Centre for Transport Studies
STOCKHOLM

How “wide” are the “wider economic impacts”?
On the overlap between standard CBA and agglomeration benefits

Presented at the conference
Advancing Metropolitan Modeling for the
Analysis of Urban Sustainability Policies, UC Riverside
Manuscript 2015-01-10

Jonas Eliasson
Centre for Transport Studies, KTH Royal Institute of Technology

CTS Working Paper 2015:X

Abstract
The presence of agglomeration effects or income taxation means that social benefits of transport projects may fall partly outside standard cost-benefit analysis (CBA), i.e. there may be so-called wider economic impacts. Assuming there is a method to calculate the total economic benefits of a transport project, the next question is to determine what share of these benefits actually fall outside standard CBA, the so-called overlap problem. This paper analyses this question, showing that the answer depends on what micro-mechanisms generate the agglomeration effect, such as better matching on the labour market or some type of spillover effects. Determining the relative contribution to agglomeration effects of different micro-mechanisms is difficult, casting doubt on the possibility of providing a robust answer to the overlap problem.

Keywords: Agglomeration effects, cost-benefit analysis, wider economic benefits, transport investments.

JEL Codes: D61, H54, R41, R48.

Centre for Transport Studies
SE-100 44 Stockholm
Sweden
www.cts.kth.se

1 INTRODUCTION

In recent years, increasing attention has been given to the fact that standard cost-benefit analyses (CBA) may not capture all economic benefits generated by transport investments and policies. This omission occurs because standard CBA only captures the benefits of improved accessibility by calculating travellers’ consumer surplus. In the presence of external agglomeration effects or income taxation, however, commuters may not reap the full benefits of accessibility improvements. Some may accrue to taxpayers in the form of increased tax revenues, and some to other workers, firms or customers e.g. through increased productivity. Economic benefits which fall outside standard CBA are often called *wider economic impacts* (WEI). WEI may refer to several distinct effects and mechanisms, but in this paper I will concentrate on one, arguably the quantitatively most important one: that labour productivity tends to increase with improved accessibility between workers and firms. There are several possible reasons for this widely observed phenomenon, including matching effects (more efficient pairing of employers and employees) and various kinds of spillover effects (such as knowledge spillovers). Other kinds of WEIs are outside the scope of the present paper, for example increased production in imperfectly competitive markets or improved linkages between intermediate and final goods suppliers.

To include WEIs in transport CBA, two things are needed: a quantitative causal relationship between accessibility improvements and productivity increases, and an estimate of the share of these benefits that actually fall outside standard CBA. This paper deals with the second issue. This is sometimes referred to as “the overlap problem”: the question is to what extent effects on total economic production overlaps with standard CBA.

Several countries have made efforts in recent years to introduce WEIs in their CBA methodologies. There is a large and rapidly growing literature on causal relationships between accessibility and productivity (and a growing appreciation of the many econometric pitfalls), and applied CBA guidelines are able to draw from these results. As to the second issue, however, there has been much less discussion. The scientific literature is comparatively scarce, and the few applied guidelines in existence are more or less silent as to how they have arrived at their recommendations.

In this paper, I will argue that it is in fact not possible to give an answer to the overlap problem based on the kind of observations and methods typically used. Using two versions of a very simple model, I will show that it is not possible to determine the share of agglomeration benefits captured by standard CBA using aggregate observations. In the first version of the model, all benefits are completely captured by standard CBA despite an observed positive (and causal) relationship between wage rates and accessibility, i.e. agglomeration effects. In this situation, adding these agglomeration benefits to the CBA is double-counting. In the second version of the model, substantial parts of the agglomeration benefits fall outside standard CBA. From an aggregate point of view of a hypothetical economist, the two versions of the model are essentially indistinguishable. I show that the share of agglomeration benefits captured in standard CBA depends on what mechanism is generating the agglomeration effects. If agglomeration effects are caused by matching effects, they will be captured by standard CBA (provided that changes in tax revenues are accounted for); this is the first version of the model. On the other hand, if agglomeration effects are caused by some type of spillover effects[1], then substantial parts of agglomeration benefits will fall outside standard CBA; this is the second version of the model.

In reality, agglomeration effects are most likely caused by several kinds of micro-mechanisms. This means that in order to assess how much of agglomeration benefits should be added to standard CBA, one would need to know the relative contribution to agglomeration effects from matching and spillover mechanisms, respectively. The problem is that based on aggregate observations of accessibility/wage relationships alone, it is virtually impossible to know what underlying mechanisms are causing this phenomenon. Hence, it is essentially impossible to know the overlap between standard CBA and agglomeration benefits, given the kind of methods typically in use currently.

I will also illustrate a second problem. Accessibility can be improved in several ways – by reducing travel times, reducing travel costs, or increasing travel comfort and convenience. The concept of *generalized travel cost* captures all such components of trip resistance, and is the cornerstone both of standard CBA and of the accessibility measures used to calculate agglomeration benefits[2]. However, I will show that the size of agglomeration benefits may depend on which *component* of the generalized travel cost that is affected by a transport project. In other words, a given change in generalized travel costs may give rise to different agglomeration benefits depending on which component of the generalized travel cost that changes. Since traveller benefits in standard CBA only depend on changes in generalized travel cost, not directly on its components, this means that it will not be possible to establish a fixed relationship between standard CBA benefits and WEIs – something which has been on the wish list of developers of applied CBA methodology.

Finally, I will illustrate that the unfortunate practice of standard CBA to omit changes in tax revenues may introduce substantial errors in the CBA, in particular when large parts of agglomeration benefits fall outside the CBA.

For convenience, I will use the term “agglomeration effect” in a rather restricted sense, referring to a causal relationship between improved accessibility and increases in average wages and wage rates[3]. With “agglomeration benefits” I will refer to the social benefits arising from such agglomeration effects.

The outline of the paper is as follows. Section 2 provides some background on agglomeration benefits and its relation to standard CBA. Section 3 presents a simple model showing the central argument of the paper, that the overlap between economic effects and standard CBA benefits depends on what microlevel mechanism that generates agglomeration effects, and that it is not enough to measure changes in generalized travel costs, since changing different components of it may generate different sizes of WEI. Section 4 discusses implications for applied CBA methodology, in particular two of the few existing applied guidelines for including WEI in CBA, namely the Swedish and British guidelines. Section 5 concludes.

