Review of the report on agglomeration and transport appraisal by Laird & Tveter

Prof Dan Graham, Imperial College London, London, SW7 2AZ, UK.
Email: d.j.graham@imperial.ac.uk

October 30, 2023

Introduction

This document provides a review of the report by Laird and Tveter (2023) entitled 'Agglomeration and transport appraisal: new developments and research directions'. The Laird and Tveter (L & T) report was commissioned by the Department for Transport to review recent developments in the literature on agglomeration and transport appraisal, and to identify evidence gaps and directions for future research.

There are three main chapters in the L & T report:

1. **Economics of agglomeration** - provides a summary of the conceptual basis for inclusion of agglomeration externalities as wider economic impacts (WEIs) in transport appraisal, including effects on production and consumption.
2. **Public sector elasticity** - provides a brief review of the evidence for agglomeration effects in public sector industries and discusses how this evidence should be operationalised in transport appraisal.
3. **Access to economic mass (ATEM)** - provides a review of issues concerning the use of ATEM measures to represent agglomeration and an assessment of possible courses of action for appraisal going forward.

In this review my comments are structured in relation to the three themes covered in the L & T report. Prior to my detailed comments, I first provide a summary of the main points.

Summary

- The review of supply-side agglomeration effects (e.g. production externalities) largely confirms what we already know and confirms existing appraisal practice.
- L & T argue that the agglomeration elasticities we currently use for transport appraisal should be revisited in light of development in both our understanding of forces of agglomeration (supply side and demand side) and in the practice of transport appraisal. I agree with this conclusion.

- L & T note that agglomeration effects often coincide with other source of market failure such as tax wedges and coordination failure. I agree with this and I question the extent to which existing appraisal practice is able to to quantify their value distinctly. This is an issue I believe is worth exploring further.
- The report argues that long distance agglomeration effects may exist but only for a subset of the economy and only from a subset of sources of agglomeration. This assessment sounds reasonable, and is worth investigating further particularly in relation to industry specific effects.
- The distinction between urbanisation (city / labour market level) and localisation (industry level) economies is not currently used in TAG. The scope for doing this should be explored, however, as Graham and Gibbons (2019) point out there are substantial conceptual and methodological difficult that will likely prove difficult to overcome.
- The report rightly emphasises the importance of consumption side agglomeration externalities, which have been less extensively studied. I agree that it is worth exploring whether and how consumption externalities could be represented as WEIs in appraisal.
- There are considerable conceptual and practical difficulties in applying the concept of agglomeration to public services.
- The assumptions that we make when modelling the link between productivity and ATEM are not easily applied in the public sector setting. For example, we generally cannot assume:
  - Profit maximising or cost minimising behaviour.
  - Approximate equivalence of wages to the value of the marginal product.
  - Freedom in locational choice of production.
  - Operation within a competitive market structure.
- L & T tend not to address these conceptual / methodological challenges, instead evaluating practical fixes for appraisal.
- The material L & T review does not, in my view, provide a robust evidence base to justify application of an economy wide agglomeration elasticity to the public sector.
- L & T raise a number of issues concerning representation of agglomeration via an ATEM measure: representation of economic mass, modal disaggregation, zonal aggregation, and functional form. These issues have been discussed for many years now and were reviewed in depth by Graham and Gibbons (2019). I agree that it is high time these issues were resolved definitively and I recommend that DfT pursue work to achieve this.
- Reviewing the current TAG agglomeration calculation, I noticed that the form proposed may underestimate agglomeration impacts when the change in ATEM is large. There may be particular reasons for adoption of this (log differential) form that I am missing, but worth revisiting to assess whether a simpler percentage change calculation could be used instead.
- L & T recommend that further work be undertaken on public sector agglomeration elasticities, distributional agglomeration effects, localisation economies & consumption externalities, and active travel & long distance benefits. I agree that these are potentially good topics for further investigation. However, I emphasise the need for theoretical / conceptual work as a necessary prelude to developing empirical solutions for appraisal.

