The impact of land use effects in infrastructure appraisal

Eliasson, Jonas and Savemark, Christian and Franklin, Joel

Linköping University, KTH Royal Institute of Technology, KTH Royal Institute of Technology

2020

The impact of land use effects in infrastructure appraisal

Jonas Eliasson[1], Christian Savemark[2], Joel Franklin[2]

[1]Linköping University
[2]KTH Royal Institute of Technology

Abstract
When benefits of proposed infrastructure investments are forecasted, residential location is usually treated as fixed, since very few operational transport models are able to forecast residential relocation. It has been argued that this may constitute a source of serious error or bias when evaluating and comparing the benefits of proposed infrastructure investments. We use a stylized simulation model of a metropolitan region to compare calculated benefits for a large number of infrastructure investments with and without taking changes in residential location into account. In particular, we explore the changes in project selection when assembling an optimal project portfolio under a budget constraint. The simulation model includes endogenous land prices and demand for residential land, heterogeneous preferences and wage offers across residents, and spillover mechanisms which affect wage rates in zones. The model is calibrated to generate realistic travel patterns and demand elasticities. Our results indicate that ignoring residential relocation has a small but appreciable effect on the selected project portfolio, but only a very small effect on achieved total benefits.

Keywords: Cost-benefit analysis, land use, wider impacts, land use/transport interaction models.

JEL Codes: R41, R43, R48.

1 INTRODUCTION

Proposed infrastructure investments are often evaluated ex ante using transport models. The transport model is used to construct a cost-benefit analysis (CBA) for the investment by calculating demand and generalized travel costs with and without the investment. Based on these results, social benefits (consumer surplus, producer surplus etc.) are calculated and compared with investment and maintenance costs.

In practice, transport models usually treat residential location as fixed: the number of residents per zone and their socioeconomic characteristics is treated as an exogenous input to the model. Job location, on the other hand, is often allowed to implicitly vary by using a destination choice model for work trips, implicitly unconstrained. While there is a fair number of integrated land use and transport models able to model changes of residential location (see reviews by e.g. Iacono et al. (2008), Moeckel et al. (2018) and Acheampong and Silva (2015)), few if any have seen widespread use in practice: when it comes to applied infrastructure appraisal, standard transport models with fixed residential location is still the norm. Since infrastructure investments in reality may induce substantial changes in residential location, there are reasons to suspect that ignoring this may constitute a source of serious error or bias in investment appraisal. This may in turn lead to considerable losses in achieved social benefits if the choice of what investments to build is based on biased analyses. A typical situation in national transport investment planning is that the planning agency faces hundreds or even thousands of potential investment candidates, of which a relatively small number can be afforded due to budget constraints – hence the need for consistent and objective appraisal methods which can be efficiently applied on a large scale. This also explains why extensive sensitivity analyses, for example of different land use scenario inputs, is usually practically infeasible.

In this paper, we explore this issue through a simulation approach. We use a stylized simulation model of a metropolitan region to compare CBA results for a large number of infrastructure investments with and without taking changes in residential location into account. Our focus is on three main questions: 1) whether ignoring residential relocation is a serious source of error in the calculated social benefits; 2) if so, whether this leads to substantial changes in what investments are selected for construction (considering that the total budget for investments is limited); and, 3) if so, whether this leads to an appreciable loss in total achieved benefits from the assembled portfolio of investments.

Our simulation model predicts not only location and workplace choices and their implied commuting patterns, but also certain general equilibrium effects (sometimes called "wider economic benefits" in CBA contexts): it includes endogenous land prices and demand for residential land, endogenous labor supply, heterogeneous preferences and wage offers across residents, and a spillover mechanism affecting the wage rates in different zones. Total economic output in the simulated region is affected by infrastructure investments through three mechanisms: endogenous labor supply, the quality of matching of workers and jobs, and spillovers affecting wage rates in each workplace zone. The model is calibrated to generate realistic travel patterns, residential densities and demand elasticities. Just as the majority of operational land use/transport interaction models (such as MEPLAN, UrbanSim or Tranus), it simulates a large metropolitan region. Our focus is hence on effects of infrastrcuture on location and land use within such a region, rather than on location effects on a national or even international scale.

The model simulates the very long run, in which residential location and densities are allowed to vary freely, i.e. there is no dependence on historic land use patterns. Clearly, this is different from reality, where existing houses and location patterns change mean that location changes only slowly. However, adopting the extreme assumption that location can change freely is useful for establishing an upper bound on how large the errors of ignoring land use changes might be, since in reality the land use effects will be smaller due to existing house and historic dependencies.

To check the robustness of our conclusions, we analyze four different stylized "metropolitan regions", i.e. model setups with different geographies, networks, location patterns and distributions of random parameters.

Our results indicate that ignoring residential relocation may sometimes induce appreciable errors in the benefit calculations, but the effect on investment selection is comparatively small, and the effect on achieved total benefits is virtually negligible. The intuition explaining this perhaps counterintuitive result is that the variation in benefits between projects is much larger than the errors in the benefit calculations, meaning that the project ranking (in terms of benefit-cost ratios) is fairly stable with respect to variations in the benefit calculations.

