find that the use of aggregation and non-optimal aggregation techniques can lead to errors in agglomeration benefit calculations of up to 300%. The guidance in TAG to aggregate to LAD level is likely to lead to agglomeration benefit results susceptible to MAUP. The

---

[17] If walking and cycling had been included this would have increased the level of ATEM. Given that the remain constant after a reduction in GC between LADs this implies that the effects are upward biased.

[18] Source: DfT (2021) TAG Wider Impacts Dataset

[19] Between 2015 and 2020 Highways England delivered 370 lane miles from 36 projects (source: Highways England 2020-25 Delivery Plan). If these are all 3-lane motorways (6 lanes in total) then the average project length is 1.7 miles (=370/36/6). If they are all single carriageway then the average project length would be 5.1miles.

aggregation process also recommended, a trip weighted average, is also not optimal. (Tveter et al., 2022).

In summary recent thinking and research suggest that there are both conceptual, analytical and practical reasons why agglomeration benefits from transport projects should be conducted at a small spatial scale. TAG could be updated to reflect this improved knowledge base.

4.4 Functional form of ATEM

The ATEM function is of critical importance to the estimation of the agglomeration benefits of transport projects. In their survey of the literature Donovan et al. (2021) identify that 104 studies of their 294 sample use a 'potential' measure with some sort of distance or time decay (i.e. just over a third). The other studies either use a pure size or density measure (with no interaction with neighbouring zones), or time/distance isochrones. Donovan et al. (2021) do not identify the different types of decay function, and we have not surveyed the papers they refer to, however we imagine that the power decay function would dominate. This is because the power decay function is popular in trade theory and other commentators, e.g. Graham and Gibbons (2019), consider it to be one of the most popular.

Other potential options for a decay function would be the negative exponential (see e.g. Knudsen et al. (2022) used to estimate Danish agglomeration elasticities) or some form of s-shaped logistic type curve (see e.g. KPMG (2013) for example using UK data). The argument for using a different shaped decay function is that agglomeration benefits require human interactions, these require trips, and these alternative decay functions better replicate observed trip patterns. Negative exponential and logistic type functions can better represent travel demands, which is why they are suggested as alternatives to the power function.

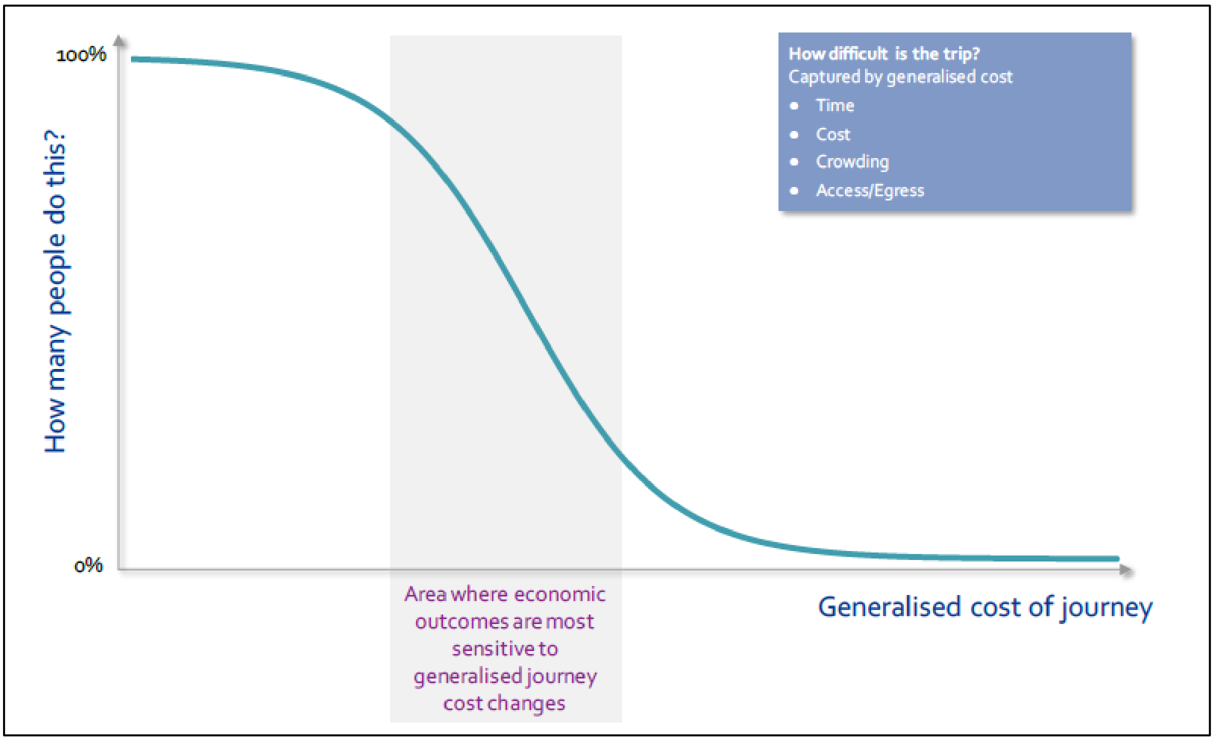
Agglomeration elasticities and distance decay parameters are inter-related. A lower decay parameter implies that the distance matters less and hence the geographical differences between zones evens out. A high decay parameter on the other hand puts more weight on the closer jobs when calculating the accessibility. As a consequence, the agglomeration elasticity is sensitive to the chosen decay parameter. Mixing and matching between elasticities and decay parameters is inappropriate. This is also the case with alternative decay functions (see for example Knudsen et al. (2022) for a practical application with the negative exponential).