1 Production and consumption benefits of agglomeration and their additionality in transport appraisal

This part of the report provide a review of agglomeration impacts and a justification for their inclusion as WEIs within transport appraisal. It considers agglomeration impacts on both production and consumption, and discusses issues related to their: additionality, overlap, spatial reach, temporal phasing and social distribution.

1.1 Production externalities

The review of agglomeration externalities in production covers empirical evidence summarised in some recent review / meta-analysis papers. The main take away conclusions from this part of the report are as follows.

- The agglomeration of economic activity, in cities or industrial clusters, induces productivity benefits for firms and workers.
- These benefits are generated via improved opportunities for learning (e.g. knowledge spillovers), matching (e.g. thick labour markets) and sharing (e.g. scale economies in input and outputs sharing).
- Transport connectivity affects *relative* agglomeration via its impact on proximity (e.g. static agglomeration) and clustering (e.g. dynamic agglomeration). A change in transport provision can induce a change in the agglomeration experienced by firms via these two routes, with implications for productivity.
- Agglomeration effects are found to vary by study context (e.g. country, industry and study design etc). A mean agglomeration elasticity value in the range 0.02 to 0.05 is often reported.
- Service sectors tend to benefit more from agglomeration than manufacturing industries, and consequently have higher agglomeration elasticities.
- Study design and methodology are critical in ensuring robustness of estimates. In particular, issues of endogeneity must be addressed properly to achieve identification.
- Studies that adopt methods to deal with endoegeneity typically report smaller agglomeration elasticities.

These points are, I believe, well known and do not represent anything particularly new. There are, however, two points made by L & T in relation to productivity effects that are worth commenting on.

1) **Use of meta-analysis estimates in appraisal** - L & T appear to endorse the finding of Donovan et al. (2021) that country specific effects are not an important determinant of the magnitude of agglomeration elasticities; and on that basis, suggest that meta-analysis estimates could be used in appraisal (p 7 para 4). I would urge caution here. Not only does the finding of invariance across countries contrast with that of other reviews based on extensive empirical evidence (e.g. Melo et al. 2009, Ahlfeldt and Pietrostefani 2019), it is also made without theoretical justification or conceptual reasoning in the L & T report and indeed in the Donovan et al. (2021) meta-analysis itself. If country specific effects are unimportant then we could safely assume that agglomeration elasticities are

constant over a broad range of ATEM. I do not believe this to be true and there is evidence for nonlinearity of agglomeration effects that supports this belief (e.g. Graham and Van Dender 2011).

2) **Bias of TFP based agglomeration elasticities** - L & T argue that agglomeration elasticities estimated using a TFP model tend to be upwards biased because journeys are shorter in denser location allowing firms to spend more time on productive activities rather that travelling. I find this claim odd for two main reasons. First, while it is true that proximity is generally greater in denser areas, diseconomies of congestion tend to extend journey times thus diluting the travel time benefits of density. Second, if there is a productivity effect due to a reduction in the requirement to travel, that is surly a source of the agglomeration externality itself, and not a bias that we want to eliminate. If a firm moves closer to other firms it will value the benefits that accrue due to a change in its proximity, but it will disregard the external benefit that its action yields for other firms.

1.2 Interaction between agglomeration and other WEIs

L & T note that agglomeration effects will often coincide with other source of market failure such as tax wedges and coordination failure. They review the relevant theory, as developed in various papers, and argue that there are three WEIs that, for the purpose of appraisal, can legitimately be regarded as additional to agglomeration

i. Tax revenues arising from labour market impacts.
ii. Imperfect competition effects that arise when output expansion yields a welfare gain in monopolistic markets
iii. Coordination failures in development that are remedied thorough public delivery of infrastructure.

I agree that these impacts are, in theory, additional to agglomeration. However, it would presumably be hard to obtain empirical measures that quantify their value distinctly, and in fact, it is likely the case that the agglomeration elasticities already incorporate productivity effects arising from at least some of these sources. For instance, if the benefits of imperfect competition and improved coordination grow with ATEM, and if they have a positive effect on the productivity of firms, then there is a problem of observational equivalence that econometric models will likely fail to solve. The agglomeration elasticity essentially captures the combined impact on productivity that arises from all sources which change systematically with ATEM.