Of course, this does not mean that modelling residential location and land use is unimportant: such changes are interesting and important in their own right and for several other reasons (Eliasson & Mattsson, 2000, 2001; Waddell, 2011). Indeed, our own backgrounds in land use/transport interaction modeling have made us appreciate the insights such models can bring, which often go far beyond merely appraising the aggregate benefits of infrastructure investments (Eliasson, 2010; Eliasson & Martìnez, 2001; Eliasson & Mattsson, 2000; Franklin et al., 2002; Waddell et al., 2007). Moreover, if the analyst has access to a model able to forecast land use effects, and the cost and effort to run such a model is not prohibitively high, then it is clearly preferable to take land use effects into account when evaluating investment benefits. However, integrated transport/land use models are still rare, and they are often more costly to run than standard transport models in terms of data preparation and running times, so an overwhelming majority of transport projects are evaluated using transport models where residential land use is fixed. This makes the question of whether this is likely to induce severe errors and large losses of benefits in the project selection worthwhile to explore. In this sense, the focus of the present paper is not on the land use effects as such, but on infrastructure appraisal – specifically, to what extent appraisal and project ranking is biased when land use effects are ignored. In certain cases, in particular for large projects yielding extremely large changes of accessibility, land use changes may be particularly important; for example, establishing the US interstate highway network after World War II, or the UK railway network in the 19th century. The focus of the present paper, however, is on selecting projects from a long list of relatively modest improvements of an existing network, which is a far more common situation for transport planners than the aforementioned ultra-large investment programmes. Moreover, our model emulates effects of infrastructure investments within a (large) metropolitan region, rather than national infrastructure that connects different regions or cities within an entire country.

Our conclusions are similar to those reached by Börjesson et al. (2014), who study the same issue but based on a handful of real infrastructure investments. The present study is different in that it is based on a simulation model which includes endogenous labor supply, wage rates, land prices and land markets, and most importantly compares a much larger number of (hypothetical) investments – hundreds of links in four different generic city simulations. The question in this paper is analogous to earlier papers studying the robustness of cost-benefit analysis with respect to other kinds of uncertainties, e.g. parameters, demand forecasts and scenario assumptions (Asplund & Eliasson, 2016; Börjesson, Eliasson, et al., 2014; Eliasson & Fosgerau, 2013)

The model we use is an extension of the theoretical framework used in Eliasson and Fosgerau (2019), in that labor supply (working hours), residential location and lot sizes are now all endogenous. We extend the results of Eliasson and Fosgerau (2019) in the sense that we show analogous results about the relation between conventional CBA benefits and so-called "wider benefits" for our extended model.

2 METHOD

The core question of this paper is to what extent the common practice to ignore residential relocation when forecasting the social benefits on transport investments affects benefit estimates, project selection and achieved total benefits from the selected project portfolio. Our approach can be summarized as follows.

We use a generic simulation model of a metropolitan region with a flexible geography, calibrated to generate realistic outcomes in terms of travel patterns and elasticities. The simulation model calculates residents' choices of residential location and job location endogenously, as well as lot sizes, working hours and land prices. To assess the robustness of our conclusions, we simulate four different "metropolitan regions" (model setups), with different geographies, networks and distributions of preference parameters and wage rates. For each such metropolitan region (or sometimes "city" for short), we simulate the effects of generic infrastructure investments, each of which is assumed to reduce the travel time on a link by 10%. Each investment project hence corresponds to an improvement of one (and only one) link. For each such generic investment, social benefits are calculated as described below. The corresponding investment cost is assumed to be proportional to the link length. The model simulation where transport improvements *do* affect residential location are considered to be the "true" outcome, and we want to compare this to restricted simulations where residential location is kept fixed.

Our main focus is how the bias, or "prediction error", resulting from keeping location fixed affects decisions about what investments to make. To simulate this, we consider the following situation. A planner wants to maximize total social benefits under a given investment budget. For each investment, the planner knows the investment cost (which we assume is proportional to the link length) and calculated benefits from the fixed-location model. Based on this information, the planner chooses the set of investments that gives highest total benefits[1]. However, in reality residential location is *not* fixed, so the planner's selection of investment is not quite the optimal one: had the planner had access to the true benefits, the planner would have selected different investments, thereby achieving a higher total benefit. Our main question is then what the difference is between the achieved total benefit and the maximal possible benefit. This measures how much benefits that are lost from not having access to the true benefits, that is, from using a prediction model with fixed residential location. The "fixed residential location" simulation may also be interpreted as a medium-term equilibrium, if job locations and wage rates change quicker than residential location and lot sizes.

[1] Of course, infrastructure decisions are determned by many other factors in addition to CBA results, but several studies have shown that there is often at least some correlation between CBA results and project selection (Bondemark et al., 2020; Eliasson et al., 2015; Eliasson & Lundberg, 2012; Nellthorp & Mackie, 2000).