The different shapes of the different decay functions also have important implications for the agglomeration benefits that are derived from a transport project. It is when a project moves zonal movements through the steep part of the decay function that the largest benefits are derived (see Figure 4-1). As the different decay curves have different shapes then it would imply that the agglomeration benefits will also be dependent on the decay curve chosen. With its rapid decay the power decay function typically favours shorter distances, whilst the negative exponential and logistic type curves may give larger benefits to zones that are further apart. This discussion has resonance with that regarding the existence or not of long distance agglomeration economies made earlier. To our knowledge the choice of decay function has not been tested empirically. Most researchers impose a decay function on their data. E.g. The TAG values derived by Graham et al. (2010) had a power decay function imposed on them, and Knudsen et al. (2022) imposed a negative exponential on their data.

When considering using a different decay function some of the pleasant properties of the power function become lost. The most important of which is its neutrality to scale. With a power decay function it does not matter whether the generalised cost impedance measure is in pence or pounds or some variation inbetween. However, this does matter with the negative exponential, with the decay parameter absorbing the scale differences. Thus the elasticity and decay parameter in a negative exponential are not only related to one another but are also related to the units of measurement (see the Appendix for an exposition). This has two important points: firstly in application the correct units have to be used, and secondly the estimation method needs to estimate the decay parameter and not just impose it on the data. The consequence of imposing it on the data, as Gibbons et al. (2019) notice this with respect to an employment elasticity to ATEM, is that the elasticity will be quite different between specifications. This is despite different versions of ATEM giving similar predictions on the percentage change in employment for a 1 standard deviation change in ATEM. It is also not clear, without further research, whether the different ATEM functions will give rise to similar or different elasticities. One would assume the latter, but how different is currently unknown. This has an important corollary, as once we move away from using a power decay function, then we cannot benchmark the agglomeration elasticity against the evidence base, such as the Melo et al. (2009) or Donovan et al. (2021) meta-analyses. A benchmarking exercise such as that used by Gibbons et al. (2019) for their employment elasticities would need to be undertaken.

The investigation into the best functional form for ATEM, remains a little studied area, but given the parallels with a number of other interest areas, such as treatment of agglomeration benefits over a wider area, it could be a fruitful area of research. At this point in time no definitive policy guidance can be offered on it, with respect to any immediate changes to TAG.

FIGURE 4-1: ILLUSTRATIVE DECAY CURVE



Source: KPMG (2013)

5 IMPLICATIONS FOR POLICY AND RESEARCH

Within the UK the empirics of agglomeration in transport appraisal have remained broadly unchanged aside from periodic updates of the parameters, primarily the GDP and employment data. The underlying equations remain as they were twenty years ago. The knowledge base has however continued to grow over this time. In places the guidance can be updated immediately to reflect improvements in this knowledge base, in other areas some further research needs to be undertaken to either understand if the topic merits further investigation or to be firm about how guidance should be updated. Other research gaps may require a more concerted effort to address. These have been classified as longer term, not because they should commence further down the line, but because the results from the research may take time before they feed into guidance.