1.3 Long distance agglomeration, urbanisation and localisation economies

The first part of this section of the report sets out the economic case for the generation of WEIs via long-distance connectivity. It argues that long distance agglomeration effects are likely only enjoyed by a subset of the economy and will arise only from a subset of sources of agglomeration (mainly sharing of inputs).

This assessment sounds reasonable, but worth emphasising that there is likely to be an industry specific element that may be worth investigating. We know that agglomeration benefits differ across sectors of the economy, as do the travel patterns associated with the sources or

mechanisms of agglomeration. It seems likely that mechanisms of agglomeration could operate at both long and short distances, or neither, depending on the sector.

The second part of this section of the report deals with the distinction between urbanisation (city / labour market level) and localisation (industry level) economies. The authors note that the current TAG agglomeration elasticities do not map to this distinction. The scope for separately evaluating localisation and urbanisation effects as WEIs was consider by Graham and Gibbons (2019). They note that while separate estimates of the elasticity of productivity with respect to urbanisation and localisation could in principle be obtained for different industrial sectors, there are conceptual and methodological difficulties that make this hard to achieve.

- For most sectors of the economy correlation between measures of localisation and urbanisation is high and likely to induce a problem of multicollinearity in econometric estimation.
- Conceptually, it is hard to imagine a situation in which a transport intervention alter localisation without simultaneously altering urbanisation. Thus, to treat the two effects as distinct additive components, rather than combining them in a general agglomeration term, may not really add any additional value.
- It is debatable whether industrial classification is really the most effective categorisation we can use to cluster like firms. There is evidence, for instance, that concentrations based on occupation / skills could be more relevant.

1.4 Consumption externalities

The report rightly emphasises the importance of consumption externalities, which have been less extensively studied as foundations of agglomeration than production externalities. This oversight has been rectified in the recent spatial economics literature which has generated evidence on the importance of these effects.

Production externalities cover supply side effects. On the demand side, the argument is that externalities in consumption induce firms to locate closer to consumers and exploit the benefits of scale, while consumers enjoy lower prices and better opportunities for consumption of amenities and goods in larger markets (e.g. variety).

L & T focus on the gains in variety element of consumption externalities, arguing that transport improvements can increase the variety available to consumers, but that not all of this gain is additional to that captured by conventional user benefits. It is additional only if the gain in variety can be sourced to correction of a market failure. The authors note that identification of gains to variety that are genuinely additional can be tricky, and there is scope for both double counting with conventional benefits and duplication with production externalities effects.

In my view, it is certainly worth exploring whether and how consumption externalities could be represented as WEIs in appraisal. This could be done through the addition of a new WEI component, much like the production side agglomeration calculation. Or it could be done through the modelling of dynamic agglomeration effects, for instance using a Quantitative Spatial Model (QSM).

2 Public sector elasticity

The literature has, to my knowledge, not reached solid conclusions on whether agglomeration delivers returns to public services. Typically, the public sector is simply excluded from analyses of the effects of agglomeration on productivity. The reason for this is that the assumptions that we typically have to make when modelling the link between productivity and ATEM are not easily applied in the public sector setting. For example, we generally cannot assume:

- Profit maximising or cost minimising behaviour.
- Approximate equivalence of wages to the value of the marginal product.
- Freedom in locational choice of production.
- Operation within a competitive market structure.

In short, there are considerable conceptual and practical difficulties in applying the concept of agglomeration to public services.

To investigate the issue of L & T review evidence from two sources.

1. **Spatial variance in public service efficiency** - studies show mixed evidence of returns to public services with a credible link to ATEM. Some of the papers they review relate public sector expenditure to urban density and apparently find efficiency gains. But it is of course hard to know if this effect arises via mechanisms of agglomeration (e.g. sharing, matching or learning) or because density simply dilutes the cost of public service supply.
2. **Agglomeration studies that have included public sector activities** - studies report some evidence of an urban wage premium in public services. Again, its not clear that this can be construed as evidence of agglomeration economies since public workers are often paid an explicit urban weighting that is not related to their productivity per se, but given to compensate for higher urban prices.