Section 2.1 describes this idea in precise terms. Section 2.2 describes the simulation model, and section 2.3 describes the four different geographies and networks we will use in the simulations. Section 2.4 describes how the model is amended to keep residential land use fixed while still maintaining equilibrium conditions.

2.1 The planner's problem

Consider a planner wanting to maximize total social benefits under a budget constraint by selecting a number of investments from a set of candidates. Each investment candidate corresponds to an improvement of one link. Call the investment candidates $k = 1, \dots, K$, the true investment benefits B_k and the corresponding investment costs C_k. Call the planner's budget C, and let $\{\delta_k\}$ be a set of indicator variables that are 1 if investment k is chosen and 0 otherwise. (For technical reasons noted below, we assume that fractions of an investment can be chosen so δ_k can be a fraction between 0 and 1; this is only relevant for the last investment to be included, however.) The optimal selection of investments is the solution to the optimization problem

$$\begin{gathered}\max_{\{\delta_k\}} \Sigma_k \delta_k B_k \\ \text{such that} \\ \Sigma_k \delta_k C_k \leq C \\ \delta_k \in [0,1] \; \forall k\end{gathered} \tag{1}$$

Call the solution to the problem $\boldsymbol{\delta}^* = \{\delta_k^*\}$, and the optimal value (the maximal attainable total benefit) $B^* = \Sigma_k \delta_k^* B_k$. This optimization problem is called a knapsack problem, and it is well known that the solution is obtained by ordering the investments according to the ratio B_k/C_k in descending order, and then selecting from the top of this list going down until the budget constraint C is met. The variables δ_k will hence always be 0 or 1 even if they are allowed to be fractions in the problem formulation, except for the last (marginal) investment where δ_k will be a fraction. It is conventional to allow fractional investments in this way for several reasons, most importantly that it makes B^* a continuous function of C, and that it ensures that for two budget constraints $C_1 < C_2$, the solution $\boldsymbol{\delta}^*(C_1)$ will be a strict subset of the solution $\boldsymbol{\delta}^*(C_2)$.

The planner, however, does not know the true benefits B_k, but only their approximate values B'_k, coming from the fixed-location simulation model. So the planner will select the set of investments $\boldsymbol{\delta}' = \{\delta'_k\}$ which solves the optimization problem (1) but with the true benefits B_k replaced with the approximate values B'_k. This means that the planner achieves less total benefits than would be maximally attainable, since the project selection is based not on the real benefits B_k but their approximate values B'_k. The relative loss of benefits due to the imperfect information available to the planner is

$$\Delta B = 1 - \frac{\Sigma_k \delta'_k B_k}{\Sigma_k \delta_k^* B_k} \tag{2}$$

We will also explore how much the choice of investments differs between the optimal choice $\boldsymbol{\delta}^*$ and the planner's choice $\boldsymbol{\delta}'$. In a way, this is actually less important, since what matters most is the difference in achieved total benefit ΔB – but in a real case, it might matter a great deal precisely what investments are chosen. Hence, it is interesting to explore how big the difference is between the two choices. Let S be the share of investments from the optimal selection that are excluded from the planner's selection:

$$S = 1 - \frac{\Sigma_k \delta_k^* \delta'_k}{\Sigma_k \delta_k^*} \tag{3}$$

It is worth pointing out that the problem formulation above uses the conventional approximation that project benefits are independent, i.e. that each project benefit is constant and does not change depending on the combination of projects that are chosen. This is clearly a simplification, but an ubiquitous one in applied planning, since the number of project combinations quickly becomes too large to be feasible to handle (although there are attempts at solution algorithms, such as Szimba and Rothengatter (2012) and Iniestra and Gutíerrez (2009)). The simplification is usually defended by noting that in practice, the task is to improve a very small share of links in a large network, so interdependencies between projects can be assumed to be small. Cases where interdependencies are clearly substantial, for example where projects depend on each other, are handled by grouping projects together into larger projects. How large errors in project selection this simplification really leads to in practice seems to be an underresearched issue, but it is out of the scope of the current paper since our focus is on exploring the size of errors in benefit calculation and the subsequent stability of the project ranking. Taking project interdependencies into account might affect the optimal project selection, however, and our simplifying assumption that project benefits are independent should be kept in mind as a caveat when interpreting results.

2.2 The simulation model

The simulated region is divided into zones containing both residents and workplaces. The zones are connected through a transport network. In the simulations, the geography of the region (zones and network) can be generated randomly, which allows us to check the robustness of our results with respect to different draws of the geometry's random parameters.

Together, the geography and the network generate travel times t_{ij} and travel costs c_{ij} between each pair of zones (i, j). Prior to investments, travel speed is constant and equal everywhere[2], so travel times and travel costs are proportional to link lengths.