---

[1] For convenience, I will abuse the term ”spillovers” a little and use it in a very general sense, including not only knowledge spillovers between workers and firms but also sharing, network and scale economies; essentially, any kind of mechanism that causes (local) productivity to increase when workers and firms are located close to each other.

[2] Many estimates of agglomeration effects in fact use cruder measures of proximity or contact costs, such as geographical distance or even aggregate city population (especially earlier studies). In order to use them for transport CBA, however, they need to be converted or reinterpreted to be based on generalized travel costs, since that is what is affected by transport projects.

[3] In section 4, which deals with implications for applied CBA guidelines, I will return to the issue of whether wage rates are an acceptable indicator of labour productivity.

2 AGGLOMERATION BENEFITS AND CBA – A SUMMARY

The positive relationship between city size and productivity was pointed out already by Smith (1776) and Marshall (1890). A large number of studies have confirmed the correlation between high accessibility and high productivity (Rosenthal & Strange, 2004). Establishing causality is a thorny issue, since at least part of the observed correlations have been shown to be due to self-selection (Combes, Duranton, & Gobillon, 2008; Graham, Melo, Jiwattanakulpaisarn, & Noland, 2010). However, substantial parts do indeed seem to be causal. For example, Gould (2007) controls for selection and endogeneity and shows that residing in a city do indeed make workers more productive, also in the long term.

The mechanisms in which accessibility can increase productivity are summarized by Duranton and Puga (2004) in the phrase “sharing, matching and learning”. “Sharing” refers to that workers and firms can share the costs of certain common resources such as education. “Matching” refers to mechanisms such as workers being more efficiently paired with employment opportunities, so the individual-specific skills of each worker are better used. “Learning” refers to the process where workers and firms learn from each other’s knowledge and innovations. Puga (2010) and Graham (2014) point out that the understanding of the relative contributions of the three mechanisms is limited. This is an important point for the discussion in the present paper, since a central conclusion will be that the share of agglomeration benefits captured in standard CBA depends on which mechanism that generates them. Of particular interest is a recent and very careful study by Graham (2014), showing that matching effects do indeed play an important part for agglomeration effects, even after controlling for endogeneity and selection.

All three mechanisms become stronger the larger the city is (urbanization economies) and the lower contact costs are within the city (localization economies). Investments in improved urban transport infrastructure will primarily increase localization economies by lowering contact costs within the city, although in the long run such investments may also increase the equilibrium size of the city and hence increase urbanization economies.

Accessibility refers to the ease with which firms and workers can access each other. It can be measured in different ways. Early studies used total city size as a proxy for accessibility. Most later studies have used more elaborate measures, where workers and firms are weighted with some measure of the generalized travel cost between them. Some studies have simply used the geographical distance as a proxy for the generalized cost, while others have used more realistic measures based on actual travel times and travel costs. Obviously, the latter method is necessary if results are to be used for estimating the impact of transport investments.

Studies have measured productivity in various ways, such as firm output, worker productivity or wage rates. In the stylized model in this paper, productivity gains are assumed to accrue to workers through increased wages. In reality, however, some of these gains might instead accrue to firms or their customers. This will also have implications for the share of agglomeration benefits captured by standard CBA; I will return to this in section 4.

The benefits of accessibility are captured in standard CBA through the consumer surplus (CS). Usually, the CS is calculated by the rule of a half, which is a good approximation as long as the demand curve is approximately linear. The CS calculation can in principle be very disaggregated across individuals, but in practice broader aggregates are used. A simple form would work like this: let c_{ij} and t_{ij} be the travel cost and travel time per origin-destination pair (extending to e.g. several modes, departure times etc. is trivial). Define the generalized cost as $d_{ij} = c_{ij} + \theta t_{ij}$, where θ is travellers' average value of time. The value of time will be the sum of the wage rate plus a term proportional to the direct disutility of the trip relative to some reference activity, and divided by the marginal utility of time. This means that increasing the comfort of travel will reduce the value of time and hence the generalized travel cost. Let T_{ij} be the number of trips per origin-destination pair. Let Δd_{ij} be a change in generalized cost and ΔT_{ij} be the resulting change in the number of trips. The rule-of-a-half approximation of the CS of this change is then

$$CS = \sum_{ij}\left(T_{ij} + \frac{\Delta T_{ij}}{2}\right)\Delta d_{ij}$$

This formula involves two kinds of approximations. First, the value of time is usually an average over a group of travellers – in extreme cases over all travellers in a country. In practice, the value of time may be differentiated with respect to mode, trip purpose and trip length, but seldom with respect to income or a specific project. Moreover, possible differences between existing and new travellers are ignored. This introduces a bias with respect to calculation of economic benefits which may be substantial. Second, income effects are neglected (i.e. it is assumed that the marshallian CS can be used), and the demand curve is assumed to be approximately linear. These approximations are usually small, although there are cases where they cannot be ignored.

Since standard CBA only captures benefits accruing to the traveller, any effect on an accessibility improvement on firms or workers who do not travel are neglected. In an influential paper, Venables (2007) showed that the neglected benefits may be substantial. Combined with pioneering studies of the size of agglomeration effects by Graham (summarised in Graham (2007)), the UK Department for Transport calculated the wider economic impacts of the large Crossrail investment, resulting in additional benefits of over 50% relative to the standard CBA benefits.

3 A MODEL OF AGGLOMERATION BENEFITS

The central arguments will be illustrated using a simple model. The model focuses only on how agglomeration effects can be generated through matching, spillovers and increased labour supply, and how these are affected by transport improvements. Hence, it ignores location choices, the land market and unemployment. Further, wage setting and the demand for labour is captured in a simplistic way. The model is sufficient to illustrate the central arguments, however; extending it in various ways is relatively easy, but would incur some analytical complexity and loss of intuition.