Immediate updates to TAG and immediate research to be undertaken for implementation in TAG in the short term:

Public sector elasticity: Certain parts of the public sector show clear evidence of benefitting from external economies of scale: education, hospitals and fire and rescue services. The evidence is less clear with regard to other parts of the public sector, and there is some evidence of diminishing returns. Agglomeration elasticities also seem insensitive as to whether workers from the public sector are included or excluded from the analysis. This can be interpreted that on average the public sector is no different from the rest of the economy. There is therefore firm ground to include a public sector elasticity in appraisals. Without direct evidence an elasticity that is equivalent to the economy wide average is likely a suitable approximation to the average elasticity the public sector will experience.

Spatial level of agglomeration calculations. zonal aggregation introduces an error into the agglomeration calculations. This is known as MAUP. It is recommended that agglomeration calculations should be undertaken at as small a level as practical. Where aggregation of zones occurs, transport costs should be aggregated using the harmonic mean. Guidance may also need to be provided regarding practical situations related to when employment and GDP data is not available at an equivalent disaggregate level. This may require some exploration of the impacts of choosing different approaches before offering firm guidance.

Research to be commissioned in the short term, as a pre-cursor to further research or a TAG update in the medium term

Exchequer impacts and private benefits of agglomeration economies: a TAG appraisal identifies the distribution of impacts by mode, commuting, non-work, business and on transport providers. It also includes a distributional analysis identifying the impacts on different sections of society. Agglomeration benefits do not occur evenly across the population. With competitive markets they are split between households (via workers and reduced prices) and the exchequer (via labour market taxes). This distribution could be made explicit in the presentation of results, particularly as the distinction between exchequer impacts and private benefits is an important component of TAG. Such a presentation will have implications for the presentation of other wider impacts. It is also noted that wider impacts do not currently feature in the TEE, PA or AMCB tables, so some thought would be needed as to how to incorporate this breakdown in the reporting of the CBA, and how to treat the other wider impacts. Further research would be needed to undertake a full distributional analysis regarding the split of private benefits between 'rich' and 'poor' households.

Localisation economies and consumptions externalities (part 1): Embryonic research suggests that localisation economies and consumption externalities both seem to be significant areas of additional benefit. Potentially adding in the order of two thirds of existing agglomeration benefits to the appraisal. It is recommended this is explored further, before consideration is given to commissioning further research on parameters that could be used in TAG. Such research should include applying existing elasticities in several case studies for which wider impact appraisal data is available. This would identify whether the two-thirds of existing benefits figure implied by the embryonic research is replicated in other contexts. This research should also include a further consideration of the robustness of the elasticities being employed. The outcome of this research would then indicate whether further research on the elasticities would be beneficial, it would also indicate whether some practical options for inclusion of these benefits as sensitive tests in existing appraisals could be made in the interim.

Research to be initiated in the short term, with outcomes available in the medium term

Active travel and long distance benefits: Given the observation that clustering is evident at very local and at regional levels, with the associated interest policy interest in both urban realm and long distance connectivity, research could be undertaken exploring the role of the functional form of ATEM in the estimation of agglomeration elasticities. This would involve exploring different decay functions and means of combining generalised cost between time periods, and modes (bringing active travel and freight into the frame). Wage data is most easily available, but evidence shows that the wage premia to agglomeration dissipates over medium distances. Wage premia could likely be used to investigate the role of active travel to agglomeration benefits, but not long distance benefits. Such analysis would require a TFP measure of productivity. Active travel measures would also require walk and cycle time access metrics. These would require a very fine spatial resolution of the employment data. Such a resolution would not be necessary for an investigation of long distance benefits.

Research to be initiated in the medium term

Localisation economies and consumptions externalities (part 2): Subject to the outcome of the investigations on the impact of localisation economies and consumption externalities on a transport appraisal, consideration should be given to conducting research on parameters that would inform these calculations: namely elasticities for localisation economies and consumption externalities.

6 REFERENCES

AHLFELDT, G. M. & PIETROSTEFANI, E. 2019. The economic effects of density: A synthesis. *Journal of Urban Economics,* 111**,** 93-107.

AHLFELDT, G. M., REDDING, S. J., STURM, D. M. & WOLF, N. 2015. The economics of density: Evidence from the Berlin Wall. *Econometrica,* 83**,** 2127-2189.

AITBIHIOUALI, L. 2022 Effects of population density on the value of amenities in the UK. Evidence from the Rail Plan for the Midlands and the north of England. London: Imperial College London.