We do not model the housing construction sector, so there is no dependence on historic land use patterns such as if existing buildings would be costly to demolish or change, and new buildings costly to build. Had such costs been included, land use and location effects would tend to be smaller. In this sense, our results represents upper bounds of the effects. The total amount of land is fixed, but in principle the demand for land in a zone may be zero, so the model allows for the total residential area to vary. However, since we use a probabilistic model and endogenous land prices, demand for land will not go to exactly zero in the simulations, although residential density can become arbitrarily low.

Travel patterns, land use and location is only determined by the exogenous travel times, travel costs and baseline wages rates, plus the endogenous land prices and wage rates. The model is a steady-state equilibrium model, so we do not model the temporal dimension or the process of going from one equilibrium to another. Clearly, these are simplifications: in reality, people's travel and location choices are affected by many more

[2] Having different but constant speeds on different links would be possible but would not add any furter insights. Modeling endogenous congestion, however, could be interesting. But we prefer not to do this, since this would introduce yet another layer of complexity when interpreting results. As will be shown in the results section, a main result is that project ranking is relatively stable. It is hard to see why this would be different with endogenous congestion in the model, since endogenous congestion tends to cancel out some part of a link improvement. If there *was* endogenous congestion in the model, however, this would cast some doubt on whether this result is valid even in the absence of congestion: it might be that project ranking is stable just because endogenous congestion cancels out some of the project benefits by "moving congestion around" in the network.

variables than just wages, prices, travel times and travel costs, and location and land use change very slowly, and these caveats should be kept in mind. Another simplification is that we do not model firms' or employers' behaviour, and we use the simplification (common in economic modeling) that wages simply reflect worker productivity. However, using a simple model setup has the advantage that its structure can be explained and calibrated relatively easy.

Workers' choice of residence, workplace, working hours and lot size

The population is fixed. Workers are divided into classes $n = 1 \dots N$ where each class has different preference parameters and wage rates. We treat each class as a continuum and normalize the number of workers per class to 1. Workers of class n optimize their utility by choosing residential zone i, workplace zone j, working hours[3] and lot size. The exposition is simplified if this is described as a two-step optimization problem in the following way. First, conditional on a residence-workplace zone pair (i, j), the optimal number of working hours W_{ij}^n and the optimal lot size L_{ij}^n are solved for, taking the wage rate w_j^n and the land price p_i as given (both wage rates and land prices are in fact endogenous):

$$\begin{aligned} v_{ij}^n = \quad & \max_{W,L,l,x} u(l, x, L, t_{ij}) && (4) \\ & \textit{such that} \\ & l + t_{ij} + W = T && \text{(time constraint)} \\ & x + c_{ij} + p_i L = (1-\tau) w_j^n W + Y && \text{(budget constraint)} \end{aligned}$$

Here, $u(\dots)$ is the worker's utility[4] function, l is leisure, x is consumption of a generalized good with unit price, T is the number of hours per day, τ is the income tax rate, and Y is exogenous income (this is convenient for expressing the marginal utility of income later on). The optimal value of this problem is v_{ij}^n, i.e. the maximal utility given residence zone i and workplace zone j. Then, the utility of choosing the residence-workplace zone pair (i, j) is

$$\tilde{v}_{ij}^n = v_{ij}^n + \xi_{ij}^n \qquad (5)$$

where ξ_{ij}^n is a random term capturing idiosyncratic taste preferences. In the simulation model, we assume that $\xi_{ij}^n = H_i^n + D_j^n + \varepsilon_{ij}^n$. H_i^n and D_j^n are constants reflecting the characteristic amenities and attractiveness of each residence and workplace zone. They are drawn from normal distributions, and kept constant across scenarios. ε_{ij}^n is a random term, different for each individual in the class (individuals are treated as a continuum), which is assumed to be redrawn when the scenario changes, e.g. following a transport investment. (The random terms are assumed to be Gumbel distributed in the simulations, yielding logit choice probabilities.) Individuals in the class choose the

---

[3] Clearly, it is a simplification that workers can choose working hours freely. However, in most countries a substantial share of the population does not work the standard "full working hours" (often 40 hours per week), and it is a common empirical finding that better accessibility tends to increase labour supply. It is this labor supply effect that is modeled by assuming that workers can choose working hours.

[4] At this stage we do not assume that this function is money-metric; in other words, we allow for varying marginal utilities of income. However, a cornerstone assumption in conventional CBA (and most applied welfare analyses) is to treat monetary transfers as welfare neutral, that is, ignoring differences in marginal utilities of income across society. One way to introduce this assumption is to define the utility function as money-metric; but we prefer to introduce that assumption at a later stage, since this makes the assumption and its consequences more transparent.

residence-workplace pair which maximizes $\tilde{v}_{ij}^n$. Call the expected achieved utility $V^n = E\left(\max_{ij} \tilde{v}_{ij}^n\right)$, and the expected share of class n choosing residence-workplace (i, j) P_{ij}^n.