Imagine a city consisting of a suburb and a downtown. All workers live in the suburb, but they can choose between working in the suburb or downtown. Workers also choose how many hours to work W. Jobs in the suburb are generic, so all workers choosing to work there get a wage rate w_0 (in dollars per hour). Downtown jobs are specialized, and workers are heterogeneous in terms of productivity, so workers are offered different wage rates w downtown, depending on their individual productivity[4]. The

[4] The assumption that everyone gets the same wage in the suburb is inessential. More generally, we can assume that each worker-job pair is heterogeneous – workers' productivity is different on different jobs – and that workers have several jobs to choose from in each zone, so w is the maximal attainable wage rate in each zone for a given worker. However, normalizing the wage distribution in one of the

distribution of wage rate offers is denoted $f(w;N_D)$. N_D is the number of workers downtown: because of spillover effects (knowledge spillovers, economies of scale and so on), wage rate offers increase with the number of downtown workers. We assume that there are many firms competing for workers, so each worker’s wage rate offer is equal to her (constant) marginal productivity, and firms do not make any profit. Workers pay income tax on their income: w and w_0 denote wage rates after tax, while gross (before tax) wage rates are $w(1+\tau)$ and $w_0(1+\tau)$.

The commuting cost from the suburb to downtown is c and the commuting time t. Commuting cost and time within the suburb is normalized to zero. Workers are also heterogeneous in the sense that they get different, idiosyncratic utilities D of commuting to a downtown job. This represents that downtown jobs may have different intrinsic utilities, different for different workers, or that the perceived disutility of commuting may vary across workers. The distribution of idiosyncratic commuting utilities is denoted $g(D)$. D can be positive or negative.

Conditional on wage rate w, commuting time t and commuting cost c, workers choose the number of hours to work W by maximizing their utility function subject to time and money constraints:

$$
\begin{aligned}
u^*(w,t,c) = \quad & \max_W u(x,L) \\
& \text{such that} \\
& c + x \leq wW + Y \qquad \text{(budget constraint)} \\
& L + W + t \leq T \qquad \text{(time constraint)}
\end{aligned}
$$

Here, x is consumption, L is leisure, Y is fixed income, and T available hours per day. A worker will choose to commute if his maximal utility u^* is higher when working downtown than when working in the suburb, i.e. if $u^*(w,t,c) + D > u^*(w_0,0,0)$. A worker’s marginal monetary value of travel time savings will be w; this is a standard result that can be obtained by using the envelope theorem.

In this model, agglomeration effects will emerge: average wages and wage rates will increase when commuting costs and times decrease. Agglomeration effects are generated by three mechanisms. First, shorter commuting time will make commuters choose to work more hours, thus increasing average wages. This is a labour supply effect. Second, decreasing the commuting time or cost will make more workers choose to commute, which in itself increases average wage rates. This is a matching effect. Third, more workers choosing to commute will increase wage rates for all downtown workers. This is a spillover effect. If the meaning of “agglomeration effect” is restricted to refer only to wage *rates*, not wages, then only the two latter mechanisms will contribute.

In the following, two extreme cases of the model are presented. In the first version, there is no spillover effect among downtown workers, so agglomeration effects on wage rates are only caused by the matching effect. In the second version, there is no heterogeneity in worker productivity, so all workers get the same wage rate offer downtown. Agglomeration effects on wage rates are hence only caused by spillover effects among downtown workers. This will have very different implications for the share of agglomeration benefits captured by standard CBA.

---

zones to a single point means no loss of generality, so the current model is a simplification of a more general case with several zones (which is treated by simulation in a forthcoming paper). We assume throughout that the wage offer distribution is such that at least some workers get a sufficiently high wage rate offer to make them choose to commute.

It is important to stress that a potential modeller cannot observe ”commuters” and ”suburbians” separately – only aggregate numbers (average wage, VMT etc). This is a representation of how it would be in reality, where we would have a continuum, not two zones, and hence could not divide workers neatly into ”commuters” and ”suburbians”. We can only observe average wages, commuting distances etc. Even if we have micro-data, this changes only slightly, but the essential problems remain, in a sense which will become clear.

The model can easily be extended with heterogeneous commuting distances and endogenous employment. For example, idiosyncratic commuting utilities can be replaced by heterogeneous commuting distances. However, this does not add anything to the points I wish to make.

In both versions of the model, standard CBA benefits are calculated with the rule-of-a-half as follows. Let N_D and N_D' be the number of commuters before and after a change, and let $dN = N_D - N_D'$. Let $\overline{w}$ be the average value of time for commuters (this is equal to the average wage rate in this simple setting). Consider a marginal change of commuting time $dt=t-t'$ and of commuting cost $dc=c-c'$. Total benefits according to standard CBA, TB_{CBA}, are then defined as

$$TB_{CBA} = N_D(\overline{w}dt + dc) + ½dN(\overline{w}dt + dc)$$

Note the approximation that all travel time savings are valued with the average value of time for existing travellers, although time savings for *new* travellers should actually be valued with the average value of time of these new travellers. This approximation is usually defensible, although there are situations where it can matter.

Model 1: Agglomeration effects caused by matching

First, consider a version of the model where there is no income taxation, idiosyncratic commuting utilities or spillover effects on downtown wage rates. In this case, there will be a cut-off wage rate offer $\hat{w}$ above which workers will commute, and below which they choose to work in the suburb. Figure 1 plots total wage per worker as a function of the wage rate offer downtown, and what happens if commuting time decreases.

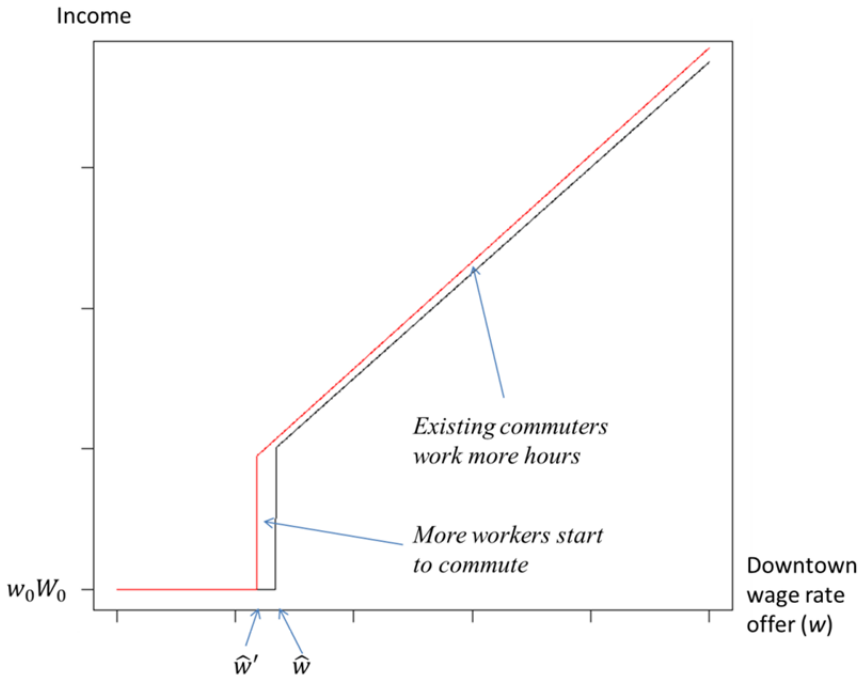

Figure 1. Total income for different levels of wage rate offers, before (black) and after (red) a reduction of commuting time.