ATKIN, D., CHEN, K. & POPOV, A. 2022. The returns to face-to-face interactions: Knowledge spillovers in Silicon Valley. NBER report NO. w30147.: National Bureau of Economic Research.

BATES, J. 2006. Economic evaluation and transport modelling: theory and practice. . *In:* AXHAUSEN, K. (ed.) *Moving Through Nets: The Physical and Social Dimensions of Travel. .* Oxford: Pergamon.

BATES, L. J. & SANTERRE, R. E. 2005. Do Agglomeration Economies Exist in the Hospital Services Industry? *Eastern Economic Journal,* 31**,** 617-628.

BEHRENS, K. & ROBERT-NICOUD, F. 2015. Agglomeration theory with heterogeneous agents. *Handbook of regional and urban economics,* 5**,** 171-245.

BYRNES, J. & DOLLERY, B. 2002. Do Economies of Scale Exist in Australian Local Government? A Review of the Research Evidence. *Urban Policy and Research,* 20**,** 391-414.

BÜTTNER, T., SCHWAGER, R. & STEGARESCU, D. 2004. Agglomeration, population size, and the cost of providing public services: an empirical analysis for German states. *ZEW-Centre for European Economic Research Discussion Paper*.

BÖNISCH, P., HAUG, P., ILLY, A. & SCHREIER, L. 2011. Municipality size and efficiency of local public services: Does size matter? : IWH Discussion papers.

BÖRJESSON, M., ISACSSON, G., ANDERSSON, M. & ANDERSTIG, C. 2019. Agglomeration, productivity and the role of transport system improvements. *Economics of Transportation,* 18**,** 27-39.

CARLSEN, F., RATTSØ, J. & STOKKE, H. E. 2016a. Education, experience, and urban wage premium. *Regional Science and Urban Economics,* 60**,** 39-49.

CARLSEN, F., RATTSØ, J. & STOKKE, H. E. 2016b. Education, experience, and urban wage premium. *Regional Science and Urban Economics,* 60**,** 39-49.

CARRUTHERS, J. I. & ULFARSSON, G. F. 2003. Urban sprawl and the cost of public services. *Environment and Planning B: Planning and Design,* 30**,** 503-522.

CHAKRABORTY, K., BISWAS, B. & LEWIS, W. 2000. Economies of scale in public education: an econometric analysis. *Contemporary Economic Policy,* 18**,** 238-247.

COHEN, J. P. & PAUL, C. M. 2008. Agglomeration and cost economies for Washington state hospital services. *Regional Science and Urban Economics,* 38**,** 553-564.

COMBES, P.-P., DURANTON, G. & GOBILLON, L. 2011. The identification of agglomeration economies. *Journal of Economic Geography,* 11**,** 253-266.

COMBES, P. P. & GOBILLON, L. 2015. The empirics of agglomeration economies. *In:* DURANTON, G., HENDERSON, J. V. & STRANGE, W. (eds.) *Handbook of Regional and Urban Economics.* Elsevier.

D'COSTA, S. & OVERMAN, H. G. 2014. The urban wage growth premium: Sorting or learning? *Regional Science and Urban Economics,* 48**,** 168-179.

D'COSTA, S., GIBBONS, S., OVERMAN, H. & PELKONEN, P. 2013. Agglomeration and labour markets: the impact of transport investments on labour market outcomes. *In:* CRESCENZI, R. & PERCOCO, M. (eds.) *Geography, Institutions and Regional Economic Performance.* Springer.

DE LA ROCA, J. & PUGA, D. 2017. Learning by Working in Big Cities. *The Review of Economic Studies,* 84**,** 106-142.

DEPARTMENT FOR TRANSPORT 2005. Transport, Wider Economic Benefits, and Impacts on GDP. *Discussion paper.* London: Department for Transport (DfT).

DEPARTMENT FOR TRANSPORT 2020. TAG Unit A2-4: Appraisal of productivity impacts. *In:* DEPARTMENT FOR TRANSPORT (ed.) *Transport Appraisal Guidance (TAG).* London: Department for Transport.

DEPARTMENT FOR TRANSPORT 2021. TAG Wider Impacts Dataset. *In:* DEPARTMENT FOR TRANSPORT (ed.) *Transport Appraisal Guidance.* London: Department for Transport.

DONOVAN, S., DE GRAAFF, T., DE GROOT, H. L. & KOOPMANS, C. C. 2021. Unraveling urban advantages—A meta-analysis of agglomeration economies. *Journal of Economic Surveys*.