Determining wage rates and land prices

To emulate external agglomeration effects in the model, wage rates are assumed to increase with the number of workers in the zone[5]. This emulates spillover effects such as sharing and learning between workers and firms. The wage rate of class n in zone j is the product of a constant baseline wage rate $\widehat{w}_j^n$ and an agglomeration factor which is an increasing function of the number of workers in the zone:

$$w_j^n = \widehat{w}_j^n * \left(1 + \beta \Sigma_{ni} P_{ij}^n\right)^\eta \tag{6}$$

The parameter η hence determines the strength of the spillover mechanism (β is just a scaling parameter). Note that the baseline wage rate is different for each class and each workplace zone (drawn from a truncated lognormal distribution in the simulations); this generates a matching effect since workers are willing to commute longer to achieve a higher wage.

The choice probabilities P_{ij}^n are hence a function of wage rates $\boldsymbol{w} = \{w_{ij}^n\}$ and land prices $\boldsymbol{p} = \{p_i\}$. Wage rates and land prices are endogenous, determined such that the following equilibrium conditions hold (letting $£_i$ be available residential land in zone i):

$$\begin{gathered} \Sigma_{nj} P_{ij}^n(\boldsymbol{w}, \boldsymbol{p}) L_{ij}^n = £_i \quad \forall i \quad \text{(land market equilibrium)} \\ w_j^n = \widehat{w}_j^n * \left(1 + \beta \Sigma_{in} P_{ij}^n(\boldsymbol{w}, \boldsymbol{p})\right)^\eta \quad \forall j, n \quad \text{(wage rate equilibrium)} \end{gathered} \tag{7}$$

Firms are assumed to operate under constant return to scale, and wages are assumed to be equal to worker productivity. Hence, all changes in productivity are transferred to workers as wage changes. Firm location is not modeled explicitly; it is reflected implicitly in the sense that the number of jobs in each zone is determined by workers' choice of job location.

The effect on economic output of a transport improvement

Wage rates are assumed to equal workers' productivities, so the total economic output Z is equal to total wages:

$$Z = \Sigma_{nij} P_{ij}^n W_{ij}^n w_{ij}^n. \tag{8}$$

This generates agglomeration effects on an aggregate scale, in the sense that higher accessibility will tend to increase economic production[6]. A reduction of travel times or travel costs will affect economic output through three mechanisms: the number of working hours may change (labor supply effects); workers may choose to change

[5] An alternative way to model this would be to let wage rates depend on the accessibility from one zone to workers in *all* other zones, not just on the number of workers in that zone. This does not change the derivations in this chapter, however; but it would tend to increase the size of spillover effects from transport improvements, since spillovers would increase not only due to workplace relocation but also due to transport cost reductions. See Eliasson and Fosgerau (2019) for such a model.

[6] For empirical studies of the relationship between accessibility and productivity, see e.g. Melo and Graham (2014) and Börjesson et al. (2019).

workplace zones given existing wage rates (matching effects); and this, in turn, will affect wage rates even for those workers who do not change workplace zones (spillover effects). To see this formally, differentiate the total wages Z with respect to an arbitrary travel time, say t_{12}:

$$\frac{\partial Z}{\partial t_{12}} = \Sigma_{nij} \frac{\partial P_{ij}^n}{\partial t_{12}} W_{ij}^n w_j^n + \Sigma_{nij} P_{ij}^n \frac{\partial W_{ij}^n}{\partial t_{12}} w_j^n + \Sigma_{nij} P_{ij}^n W_{ij}^n \frac{\partial w_j^n}{\partial t_{12}} \tag{9}$$

The first term is the matching effect: the change in commuting patterns under constant wages and working hours. The second term is the labor supply effect: the change in the number of working hours, under constant commuting patterns and wage rates. The third term is the spillover effect: the change in wage rates due to changes in the number of workers per zone, under constant commuting patterns and working hours.

However, a transport improvement will not necessarily increase overall production, for example if a transport improvement will cause workers to choose workplaces in a way which reduces spillover effects and hence average wage rates.

The social benefits of a transport improvement

A transport improvement will generate social benefits in four ways. First, there is a direct effect for commuters: the consumer surplus generated by the reduction of generalized travel costs. Second, the change in commuting patterns will affect wage rates through the spillover mechanism, which will generate (positive or negative) benefits also for those workers whose travel costs did not change. Third, part of the benefits will be capitalized in land markets, when changes in the demand for land induce changes in land prices. This is a transfer of benefits from workers to land owners (which are assumed to be absent), and is therefore not an addition to total benefits (provided that monetary transfers are treated as neutral from a social welfare perspective). Fourth, part of the change in total wages is taken from workers as tax revenues. We now show this formally.