In summary, L & T find mixed evidence of returns from agglomeration to public services, and they are unable to infer much from existing agglomeration studies. The discussion they present is entirely empirical, and does not present a conceptual case for the existing of agglomeration effects on public sectors. In my view, if this issue is to be taken further, it will be first necessary to develop this theoretical / conceptual case and crucially define the outcomes that we expect agglomeration to be capitalised within. This can then for the foundations upon which an empirical strategy for identification of an elasticity can be based.

L & T conclude that the agglomeration effect in the public sector "is closer to the economy weighted average than it is to zero". This may be true, but given that the authors really have no insights at all into the magnitude of any effect, should it exist, it seem to me an odd statesman to make. Certainly, the material they review does not provide a robust evidence base to justify application of an economy wide agglomeration elasticity to the public sector.

3 ATEM

3.1 B2B and B2C

The authors note that agglomeration effects arise via interactions between firms and between workers and firms. They refer to the former as B2B and the latter as B2C.

The issue L & T raise is whether econometric models of agglomeration should include two separate ATEM measures : one that uses employment as the measure of mass (B2B) and one that uses population as the measure of mass (C2C). They note that Graham and Gibbons (2019) find high correlations between these measures which would cause issues in achieving separate identification econometrically.

L & T argue that a choice has to be made between one measure or the other, and that B2B is preferable as it relates mor to the microfoundations of agglomeration. I would like to make tow points on this claim.

- In practice, the correlation between the ATEM measures is so hight that we are unlikely learn anything additional about agglomeration effects using one mass measure or the other.
- There is no reason why we cannot construct a single ATEM that uses both population and employment as measures of mass, e.g.

$$\rho_i^{PE} = \frac{1}{n}\sum_{j=1}^{n}(E_j + P_j)f(d_{ij})$$

3.2 Modal ATEM

The report discusses the issue of modal proximity and whether there is scope for estimating agglomeration impacts separately for measures of ATEM distinguished by mode. This issue was also considered in Graham and Gibbons (2019) who concluded that the correlations between ATEM measures with impedance measured by different modes was so high that an econometric model would be unable to identify separate elasticities.

L & T note that a study by D'Costa et al. (2013) estimated elasticities of wages with respect to ATEM, where ATEM was calculated for two modes (car and rail) using GC as the impedance factor. These elasticities, which were not proposed by the authors for use in appraisal, are shown below.

Mode	*Elasticity*
Car	0.069
Rail	0.055

D'Costa et al. (2013) were able to get around the multicollinearity problem and estimate separate car and rail elasticities in their regressions by using different levels of spatial aggregation in their ATEM measures. The rail ATEM measure was calculated at inter-Local Authority level and the car ATEM measure at Output Area level. Consequently, the rail measure captures changes in accessibility at an aggregate level and the road measure at a more localised

level. The net result is that the elasticities do not actually provide distinct *modal* elasticity values per se, but rather some combined effects of mode and spatial aggregation.

3.3 Zonal aggregation

I agree that the smallest level of zonal aggregation should be used to calculate agglomeration benefits.

3.4 Functional form of ATEM

I agree that more research is required on the functional form of ATEM measures, including use of 'nonparametric' forms in the econometric models. Graham and Gibbons (2019) discussed this issue at length and also recommended that further research be undertaken.

On page 3 of the report, and in the appendix, L & T make some comments on the TAG agglomeration calculation. Using the notation of Graham and Gibbons (2019), this calculation can be written

$$\omega^1 - \omega^0 = \left[\left(\frac{\rho^1}{\rho^0}\right)^{\eta} - 1\right]\omega^0,$$

where ω is a measure of productivity (or GDP), ρ is an effective density (ATEM), η is the agglomeration elasticity (e.g. $\eta = \partial \log \omega / \partial \log \rho$), and superscripts 0 and 1 denote the baseline and intervention scenarios respectively.