The black line shows income per worker before the improvement. Below the cut-off wage rate $\hat{w}$, workers choose to work in the suburb at wage rate w_0 and working hours W_0, getting total income $w_0 W_0$. Workers who get wage rate offers above $\hat{w}$ choose to commute. Working hours for commuters depend on their wage rates, among other things. They can either increase or decrease in w, depending on the parameters of the model. Total income in the city is equal to the area under the black curve weighted by the density of wage rate offers $f(w)$.

If commuting time is reduced, the cut-off wage rate decreases from $\hat{w}$ to $\hat{w}'$, so more workers choose to commute. Existing commuters use some part of the travel time gain to work more hours, so their income increases as well. Hence, a reduction of commuting time increases total income through two mechanisms: a matching effect where $\hat{w}$ decreases and hence the overall average wage rate increases, and a labour supply effect, where all commuters work more hours. An observer of the city will hence conclude that there are agglomeration effects going on, both in terms of average wages and wage rates.

The aggregate monetary benefit (equivalent variation) of a reduction of commuting time dt can be written (with a slight approximation in the second term in the last equality[5], and denoting the marginal utility of income $\frac{\partial u^*}{\partial Y} = \lambda$ to save space)

$$TB = \int_{\widehat{w'}}^{\infty} \frac{1}{\lambda} u^*(w, t-dt, c) f(w) dw - \int_{\hat{w}}^{\infty} \frac{1}{\lambda} u^*(w, t, c) f(w) dw =$$

$$\int_{\hat{w}}^{\infty} w dt f(w) dw + \int_{\widehat{w'}}^{\hat{w}} \frac{1}{\lambda} u^*(w, t-dt, c) f(w) dw = N_D * \bar{w} dt + ½ dN * \hat{w} dt$$

[5] This is derived by Taylor expanding the first integrand around t and invoking the envelope theorem.

This is almost exactly equal to the benefits of standard CBA TB_{CBA}. The difference is only that the proper calculation of benefits TB values time savings for new commuters at $\hat{w}$ rather than $\bar{w}$ which TB_{CBA} does. This error is negligible in the numerical simulations below.

Hence, in this model all benefits are captured by standard CBA, *despite* the existence of agglomeration effects. The reason is that these agglomeration effects are caused by matching and labour supply effects – not by spillovers.

Model 2: Agglomeration effects caused by spillovers

Consider a second extreme case of the model. Assume that there is no heterogeneity in productivity among workers, and hence all workers get the same wage rate offer downtown. However, this wage rate increases with the number of downtown workers because of positive externalities, i.e. spillover effects. What makes some workers choose to commute and some not is their idiosyncratic utilities of commuting. Figure 2 illustrates how total income depends on idiosyncratic utilities D, and what happens if commuting time is reduced.

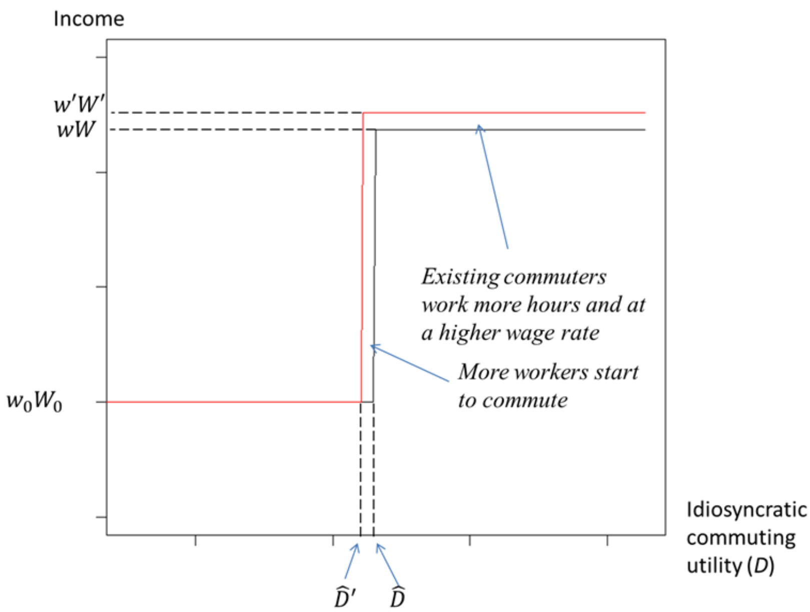

Figure 2. Total income for different idiosyncratic commuting utilities, before (black) and after (red) a reduction of commuting time.

Workers with idiosyncratic commuting utility less than $\hat{D}$ will choose to work in the suburb at wage rate w_0, working W_0 hours. Workers with idiosyncratic utility larger than $\hat{D}$ will choose to commute, and work W hours at wage rate w. This wage rate depends on the number of downtown workers, so there will be a dynamic process leading to an equilibrium. If commuting time is reduced, a new equilibrium will be reached with a new cut-off idiosyncratic utility $\hat{D}'$. As more workers choose to commute, the downtown wage increases to w'. This, and the shorter commuting time, changes the number of working hours downtown to W'.

In this version of the model, agglomeration effects in terms of average wage rates emerge because of spillovers in downtown. In addition, agglomeration effects on average wages emerge because of labour supply effects.