Aggregate social benefits in monetary terms are denoted B, and are defined as the sum of workers' expected utility in monetary terms[7] (letting $\lambda_{ij}^n = \frac{\partial v_{ij}^n}{\partial Y}$ denote the marginal utility of income and $\lambda^n = \Sigma_{ij} P_{ij}^n \lambda_{ij}^n$ be the expected marginal utility of income), land owner revenues and tax revenues:

$$B = \sum_n \frac{V^n}{\lambda^n} + \Sigma_{nij} p_i P_{ij}^n L_{ij}^n + \tau Z. \tag{10}$$

To derive the change in total benefits dB caused by a change in travel times or travel costs, we need some general results. First, we get from Roy's identity relationships between changes in land prices and demand for land, and between wage rates and working hours:

$$\begin{aligned} \frac{\partial v_{ij}^n}{\partial p_i} &= -\lambda_{ij}^n L_{ij}^n \\ \frac{\partial v_{ij}^n}{\partial w_j^n} &= \lambda_{ij}^n (1-\tau) W_{ij}^n \end{aligned} \tag{11}$$

[7] We define the welfare change to be the amount of money that keeps expected utility constant before and after a change. See Appendix for a derivation.

Second, a general property of additive random utility models is that the derivative of the expected maximal utility with respect to the utility of an alternative equals the corresponding choice probability (Fosgerau et al., 2013). This means that

$$\frac{\partial V^n}{\partial v_{ij}^n} = P_{ij}^n \tag{12}$$

Consider a small change of an arbitrary travel time, say dt_{12} (the corresponding case for a travel cost change is a trivial modification). This will induce second-order changes in land prices dp_i and wage rates dw_j^n, and a change in economic output (total wages) dZ. The monetized social benefits dB of all these changes can then be obtained by total differentiation of B:

$$\begin{aligned} dB = &\sum_{nij} \frac{1}{\lambda^n} \frac{\partial V^n}{\partial v_{ij}^n} \left(\frac{\partial v_{ij}^n}{\partial t_{12}} dt_{12} + \frac{\partial v_{ij}^n}{\partial p_i} dp_i + \frac{\partial v_{ij}^n}{\partial w_j^n} dw_j^n \right) + \\ &+ \sum_{nij} \left(P_{ij}^n L_{ij}^n + p_i \frac{\partial}{\partial p_i} \left(P_{ij}^n L_{ij}^n \right) \right) dp_i + \tau dZ \end{aligned} \tag{13}$$

Since total land supply is fixed for each zone, $\sum_{nj} P_{ij}^n L_{ij}^n$ is constant and hence $\frac{\partial}{\partial p_i}\left(\sum_{nj} P_{ij}^n L_{ij}^n\right) = 0$ for all i. Note that $\frac{\partial v_{ij}^n}{\partial t_{12}} = 0$ for all $(i,j) \neq (1,2)$ and $\frac{\partial v_{ij}^n}{\partial t_{12}} = \frac{\partial v_{12}^n}{\partial T}$ for $(i,j) = (1,2)$. Define the monetary value of travel time savings as $\theta_{ij}^n = -\frac{\partial v_{ij}^n}{\partial T} / \lambda^n = w_{ij}^n + \frac{\partial u_{ij}^n / \partial t}{\lambda^n}$ (see e.g. Jara-Díaz and Guevara (2003) for a derivation of this). Using this and (11)-(12), we get

$$\begin{aligned} dB = &\sum_n P_{12}^n \theta_{12}^n dt_{12} + \sum_{nij} \frac{\lambda_{ij}^n}{\lambda^n} (1-\tau) P_{ij}^n W_{ij}^n dw_j^n + \sum_{nij} \left(1 - \frac{\lambda_{ij}^n}{\lambda^n}\right) P_{ij}^n L_{ij}^n dp_i \\ &+ \tau dZ \end{aligned} \tag{14}$$

It is conventional in applied CBA to ignore differences in marginal utilities of income, and treat monetary transfers as neutral from a welfare perspective. There is a large literature and long-standing discussion on this practice (Börjesson & Eliasson, 2019; Galvez & Jara-Díaz, 1998; Harberger, 1978; Pearce & Nash, 1981; Sugden, 1999). For the purposes of this paper, it is both a convenient assumption and in line with how CBA is applied, so we will adopt it and assume that marginal utilities of income do not vary with the choice of workplace/residence, so $\frac{\lambda_{ij}^n}{\lambda^n} = 1$ for all (n,i,j). An equivalent assumption would be to assume from the start that the utility function $u^n(x,l,L,t)$ is money-metric (i.e. all $\lambda_{ij}^n = 1$ by definition); but in our view, this is less transparent than explicitly ignoring differences in the marginal utilities of income in this step.[8] Thus, ignoring differences in marginal utilities of income leads to the following welfare measure:

[8] There is a long debate about whether and in that case how differences in marginal utilities of income (MUI) should be taken into account in applied CBA, and for transportation projects in particular. Authors such as Harberger (1978) and Sugden (1999) have argued that the only practical way is to ignore them, and treat redistributive effects as welfare-neutral, while authors such as Galvez and Jara-Díaz (1998) and Pearce and Nash (1981) have argued that differences in MUI should be taken into account, for example by removing income effects from the value of time. See Börjesson and Eliasson (2019) for further discussion. Here, we follow the standard practice of ignoring differences in MUI:s. To be precise, this assumption is used in two places in the derivation. The first is when total benefits

$$dB = \sum_{n} P_{12}^{n}\theta_{12}^{n}dt_{12} + \sum_{nij}(1-\tau)P_{ij}^{n}W_{ij}^{n}dw_{j}^{n} - \sum_{nij} P_{ij}^{n}L_{ij}^{n}dp_{i} + \sum_{nij} P_{ij}^{n}L_{ij}^{n}dp_{i} + \tau dZ \qquad (15)$$