It is worth noting that this equation will underestimate agglomeration impacts when the change in ATEM is large. I have included a note at the end of this report that explains why this is the case.

Instead, we could use a perhaps simpler calculation based on the discrete approximation

$$\omega^1 - \omega^0 \approx \eta \; \%\Delta \; \rho \; \omega^0.$$

See the note at the end for a comparison of impacts under this calculation and the current TAG version.

4 Implications for policy and research

4.1 Immediate updates to TAG

1. **Public sector elasticities** - L & T recommend that for the purpose of appraisal an economy wide agglomeration elasticity should be applied to public sector SICs. For reasons explained above, I do not think the authors have shown sufficient evidence to justify this approach, and I believe it would be prudent to undertake further conceptual and empirical work before current appraisal practice is changed. Crucially, in my view, there would be real value in developing a theoretical analysis that can
    - Show how gains from agglomeration differ between the public and private sectors, and suggest mechanisms that could generate agglomeration in a non-market production structure but within a competitive labour market.
    - Provide a conceptual foundation upon which empirical approaches can be developed to approximate agglomeration effects for the public sector.
2. **Spatial level of agglomeration calculations** - I agree that a small spatial scale should be used if possible, but of course the level of aggregation is often determined by data availability.

4.2 Research to be commissioned

L & T recommend future work on distributional agglomeration effects, localisation economies & consumption externalities, and active travel & long distance benefits.

I agree that these are potentially good topics for further investigation. However, I do not agree that the way forward is to plug in some parameters or decay functions see what transpires. To develop these components properly, and provide a robust justification for their separate inclusion in appraisal, theoretical / conceptual work is a necessary prelude to empirical estimation and practical calculation.

References

Ahlfeldt, G. M. and E. Pietrostefani (2019). The economic effects of density: A synthesis. *Journal of Urban Economics 111*, 93–107.

D'Costa, S., G. S, O. H, and P. Pelkonen (2013). *Geography, Institutions and Regional Economic Performance*, Chapter Agglomeration and labour markets: the impact of transport investments on labour market outcomes. Springer.

Donovan, S., T. de Graaff, H. L. de Groot, and C. C. Koopmans (2021). Unraveling urban advantages—a meta-analysis of agglomeration economies. *Journal of Economic Surveys*.

Graham, D. J. and S. Gibbons (2019). Quantifying wider economic impacts of agglomeration for transport appraisal: Existing evidence and future directions. *Economics of Transportation 19*, 100121.

Graham, D. J. and K. Van Dender (2011). Estimating the agglomeration benefits of transport investments: some tests for stability. *Transportation 38*, 409–426.

Melo, P. C., D. J. Graham, and R. B. Noland (2009). A meta-analysis of estimates of urban agglomeration economies. *Regional Science and Urban Economics 39*, 332–342.

Note on the TAG equations for agglomeration impacts

Using the notation of Graham and Gibbons (2019), the calculation for agglomeration impacts suggested in TAG is

$$\omega^1 - \omega^0 = \left[\left(\frac{\rho^1}{\rho^0}\right)^{\eta} - 1\right]\omega^0, \tag{1}$$

where ω is a measure of productivity (or GDP), ρ is an effective density (ATEM), η is the agglomeration elasticity (e.g. $\eta = \partial \log \omega / \partial \log \rho$), and superscripts 0 and 1 denote the baseline and intervention scenarios respectively.

From this equation, the percentage change in productivity is defined

$$\frac{\omega^1 - \omega^0}{\omega^0} \approx \left(\frac{\rho^1}{\rho^0}\right)^{\eta} - 1.$$

This is a log differential form, and it provides a good approach to calculating percentage change when ρ^1/ρ^0 is in the neighbourhood of 1.0 (e.g. when $\Delta\rho$ is small), but is less satisfactory the further we get from 1.0 (see explanation in box below).