Just as in the previous version of the model, a relationship between accessibility and productivity emerges. However, the mechanism generating the agglomeration effect on the microlevel is no longer matching but spillovers. Contrary to the first version of the model, this means that some benefits will fall outside standard CBA, and there will in fact be “wider” economic benefits of a transport improvement. To see this, write the total benefits of a travel time reduction dt as follows (setting $dw = w' - w$ and $\widetilde{w} = ½(w + w')$, and using a slight approximation at the second equality sign)

$$TB = \int_{\widehat{D'}}^{\infty} \frac{1}{\lambda} u^*(w', t - dt, c) f(w) dw - \int_{\widehat{D}}^{\infty} \frac{1}{\lambda} u^*(w, t, c) f(w) dw =$$
$$= (Wdw + wdt)N_D + ½dN * \widetilde{w}dt = TB_{CBA} + (WN_D + ½dNdt)dw$$

The difference is that standard CBA benefits include the time saving for existing commuters, valued at the wage rate *before* the improvement, and half the time saving for new commuters, valued at the same wage rate. This neglects the wage rate increase dw for existing and new commuters – the last term. If the wage rate elasticity is high, this omitted benefit can be substantial.

Numerical simulations

Let us compare the two models using numerical illustrations. The point I wish to make is that the two versions of the model can be virtually indistinguishable when observing the city at an aggregate level. This implies that using aggregate observations, it is essentially impossible to determine the relative contributions from matching and spillovers, respectively. Hence, there is no way to determine the overlap between economic effects and standard CBA benefits.

Let the utility function be $u(x, L) = 0.5 \log(x) + 0.5 \log(L)$, and set T=16 hours, t=1 hour, c=5$ and w_0=5$/h. In the first model, let $f(w)$ be a uniform distribution between 5$/h and 10$/h. In the second model, let downtown wage be a constant-elastic function of the number of commuters with elasticity 0.25, and let $g(D)$ be a uniform distribution between -2.7 and 0.3 ($g(D)$ is calibrated to make the outcomes of the model similar).

Given these parameters, the model produces some sufficiently plausible numbers: workers work a little less than 8 hours per day on average, at an average wage rate of 7.32$/h in model 1 and 5.42$/h in model 2. The elasticity of travel (total kilometres travelled) with respect to travel time is around -0.2 in both models, broadly in line with empirical studies. The elasticity of the average wage rate (i.e. worker productivity) with respect to accessibility (generalized travel cost[6]) is 0.044 in model 1 and 0.047 in model 2, also broadly in line with empirical studies. Most importantly, the two elasticities are almost equal. In other words, the two models will produce agglomeration benefits of the same magnitude, and cannot be distinguished using aggregate data.

The last row shows the extent of the “wider” economic benefits for a reduction of commuting time with 20%. In model 1, standard CBA benefits approximate true benefits very well (the slight overestimation is because time savings for new commuters are valued with the average value of time for existing commuters). In model 2, on the other hand, standard CBA underestimate true benefits substantially: standard CBA benefits should be increased by 42% to reflect true benefits.

[6] The generalized travel cost is $GC = \overline{w}t + c$.

Table 1. Key characteristics of the two versions of the model.

	Model 1	**Model 2**
Mean wage rate ($/h)	7.32	5.42
Mean working hours (h)	7.86	7.97
Mean income ($/day)	57.41	43.12
Elasticity of travel wrt. time	-0.22	-0.23
Elasticity of mean wage rate wrt. accessibility	-0.044	-0.047
Wider economics benefits: benefits outside CBA relative to standard CBA benefits (TB/TB_{CBA}-1)	-1%	+42%

The two versions of the city are essentially indistinguishable on an aggregate level: key indicators are similar (and the models can be calibrated to make them even more similar if necessary), and most importantly, they exhibit similar elasticities in terms of travel and average wage rates. However, a standard CBA will capture virtually all benefits of a travel time reduction in one version of the city, but will substantially underestimate benefits in the other. This means that it is impossible to assess how much of productivity benefits that fall outside standard CBA, without knowing which microlevel mechanism generates the agglomeration effects – matching or spillovers. Even worse, these different microlevel mechanisms can generate agglomeration effects that on an aggregate scale are indistinguishable.

Income taxation

Next, let us introduce income taxation into the models. If income is taxed, then any transport improvement which affects total income will also affect tax revenues. However, this effect is seldom accounted for in standard CBA: standard practice is to capture the value of changes in generalized travel costs only through travellers' consumer surplus[7]. Ignoring changes in revenues from income tax is only correct if two conditions hold: 1) no part of a travel time reduction is used to work more (paid) hours; 2) reductions of generalized costs are not in any way used to reach higher-paying jobs, or go from unemployment to employment. Both assumptions seem unlikely, and there is considerable empirical evidence contradicting them. Still, the practice prevails, despite that already Forsyth (1980) pointed out that the value of time used in CBA should include a term capturing this effect (the marginal income tax rate multiplied by the share of a time saving that is used to increase income). This practice is likely due to the formidable empirical difficulties of establishing how an accessibility improvement affects aggregate income. However, this is essentially the same empirical difficulty encountered when trying to estimate the effect of accessibility on productivity, and considerable progress has been made in this area the last few years.

With a tax rate of $\tau/(\tau+1)$, the change in tax revenues will be τ times the change in total net income. Let φ be the share of a time saving that is used to work more hours (φ will vary across workers, depending on w and several other parameters). The increase in tax revenues TR of a travel time reduction dt in the two models will be (with self-explanatory notation)

$$TR_1 = \tau N_D \overline{w\varphi} dt + \tau dN * ½\hat{\varphi}\hat{w}dt$$
$$TR_2 = \tau(W\varphi dw + w\varphi dt)N_D + ½dN * \tilde{\varphi}\tilde{w}dt$$

[7] Other kinds tax and government revenue effects are accounted for in most CBA guidelines, however, for example changes in fuel tax revenues and public transport subsidies and fare revenues. The discussion here refers to revenues from income taxes.

Table 2 shows results of numerical simulation assuming an income tax rate of 30%. In model 1, increased tax revenues add 31% to the benefits relative to standard CBA. In model 2, the corresponding figure is 92% - that is, almost half of the benefits fall outside standard CBA.

Table 2. Tax revenues relative to other benefits in the two models.

	Model 1	**Model 2**
Tax revenues relative to standard CBA benefits	32%	62%
Tax revenues relative to total non-tax benefits	33%	48%
Wider benefits as share of standard benefits	31%	92%

The reason that so much of the total benefits fall outside standard CBA in Model 2 is of course the multiplicative effect of the omitted tax revenues and the omitted benefit of increased wage rate for existing commuters.