The benefits consist of five terms. The first three accrue to workers: travel time savings, i.e. the conventional consumer surplus (the first term), changes in wage rates (the second term) and changes in land prices (the third term). The fourth term is the change in land owner revenues, and the fifth term is the change in tax revenues. Note that the third and fourth terms cancel out; they represent the transfer from workers to land owners. Hence, if one is interested only in aggregate benefits, and not how they are divided between workers and land owners, changes in land prices can be ignored.

In the simulations, we calculate the consumer surplus with the common rule-of-a-half approximation. This gives the following benefits of a vector of travel time changes $\{\Delta t_{ij}\}$ and induced changes in working hours, wage rates and land prices. Indices 0 and 1 indicate before and after the change:

Conventional consumer surplus	$\Sigma_{nij}\dfrac{P_{ij}^{n0}+P_{ij}^{n1}}{2}\theta_{ij}^{n}\Delta t_{ij}$	(16)
Wider economic benefit from spillovers, net of tax (accrues to workers)	$(1-\tau)\Sigma_{nij}\dfrac{P_{ij}^{n0}W_{ij}^{n0}+P_{ij}^{n1}W_{ij}^{n1}}{2}\left(w_{ij}^{n1}-w_{ij}^{n0}\right)$	
Increased land owner revenues (transfer from residents, cancels out in total benefits)	$\Sigma_{nij}P_{ij}^{n1}L_{ij}^{n1}p_{i}^{1}-\Sigma_{nij}P_{ij}^{n0}L_{ij}^{n0}p_{i}^{0}$	
Increased tax revenues	$\tau\left(\Sigma_{nij}P_{ij}^{n1}W_{ij}^{n1}w_{j}^{n1}-\Sigma_{nij}P_{ij}^{n0}W_{ij}^{n0}w_{j}^{n0}\right)$	

"Wider economic benefits": benefits over and above the consumer surplus

The first term in (16) as well as (17) is the conventional consumer surplus for commuters, which in conventional transport CBA is assumed to capture all benefits of a transport improvement. As (16) shows, this is true only as long as wage rates are constant in each zone (all $dw_{ij}^{n}=0$), and when tax revenues do not change (either because $\tau=0$ or because $dZ=0$). When wage rates change due to spillover effects, or when tax revenues change, these additional terms (called *wider economic benefits*) should be added to the benefits in the CBA. Note that the changes in wage rates captured in this term is not the changes in workers' average wage rates resulting from changes in workplace choices (destinations): that would correspond to a matching effect, and is already included in the conventional consumer surplus. The second term instead captures the effects of changes in *wage rate offers at each workplace*. The reason that this enters the total benefits is that the wage rate offer in each workplace zone depends on

---

are defined as the sum of monetary benefits in (11); if we did not ignore differences in the MUI:s, it would matter to whom eventual benefits accrued, and we would for example need to assume something about the MUI of the eventual beneficiaries of tax revenues. The second is in the step from (15) to (16), where differences in the MUI associated with different between $(i.j)$-combinations for a given group n are ignored. The first assumption hence ignores inter-individual MUI differences, the second step ignores intra-individual MUI differences. As noted in the text, an alternative way to introduce both of these assumptions is to assume from the start that the utility function (5) is money-metric, but this would hide precisely where the assumption is needed, so we prefer to make this assumption explicitly later in the derivation instead.

the number of workers in the zone (the spillover agglomeration effect). This introduces an externality: a worker changing work zone changes not only his own wage rate (by changing zone) but also the wage rates of everyone else working in the old and new zones. These changes are the wage rate changes dw_j^n.

Further, note that the second term, the "wider economic impact term", is not equal to the total change in economic output (9). Comparing (9) to the final row of (14), we see that the net-of-tax parts of the first two terms of the change in economic output (9) – the change in workplace choice (matching effects) and the change in working hours (labor supply effects), net of tax – are already included in the consumer surplus (the first term in (14)). This means that the so-called "wider economic benefits" – the benefits not captured by the conventional consumer surplus – only consist of the change in tax revenues and the net-of-tax part of the change in output that results from spillover effects on wage rates. These results are further discussed in Eliasson & Fosgerau (2019). The derivation in the current paper shows that the same results hold also with endogenous lot sizes, residential location and working hours (which are not included in the model in Eliasson & Fosgerau (2019)). A similar result was obtained by Venables (2007), although without the distinction between labour supply, matching and spillover effects.

2.3 Four metropolitan regions (model setups)

In the simulations, we will use four different metropolitan regions (or sometimes "cities" for short), i.e. model setups with different geographies, networks and distributions of parameters. The geographies of the metropolitan regions are shown in Figure 1.