DUNCOMBE, W. & YINGER, J. 1993. An analysis of returns to scale in public production, with an application to fire protection. *Journal of Public Economics,* 52**,** 49-72.

DURANTON, G. & PUGA, D. 2004. Chapter 48 Micro-foundations of urban agglomeration economies. *In:* HENDERSON, J. V. & JACQUES-FRANÇOIS, T. (eds.) *Handbook of Regional and Urban Economics.* Elsevier.

DURANTON, G. & PUGA, D. 2020. The economics of urban density. *Journal of Economic Perspectives,* 34**,** 3-26.

DURANTON, G. & TURNER, M. A. 2011. The fundamental law of road congestion: Evidence from US cities. *American Economic Review,* 101**,** 2616-52.

ELIASSON, J. & FOSGERAU, M. 2019. Cost-benefit analysis of transport improvements in the presence of spillovers, matching and an income tax. *Economics of Transportation,* 18**,** 1-9.

FRIEDSON, A. I. & LI, J. 2015. The impact of agglomeration economies on hospital input prices. *Health Economics Review,* 5**,** 38.

GEURS, K., ZONDAG, B., DE JONG, G. & DE BOK, M. 2010. Accessibility appraisal of land-use/transport policy strategies: More than just adding up travel-time savings. *Transportation Research Part D: Transport and Environment,* 15**,** 382-393.

GIBBONS, S., LYYTIKÄINEN, T., H., O. & SANCHIS-GUARNER, R. 2019. New road infrastructure: the effects on firms. *Journal of Urban Economics.,* 110**,** 35-50.

GIBBONS, S. & SILVA, O. 2008. Urban density and pupil attainment. *Journal of Urban Economics,* 63**,** 631-650.

GIBBONS, S. & WU, W. 2020. Airports, access and local economic performance: evidence from China. *Journal of Economic Geography,* 20**,** 903-937.

GLAESER, E. L. & GOTTLIEB, J. D. 2006. Urban resurgence and the consumer city. *Urban studies,* 43**,** 1275-1299.

GLAESER, E. L., KOLKO, J. & SAIZ, A. 2001. Consumer city. *Journal of economic geography,* 1**,** 27-50.

GOTWAY, C. A. & YOUNG, L. J. 2002. Combining incompatible spatial data. *Journal of the American Statistical Association,* 97**,** 632-648.

GRAHAM, D. J. 2007. Agglomeration, productivity and transport investment. *Journal of transport economics and policy (JTEP),* 41**,** 317-343.

GRAHAM, D. J. 2009. Identifying urbanisation and localisation externalities in manufacturing and service industries. *Papers in Regional Science,* 88**,** 63-84.

GRAHAM, D. J. & GIBBONS, S. 2018. Quantifying wide economic impacts of agglomeration for transport appraisal: existing evidence and future directions. *CEP Discussion Papers (CEPDP1561).* London, UK: Centre for Economic Performance, London School of Economics and Political Science.

GRAHAM, D. J. & GIBBONS, S. 2019. Quantifying Wider Economic Impacts of agglomeration for transport appraisal: Existing evidence and future directions. *Economics of Transportation,* 19**,** 100121.

GRAHAM, D. J., GIBBONS, S. & MARTIN, R. 2009. Transport investment and the distance decay of agglomeration benefits. London.

GRAHAM, D. J., GIBBONS, S. & MARTIN, R. 2010. Transport investment and the distance decay of agglomeration benefits. London: Department for Transport.

GROVER, A., LALL, S. V. & TIMMIS, J. 2021. Agglomeration Economies in Developing Countries. Policy Research Working Paper;No. 9730. Washington DC: World Bank.

HAUNER, D. 2008. Explaining Differences in Public Sector Efficiency: Evidence from Russia's Regions. *World Development,* 36**,** 1745-1765.

HOLMGREN, J. & WEINHOLT, Å. 2016. The influence of organisational changes on cost efficiency in fire and rescue services. *International Journal of Emergency Management,* 12**,** 343-365.

HORTAS-RICO, M. & SOLÉ-OLLÉ, A. 2010. Does urban sprawl increase the costs of providing local public services? Evidence from Spanish municipalities. *Urban studies,* 47**,** 1513-1540.

HYMEL, K. 2019. If you build it, they will drive: Measuring induced demand for vehicle travel in urban areas. *Transport policy,* 76**,** 57-66.