However, note that if tax revenues are included in the CBA in Model 2, all benefits will be captured. The conclusion from the above hence remains: the share of wider benefits captured in the CBA depends on whether the agglomeration effect is caused by matching or spillovers. But the simulation also illustrates the importance of taking changes in income tax revenues into account in situations when transport improvements affect total income.

3.2 The insufficiency of measuring generalized travel cost

The previous section explained the difficulties of assessing the share of agglomeration benefits captured in standard CBA. This section will explain an additional problem, namely that it matters in what way generalized travel costs change.

In standard CBA, all aspects of travel disutility are summarized in the generalized travel cost. The generalized travel cost reflects not only total travel time and travel cost, but also the comfort levels of different parts of the trip (through different valuation of various time components), risk for delays and sometimes even more dimensions such as safety. For example, waiting times are usually perceived as more onerous per minute than in-vehicle time, traveling in a crowded train is more onerous than traveling in a train with ample space, and so on. The generalized travel cost is an extremely powerful abstract concept, which enables analysts to calculate welfare changes not only of travel times and costs, but also of changes in comfort, reliability, crowding, service frequencies and so on. The generalized cost can also reflect that different socioeconomic groups may value certain factors differently – elderly people may value seating higher than younger people, for example. Generalized costs can be defined for individual travel modes, but can also be aggregated across modes (using logsums), so that a general, “multi-modal” or “mode-free” generalized travel cost is obtained. An important advantage is that the generalized cost concept easily handles situations where costs and times change in opposite directions, for example when analysing congestion pricing. In particular in complicated networks, where different generalized routes may have different travel times, travel costs and departure times (think of railway services or tolled road networks, for example), working at the level of generalized travel cost becomes necessary. In most real applications, specialized network programs are used to calculate generalized travel costs between pairs of origins and destinations, since the complexity quickly becomes too large to allow manual calculations.

The generalized travel cost is also the basis of detailed accessibility calculations, since it is, from a behavioural perspective, the perfect measure of the contact cost between two points in space. In fact, it is necessary to use generalized travel cost as the basis of calculations of how transport projects affect agglomeration effects, since measures such as geographical distance or total city size are very seldom affected by transport projects. Thus, in practice, agglomeration benefits of a transport project are calculated by applying some elasticity of wage (or wage rates) to the relative decrease in generalized costs that the project causes.

The generalized travel cost is hence the basic building block of transport CBA, both for calculating consumer surplus and for calculating agglomeration benefits, and is virtually indispensable especially in complex multimodal networks. However, all decreases in generalized travel costs are in fact not created equal when it comes to agglomeration benefits and changes in tax revenues. A trivial example is that a reduction of travel time frees up time that can be used for working more hours, which will increase tax revenues. Reductions of travel costs or increased trip comfort will not have the same kind of effect, although they increase the willingness to commute, and may hence increase the average wage rate. This has implications for the size of the wider economic benefits, as we will now illustrate.

To account for trip comfort, let us add a term to the worker utility function so it becomes $u(x,L) - \gamma t$, where γ reflects the direct disutility of traveling. With $\gamma>0$, traveling will be less enjoyable than spending time at work (in other words, the direct utility of working is normalized to zero). In the simulations, we set $\gamma = .02$, which will be equivalent to a disutility per minute of around 25% of the average wage rate of commuters. In the consumer surplus calculation, this changes the generalized travel cost to the sum of travel cost and the travel time weighted with the wage rate plus the direct disutility of traveling. This of course applies both to the exact measure and the approximation used in standard CBA, so in this sense the approaches do not differ.

First, consider model 1, in which all agglomeration benefits are captured in standard CBA. Columns 1-3 of Table 3 show summary results of three improvements which are equivalent from a generalized cost perspective. In column 1, travel time is reduced 20%; in column 2, the direct disutility of travel is reduced 20%; in column 3, travel cost is reduced 43%. Each of these improvements will reduce the generalized travel cost approximately 14%, and are hence equivalent from the average traveller's point of view and will give approximately equal traveller benefits.

Elasticities of travel and wage rate are broadly in line with empirical evidence. Note that the elasticity of mean wage rate is higher with respect to travel cost than with respect to travel time or travel disutility. This means that even in this starkly simplified setting, there is no single elasticity of productivity with respect to generalized cost; it matters which component of the generalized cost is changing.

However, it is the increase in tax revenues that differ most substantially between the three types of improvements. Reducing the travel time increases tax revenues 2.5 times more than reducing travel disutility. The reason is of course that the travel time gain is partly used to work more hours, which increases tax revenues. Reducing the travel cost actually *reduces* tax revenues. This is because the net income of existing commuters is increased by this, causing them to reduce their working hours (note that the elasticity of working hours with respect to travel cost is positive), and this more than offsets the increase in average wage rate due to more workers choosing to commute.

Table 3. Changes in tax revenues in Model 1 for changes in travel time, travel disutility and travel cost which are equivalent from a generalized travel cost perspective.

Reduction of:	Travel time (t)	Travel disutility (γ)	Travel cost (c)
Elasticity of travel	-0.31	-0.30	-0.19
Elasticity of mean wage rate	-0.06	-0.06	-0.08
Elasticity of work hours	-0.06	0.00	0.10
Increased tax revenues, relative to traveller benefits	30%	13%	-6%
Wider economics benefits: benefits outside CBA relative to standard CBA benefits	28%	19%	-7%
Wider benefits IF tax revenues are included in the CBA	-1%	5%	-2%

Hence, although these three improvements are equally beneficial from the point of view of the travellers (and therefore standard CBA), their total social benefits vary considerably. However, *if* the increase in tax revenues is included in the CBA, then the problem disappears. The difference between the rule-of-a-half approximation used in CBA and exact benefits is negligible, once tax revenues are taken into account. The point is that tax revenues will increase with different amounts for different types of reductions of the generalized travel cost.

Next, consider model 2, in which substantial parts of the agglomeration benefits fall outside standard CBA. Table 4 shows numerical simulations for the same improvements as above.

Table 4. Changes in tax revenues in Model 2 for changes in travel time, travel disutility and travel cost which are equivalent from a generalized travel cost perspective.