Instead, we could simply use the presumably more familiar calculation based on a discrete approximation

$$\frac{\omega^1 - \omega^0}{\omega^0} \approx \eta \frac{\rho^1 - \rho^0}{\rho^0} \quad \text{e.g.} \quad \%\Delta\ \omega \approx \eta\ \%\Delta\ \rho,$$

This approximation assumes that the elasticity is constant, in the sense that it is good for approximating change in ω over both big and small changes in ρ (e.g. over intervals as well as around points). Since this assumption corresponds with that implicitly made in the econometric models, there seems no harm in continuing it in the practical calculations. Furthermore, the notion that ω^0 corresponds with ρ^0 and ω^1 with ρ^1 is made explicit, and this is I believe what we want in appraisal.

Thus, the guidance could replace (1) with

$$\omega^1 - \omega^0 \approx \eta\ \%\Delta\ \rho\ \omega^0. \tag{2}$$

The table below illustrates the difference in agglomeration impact calculation achieved using a discrete approximation versus log differential (e.g. TAG) form, for different values of ρ^1/ρ^0. The calculations use an elasticity of $\eta = 0.04$, and a base productivity (GDP) of $\omega = 500$ units.

Table 1: Difference in agglomeration impact estimation using a discrete approximation versus log differential (e.g. TAG) form

	Discrete approx.		TAG	
ρ^1/ρ^0	$\%\Delta\ \omega$	$\Delta\ \omega$	$\%\Delta\ \omega$	$\Delta\ \omega$
1.0	0.000	0.000	0.000	0.000
1.1	0.400	2.000	0.382	1.910
1.2	0.800	4.000	0.732	3.660
1.3	1.200	6.000	1.055	5.275
1.4	1.600	8.000	1.355	6.775
1.5	2.000	10.000	1.635	8.175
1.6	2.400	12.000	1.898	9.489
1.7	2.800	14.000	2.145	10.726
1.8	3.200	16.000	2.379	11.895
1.9	3.600	18.000	2.601	13.003
2.0	4.000	20.000	2.811	14.057

Log Differential form of the TAG Agglomeration Calculations

From equation (1), we can write the intervention scenario productivity as

$$\omega^1 = \omega^0 \left(\frac{\rho^1}{\rho^0}\right)^{\eta}.$$

Taking logs,

$$\log \omega^1 = \log \omega^0 + \eta \log(\rho^1/\rho^0)$$

and writing in differenced form, we have

$$\log \omega^1 - \log \omega^0 = \eta \left(\log \rho^1 - \log \rho^0\right)$$

For an infinitesimally small change in $\log \rho$, e.g. $\log \rho^1 - \log \rho^0 = \Delta \log \rho \to 0$, then $\log \omega^1 - \log \omega^0 = \Delta\omega \to 0$, and we can write in differential form as

$$d \log \omega = \frac{\partial \log \omega}{\partial \log \rho} d \log \rho.$$

A percentage change calculation $(\omega^1 - \omega^0)/\omega^0$, gets close to the log differential (and the difference in logs) when the ratio ω^1/ω^0 is close to 1.0.

To see this, note that the natural log of ω can be approximated by

$$\log(\omega) \approx \omega - 1$$

in the neighbourhood of 1.0. This follows from Taylor's theorem. The first order Taylor expansion of $\log(\omega)$ around $\omega = 1$ is

$$\log(\omega) \approx \log(1) + \frac{d}{d\omega} log(\omega) \mid_{\omega=1} (\omega - 1),$$

implying

$$\log(\omega) \approx 0 + \frac{1}{1}(\omega - 1).$$

Considering the log difference,

$$\log \omega^1 - \log \omega^0 = \log \left(\frac{\omega^1}{\omega^0} \right),$$

the we can use Taylor's approximation around $\omega^1/\omega^0 = 1$ to write

$$\log \omega^1 - \log \omega^0 \approx \frac{\omega^1}{\omega^0} - 1 = \frac{\omega^1}{\omega^0} - \frac{\omega^0}{\omega^0} = \frac{\omega^1 - \omega^0}{\omega^0}.$$

As $\Delta \log \rho$ increases, ω^1/ω^0 moves away from the immediate neighbourhood of 1.0, and the correspondence between percentage change and the log differential (difference) reduces.