JACOBS, J. 1984. *Cities and the wealth of nations. Principles of economic life. ,* New York, Vintage.

KANEMOTO, Y. 2013. Second-best cost–benefit analysis in monopolistic competition models of urban agglomeration. *Journal of urban economics,* 76**,** 83-92.

KIDOKORO, Y. 2015. Cost–benefit analysis for transport projects in an agglomeration economy. *Journal of Transport Economics and Policy (JTEP),* 49**,** 454-474.

KIRJAVAINEN, T. & LOIKKANEN, H. A. 1998. Efficiency differences of Finnish senior secondary schools: an application of DEA and Tobit analysis. *Economics of Education Review,* 17**,** 377-394.

KNUDSEN, E. S., HJORTH, K. & PILEGAARD, N. 2022. Wages and accessibility–Evidence from Denmark. *Transportation Research Part A: Policy and Practice,* 158**,** 44-61.

KPMG 2013. HS2 Regional Economic Impacts. Report for HS2 Ltd. Report dated Sept 2013.

LADD, H. F. 1994. Fiscal impacts of local population growth: A conceptual and empirical analysis. *Regional Science and Urban Economics,* 24**,** 661-686.

LAIRD, J. & TVETER, E. 2021. Agglomeration under covid. Report to the Department for Transport. Inverness: Peak Economics.

LAIRD, J., VENABLES, A. & JOHNSON, D. 2020. Move to More/Less Productive Jobs. report to the Department for Transport. Report dated February 2020. London: Department for Transport.

LAIRD, J. J. & VENABLES, A. J. 2017. Transport investment and economic performance: A framework for project appraisal. *Transport Policy,* 56**,** 1-11.

LIU, C. H., ROSENTHAL, S. S. & STRANGE, W. C. 2020. Employment density and agglomeration economies in tall buildings. *Regional Science and Urban Economics,* 84**,** 103555.

MARÉ, D. C. & GRAHAM, D. J. 2013. Agglomeration elasticities and firm heterogeneity. *Journal of Urban Economics,* 75**,** 44-56.

MARSHALL, A. 1890. *Principles of Economics,* London, Macmillan.

MATĚJOVÁ, L., NEMEC, J., KŘÁPEK, M. & KLIMOVSKÝ, D. 2017. Economies of scale on the municipal level: fact or fiction in the Czech Republic? *Network of Institutes and Schools of Public Administration in Central and Eastern Europe. The NISPAcee Journal of Public Administration and Policy,* 10**,** 39.

MELO, P. C., GRAHAM, D. J. & NOLAND, R. B. 2009. A meta-analysis of estimates of urban agglomeration economies. *Regional science and urban Economics,* 39**,** 332-342.

PIERCE, D., SHEPHERD, S. & JOHNSON, D. 2019. Modelling the Impacts of Inter-City Connectivity on City Specialisation. *International Journal of System Dynamics Applications (IJSDA),* 8**,** 47-70.

PIERCE, D. E. 2020. *Modelling the Economic Impacts of Inter-City Connectivity.* University of Leeds.

PRIETO, Á. M., ZOFÍO, J. L. & ÁLVAREZ, I. 2015. Cost economies, urban patterns and population density: The case of public infrastructure for basic utilities. *Papers in Regional Science,* 94**,** 795-816.

PUGA, D. 2010. The magnitude and causes of agglomeration economies. *Journal of regional science,* 50**,** 203-219.

RICE, P., VENABLES, A. J. & PATACCHINI, E. 2006. Spatial determinants of productivity: Analysis for the regions of Great Britain. *Regional science and urban economics,* 36**,** 727-752.

ROCA, J. D. L. & PUGA, D. 2017. Learning by Working in Big Cities. *The Review of Economic Studies,* 84**,** 106-142.

ROHANI, M. & LAWRENCE, G. 2017. *The Relationship Between Pedestrian Connectivity and Economic Productivity in Auckland's City Centre: Network Scenario Analysis*, Auckland Council, Te Kaunihera o Tāmaki Makaurau.

ROSENTHAL, S., S. & STRANGE, W., C. 2004. Evidence on the nature and sources of agglomeration economies. *In:* HENDERSON, J. V. & THISSE, J.-F. (eds.) *Handbook of regional and urban economics* Elsevier.

ROSENTHAL, S. S. & STRANGE, W. C. 2001. The determinants of agglomeration. *Journal of urban economics,* 50**,** 191-229.