Reduction of:	Travel time (t)	Travel disutility (γ)	Travel cost (c)
Elasticity of travel	-0.37	-0.37	-0.17
Elasticity of mean wage rate	-0.064	-0.063	-0.061
Elasticity of work hours	-0.010	0.002	0.017
Increased tax revenues, relative to traveller benefits	49%	37%	22%
Wider economics benefits: benefits outside CBA relative to standard CBA benefits	128%	110%	81%
Wider benefits IF tax revenues are included in the CBA	30%	34%	37%

Despite standard CBA benefits being approximately equal compared to Model 1, several things change. Some are simply due to parameter choices, such as the slight changes in elasticities of travel and wage rate. It is interesting, however, that the wage rate elasticity is no longer higher with respect to travel cost than with respect to travel time.

The main observation, however, is that it matters which component of the generalized cost that change. Just as before, substantial benefits fall outside standard CBA, even if tax revenues are included – and the relative sizes of the wider benefits are different in the three cases. In other words, it is not enough to know the change in total generalized cost to compute the wider benefits.

4 IMPLICATIONS FOR APPLIED CBA

In an influential paper, Venables (2007) showed that standard CBA could omit substantial benefits in the presence of agglomeration effects. In his model, workers can either live in the city and work in the CBD, or live and work outside the city. The wage rate is higher in the city, and increases with the city population; all city workers get the same wage. Workers choose to move to the city if the higher wage is enough to compensate the costs incurred by living in the city (commuting costs etc.). The model is similar to the second version of the model in this paper, where the downtown wage rate increases due to external spillover effects[8]. The essential feature of both models is that agglomeration effects are generated by spillovers. As was shown above, this means that there will be benefits outside standard CBA, i.e. wider economic impacts.

Several countries have in recent years developed methods to capture such wider impacts. This section discusses the implications of the conclusions above for two such methods, namely the Swedish and the English methods.

Swedish WEI methodology

Sweden uses a relationship between accessibility and average wage per worker, estimated using differences-in-differences on accessibility and average wage per zone at two points in time, controlling for socioeconomic variables (Anderstig, Berglund, Eliasson, Andersson, & Pyddoke, 2012). The relationship is used to evaluate transport projects by calculating the improvement of accessibility from the project, and then use the relationship to calculate an increase in aggregate wage per zone. It is an unresolved issue to what extent this should be added to the standard CBA. The current guideline recommends subtracting the consumer surplus from the wage effects, adding the remainder of the wage effects. This is equivalent to assuming that the reduction of generalized travel costs is completely used for working more hours or reaching a higher-paying job; nothing is used for e.g. more leisure or moving to a more attractive residential location.

Figure 3 describes the problem of overlap between wage effect and consumer surplus in simplified form. The consumer surplus consists of two parts: reductions of the current generalized cost (the upper part of the consumer surplus), and a second part where the lower generalized cost has been traded in exchange for longer trips (the lower part of the consumer surplus), making it possible to reach for a better residence and/or a more enjoyable and better paid job. The reduction in generalized cost can consist of lower travel costs, a more comfortable journey, and a shorter travel time. The shorter travel time can, in its turn, be exchanged for a combination of more leisure and more hours worked. Two parts of the consumer surplus are included in the wage effect: more hours worked and the higher wage rate. The wage effect, however, also include *external* wage rate effects, i.e. a possible increase in other workers’ wages because of spillover effects. It also includes increased tax revenues, both from the internal and external wage effects.

[8] The main difference between Venables’ model and Model 2 in this paper is technical and not important for the conclusions: it is the mechanisms making some workers choose not to work in the CBD/downtown. In the Venables model, the cost of living in the city (commuting and residential costs) increases with the city population, until workers are indifferent between living in the city or outside. In the model in this paper, it is idiosyncratic preferences (perceived commuting costs) that cause some workers to commute and others not to. Idiosyncratic commuting preferences can easily be replaced by heterogeneity in commuting distance, for example.

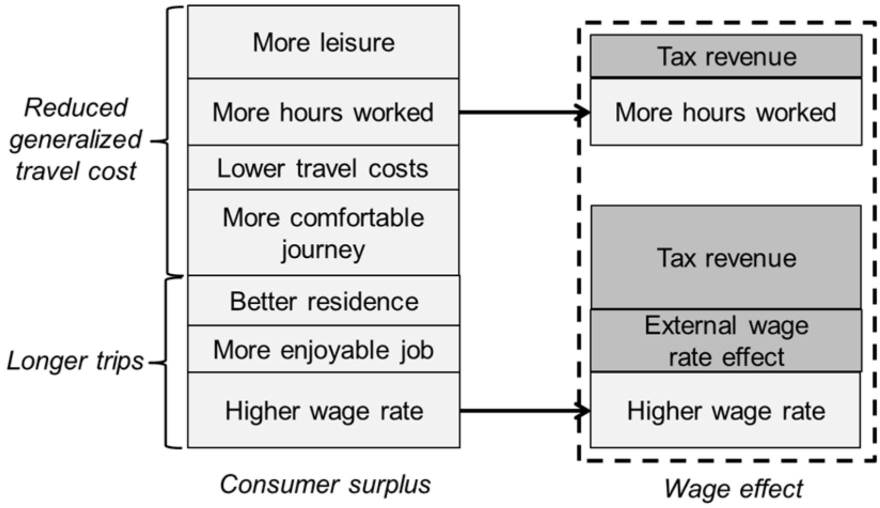

Figure 3. Consumer surplus, wage effect, and the overlap.

The “overlap problem” is that there is no way to know to what extent the wage effect and the consumer surplus overlap, since we do not know the components of them, only their total size.

Experience from using the relationship indicates that the wage effect increase is roughly equal to the consumer surplus multiplied by one plus the tax revenue. In other words, the size of the wage effect is as though the entire consumer surplus is exchanged for a higher wage through increased working hours and/or wage rates, but without any external increases in wage rates. Of course, there is no way to know whether this is the correct interpretation: it might just as well be that nearly all of the wage effect is external to the worker and should hence be added to the standard CBA.