ROSENTHAL, S. S. & STRANGE, W. C. 2020. How close is close? The spatial reach of agglomeration economies. *Journal of Economic Perspectives,* 34**,** 27-49.

SOUKOPOVÁ, J., NEMEC, J., MATEJOVÁ, L. & STRUK, M. 2014. Municipality size and local public services: do economies of scale exist? *Network of Institutes and Schools of Public Administration in Central and Eastern Europe. The NISPAcee Journal of Public Administration and Policy,* 7**,** 151.

SUN, D., ZENG, S., MA, H. & SHI, J. J. 2021. How do high-speed railways spur innovation? *IEEE Transactions on Engineering Management*.

TVETER, E. 2018. Commuting as an Indicator of Wider Economic Benefits of Transport Improvements: Evidence from the Eiksund Connection.

TVETER, E. 2020. Explaining differences in ex-ante calculations of wider economic impacts: A review of 55 calculations. *Case studies on transport policy,* 8**,** 1401-1411.

TVETER, E. & LAIRD, J. 2018. Agglomeration. How long until we see the benefits? *Scottish Transport Applications and Research.* Glasgow Caledonian University.

TVETER, E., LAIRD, J. J. & AALEN, P. 2022. Spatial aggregation error and agglomeration benefits from transport improvements. *Transportation Research Part A: Policy and Practice,* 164**,** 257-269.

TVETER, E., WELDE, M. & ODECK, J. 2017. Do Fixed Links Affect Settlement Patterns: A Synthetic Control Approach. *Research in Transportation Economics,* 63**,** 59-72.

VENABLES, A. J. 2007. Evaluating urban transport improvements: cost–benefit analysis in the presence of agglomeration and income taxation. *Journal of Transport Economics and Policy,* 41**,** 173-188.

VENABLES, A. J. 2017. Expanding cities and connecting cities: the wider benefits of better communications. *Journal of Transport, Economics and Policy,* 51**,** 1-19.

VENABLES, A. J., LAIRD, J. & OVERMAN, H. 2014. Transport investment and economic performance: Implications for project appraisal.

APPENDIX

If we take it that productivity in zone i is a function of $ATEM_i$ and a set of other influencing factors, and this is estimated in a double-log format then the agglomeration elasticity ρ is the coefficient on the ATEM term. If we also take it that the transport cost c_{ij} between zones can be expressed in different metrics (e.g. pence, pounds, seconds, minutes), which are scalable by a factor $\boldsymbol{a}$, then our interest is whether $\boldsymbol{a}$ influences the size of ρ.

The estimated agglomeration elasticity is scale invariant when using a power function for ATEM. This is shown below:

$$\rho \text{In}(ATEM_i) = \rho \ln \sum_j \frac{E_j}{(\boldsymbol{a}c)_{ij}^{\alpha}}$$

$$= -\rho\alpha \ln(\boldsymbol{a}) + \rho \text{In} \sum_j \frac{E_j}{c_{ij}^{\alpha}}$$

$$= constant + \rho \ln \sum_j \frac{E_j}{c_{ij}^{\alpha}} \qquad (5)$$

where $constant = -\rho\alpha \ln(\boldsymbol{a})$.

Thus, in estimation, the scaling factor is factored out and does not affect the estimate of the agglomeration elasticity ρ.

In contrast, scale invariance is not preserved when using a negative exponential function for ATEM. To see this, we replicate the above calculation using an exponential function when constructing ATEM, with γ as the exponential (decay) parameter.

$$\rho \text{In}(ATEM_i) = \rho \ln \sum_j E_j * e^{-\boldsymbol{a}\gamma c_{ij}}$$

$$= \rho \ln \sum_j E_j * \left(e^{-c_{ij}}\right)^{\boldsymbol{a}\gamma} \qquad (6)$$

Hence, there is no general way of factoring out the scaling factor $\boldsymbol{a}$. The implication is that for any given γ the agglomeration elasticity will be a function of the scaling factor $\boldsymbol{a}$.