UK WEI methodology

The UK guidelines (Department for Transport, 2013) are by far the most ambitious, and have also been emulated by several other countries. The guidelines distinguish between four kinds of wider impacts, of which only one fall within the scope of this paper[9]: the increase in (firm) productivity from higher accessibility. The relationship between accessibility and productivity comes from pioneering work by Daniel Graham, summarized in Graham (2007). The CBA guidelines then assume that the increase in productivity is completely external to workers, so all productivity benefits should be added to the CBA. (Note that this is the complete opposite to the Swedish assumption.)

This is a difficult issue which deserves to be described in some detail. The basis is an estimated relationship between firm outputs (measured as their turnover) and a number of explanatory variables, of which the most important are labour input and “effective density”, which is essentially an accessibility measure. Simplified down to the essentials, the relationship can be written as

$$Y \sim L^{\alpha} A^{\beta}$$

[9] The other three are WI2) increased production in imperfectly competitive markets, which is calculated by a markup of 10% on business travel time savings. I have been unable to find the underpinning of this rule-of-thumb. WI3a) increased tax revenues from decreased unemployment WI3b) increased firm productivity from firm relocation. To calculate this, an integrated land use/transport model must be used to calculate firm relocation, after which relocating firms are supposed to get the same productivity as firms that had already located in a certain area.

Where Y is output, L is labour and A effective density. The difference between this relationship and the previous discussion in this paper is that the dependent variable is not average wage or wage rate, but firm output (or more precisely turnover). The question is now to whom benefits accrue of an increase in firm output due to increased accessibility. If such benefits are in no way passed on to workers, but instead accrue to e.g. shareholders and/or customers, then the issue is settled: agglomeration benefits are indeed completely external to workers, and hence fall outside standard CBA. This seems to be a rather strong assumption, however, in particular since the original estimation equation does not control for wage rates of different firms. If productivity benefits in fact are at least partly passed on to workers, in the form of higher wages, then the problems illustrated in this paper remain: there is an unknown overlap between consumer surplus and the agglomeration benefits.

From the point of view of the model used in this paper, benefits accruing to shareholders or customers is equivalent to a tax from the point of view of workers. The difference is that in reality, it is difficult to empirically observe the equivalent of the “tax rate”, i.e. the share of benefits which go to shareholders and consumers. This means that there are now two difficult questions: first, to what extent agglomeration effects are caused by matching, and second, to what extent benefits of firm productivity are passed on to workers as wages.

In standard CBA, which tacitly assumes that distortions can be ignored, it does not matter how benefits from increased production are divided between workers, firms and consumers: it is enough to know their total size, so they can be calculated as if they were captured by travellers’ consumer surplus. Likewise, it does not matter in standard CBA how travellers spend a decrease in generalized cost – whether to work more hours, increase the wage rate, live further away or get more leisure. But in the presence of distortionary taxes and external productivity effects mean, the division of the spoils start to matter.

5 CONCLUSIONS

There is a growing and maturing literature on the economic benefits of improvements in accessibility. However, the question remains to what extent such economic benefits are already included in standard CBA. This is the “overlap” problem: to what extent standard CBA, which aims to capture not just travellers’ economic gains but also e.g. increased leisure, overlaps with calculations of economics benefits, which aims to include not just traveller benefits but also benefits accruing to others.

In this paper, I show that answering this question requires knowledge of what micro-level mechanisms generate the agglomeration effects. If agglomeration effects are generated by matching – more efficient pairing of employers and employees – then all benefits are in fact included in standard CBA, provided that changes in tax revenues are accounted for. (This last provision, however, is an important caveat: currently, there is very little evidence as to how to determine changes in tax revenues from an accessibility improvement.) If agglomeration effects are generated by various kinds of spillovers, or if the economic benefits of increased productivity completely accrue to others than workers, then agglomeration benefits fall outside standard CBA. Currently, only limited evidence exists to resolve this question.

6 ACKNOWLEDGMENTS

I am grateful to Maria Börjesson, Gunnar Isacsson and Sergio Jara-Díaz for useful discussions.

7 REFERENCES

Anderstig, C., Berglund, S., Eliasson, J., Andersson, M., & Pyddoke, R. (2012). *Congestion charges and labour market imperfections:“Wider economic benefits” or “losses”?* (No. 2012:4). CTS Working Paper. Centre for Transport Studies, KTH Royal Institute of Technology.

Combes, P.-P., Duranton, G., & Gobillon, L. (2008). Spatial wage disparities: Sorting matters! *Journal of Urban Economics*, *63*(2), 723–742.

Department for Transport. (2013). *WebTAG: TAG unit A2-1 wider impacts - Publications - GOV.UK*.

Duranton, G., & Puga, D. (2004). Chapter 48 Micro-foundations of urban agglomeration economies. In J. Vernon Henderson and Jacques-François Thisse (Ed.), *Handbook of Regional and Urban Economics* (Vol. Volume 4, pp. 2063–2117). Elsevier.

Forsyth, P. J. (1980). The Value of Time in an Economy with Taxation. *Journal of Transport Economics and Policy*, *14*(3), 337.

Gould, E. D. (2007). Cities, Workers, and Wages: A Structural Analysis of the Urban Wage Premium. *The Review of Economic Studies*, *74*(2), 477–506.

Graham, D. J. (2007). Agglomeration, productivity and transport investment. *Journal of Transport Economics and Policy*, *41*(3), 317–343.

Graham, D. J., Melo, P. S., Jiwattanakulpaisarn, P., & Noland, R. B. (2010). Testing for Causality Between Productivity and Agglomeration Economies. *Journal of Regional Science*, *50*(5), 935–951.

Marshall, A. (1890). *Principles of Economics*. London: Macmillan and Co., Ltd.

Melo, P. C., & Graham, D. J. (2014). Testing for labour pooling as a source of agglomeration economies: Evidence for labour markets in England and Wales. *Papers in Regional Science*, *93*(1), 31–52.

Puga, D. (2010). The Magnitude and Causes of Agglomeration Economies*. *Journal of Regional Science*, *50*(1), 203–219.

Rosenthal, S. S., & Strange, W. C. (2004). Chapter 49 Evidence on the nature and sources of agglomeration economies. In J. Vernon Henderson and Jacques-François Thisse (Ed.), *Handbook of Regional and Urban Economics* (Vol. Volume 4, pp. 2119–2171). Elsevier.

Smith, A. (1776). *An Inquiry into the Nature and Causes of the Wealth of Nations*. London: Methuen & Co., Ltd.

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*(2), 173–188.