Another way to see the relevance of the scale dependency is to observe whether the relative change in ATEM is scale dependent or independent with different functional forms for ATEM. This property can be seen by considering equation (3) in the main body of the report. This equation is reproduced below and is an essential step in the calculation of agglomeration benefits.

$$\theta_i^I = \left(\frac{ATEM_i^{I,DS}}{ATEM_i^{I,DM}}\right)^{\rho^I} - 1 \qquad (3)$$

Plugging in the definition of ATEM with a power function from equation (2) and multiplying travel costs by a scaling parameter $\boldsymbol{a}$ gives

$$\theta_{i,\boldsymbol{a},pow}^{I} = \left(\frac{\sum_j \sum_m \frac{E_j^{DS}}{\left(a\overline{GTC}_{ij}^{m,DS}\right)^{\alpha^I}}}{\sum_j \sum_m \frac{E_j^{DM}}{\left(a\overline{GTC}_{ij}^{m,DM}\right)^{\alpha^I}}} \right)^{\rho^I} - 1 \qquad (7)$$

Factoring on $\boldsymbol{a}$

$$\theta_{i,\boldsymbol{a},pow}^{I} = \left(\frac{\boldsymbol{a}^{\alpha^I}}{\boldsymbol{a}^{\alpha^I}} \frac{\sum_j \sum_m \frac{E_j^{DS}}{\left(\overline{GTC}_{ij}^{m,DS}\right)^{\alpha^I}}}{\sum_j \sum_m \frac{E_j^{DM}}{\left(a\overline{GTC}_{ij}^{m,DM}\right)^{\alpha^I}}} \right)^{\rho^I} - 1$$

$\boldsymbol{a}$ cancels out to give:

$$\theta_{i,\boldsymbol{a},pow}^{I} = \left(\frac{\sum_j \sum_m \frac{E_j^{DS}}{\left(\overline{GTC}_{ij}^{m,DS}\right)^{\alpha^I}}}{\sum_j \sum_m \frac{E_j^{DM}}{\left(a\overline{GTC}_{ij}^{m,DM}\right)^{\alpha^I}}} \right)^{\rho^I} - 1$$

Implying that for the power function θ (where the transport costs are scaled by $\boldsymbol{a}$**)** equals the θ as if transport costs had not been scaled by $\boldsymbol{a}$. That is:

$$\theta_{i,\boldsymbol{a},pow}^{I} = \theta_{i,pow}^{I}$$

Doing the same exercise with an exponential function gives:

$$\theta_{i,\boldsymbol{a},\boldsymbol{exp}}^{I} = \left(\frac{\sum_j \sum_m \frac{E_j^{DS}}{exp\left(a\gamma\overline{GTC}_{ij}^{m,DS}\right)}}{\sum_j \sum_m \frac{E_j^{DM}}{exp\left(a\gamma\overline{GTC}_{ij}^{m,DM}\right)}} \right)^{\rho^I} - 1 \qquad (8)$$

Since we cannot factor the scaling parameter outside the summation operator, the productivity difference will also depend on the scale. Thus:

$$\theta_{i,\boldsymbol{a},\boldsymbol{exp}}^{I} \neq \theta_{i,\boldsymbol{exp}}^{I}$$

Equation 8 appears rather complex, and the finding might be more easily seen for a simple 1 zone, 1 mode, 1 industry scenario where employment does not change between scenarios. In this case, Equation 8 simplifies to:

$$\theta_{\boldsymbol{a},\boldsymbol{exp}} = \left(\frac{\frac{E_j}{exp\left(\boldsymbol{a}\gamma\overline{GTC}_{ij}^{DS}\right)}}{\frac{E_j}{exp\left(\boldsymbol{a}\gamma\overline{GTC}_{ij}^{DM}\right)}} \right)^{\rho} - 1$$

Cancelling through by E_j and re-arranging:

$$\theta_{\boldsymbol{a},\boldsymbol{exp}} = \left(\frac{exp\left(\boldsymbol{a}\gamma\overline{GTC}_{ij}^{DM}\right)}{exp\left(\boldsymbol{a}\gamma\overline{GTC}_{ij}^{DS}\right)} \right)^{\rho} - 1$$

Factoring by $\boldsymbol{a}$ and gives:

$$\theta_{\boldsymbol{a},\boldsymbol{exp}} = \left(\frac{exp\left(\overline{GTC}_{ij}^{DM}\right)}{exp\left(\overline{GTC}_{ij}^{DS}\right)} \right)^{\boldsymbol{a}\gamma\rho} - 1$$

Showing that, for any given γ and ρ the exponential function, $\theta_{\boldsymbol{a},\boldsymbol{exp}}$ is dependent on the scaling factor $\boldsymbol{a}$. Therefore:

$\theta_{\boldsymbol{a},\boldsymbol{exp}} \neq \theta_{\boldsymbol{exp}}$