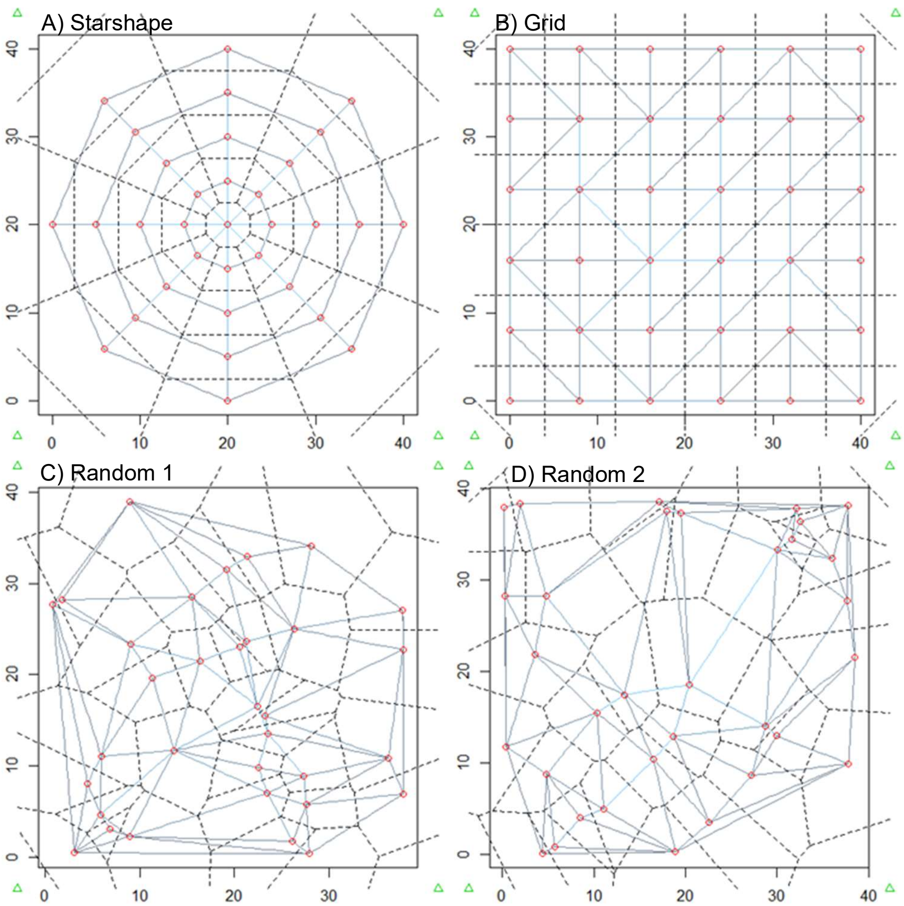

Figure 1. Four networks with roughly the same number of links, zones and total area. Dashed lines represent borders of zones. A) Star-shaped. B) Grid. C) Random 1. D) Random 2.

The regions have roughly the same number of links, zones and total area, to be comparable. The upper left network is a symmetric and star-shaped network with radial and concentric links. The upper right one has its nodes evenly distributed on a grid and the last two networks are randomly generated. Some attributes of the networks are listed in Table 1.

	#Links	#Zones	Total area (km^2)	Largest zone (km^2)	Smallest zone (km^2)
A) Star-shaped	64	33	2110	162	21
B) Grid	85	36	2272	64	56
C) Random 1	89	33	1917	136	11
D) Random 2	87	33	2023	121	10

Table 1. Some properties of the four metropolitan regions.

2.4 Model calibration

The parameters of the model are calibrated to give elasticities and key values that correspond to empirical values from Stockholm, e.g. average travel distance and travel time, wage distribution, average number of working hours, distribution of residential density, travel time and productivity elasticities etc. The utility function is a Cobb-Douglas function.

Table 2. Model parameters.

Number of classes	1000	Taxation rate		$\tau = 0.3$
Idiosyncratic tastes	$\varepsilon_{ij}^{n} \sim \boldsymbol{Gu}(0, 0.01)$	Spillover		$\eta = 1$
Exogenous income	$Y = 0$	Available time per day		$T = 16$
Home preferences	$H_i^n \sim \boldsymbol{N}(0, 0.25)$	Wage rates	Log(mean)	12.84
Destination preferences	$D_i^n \sim \boldsymbol{N}(0, 0.25)$		Log(sd)	4.60
Cobb-Douglas parameter for consumption	0.40		Lower bound	150000 SEK/year
Cobb-Douglas parameter for lot size	0.12		Upper bound	2000000 SEK/year
Cobb-Douglas parameter for leisure	0.48		Work days per year	228
Dispersion parameter in the logit function	0.01		Hours per day	8
Base travel speed	40 km/h	Travel cost		1.5 SEK/km
Travel discomfort parameter	0.05			

The diagrams below illustrate a number of outcomes from the model.

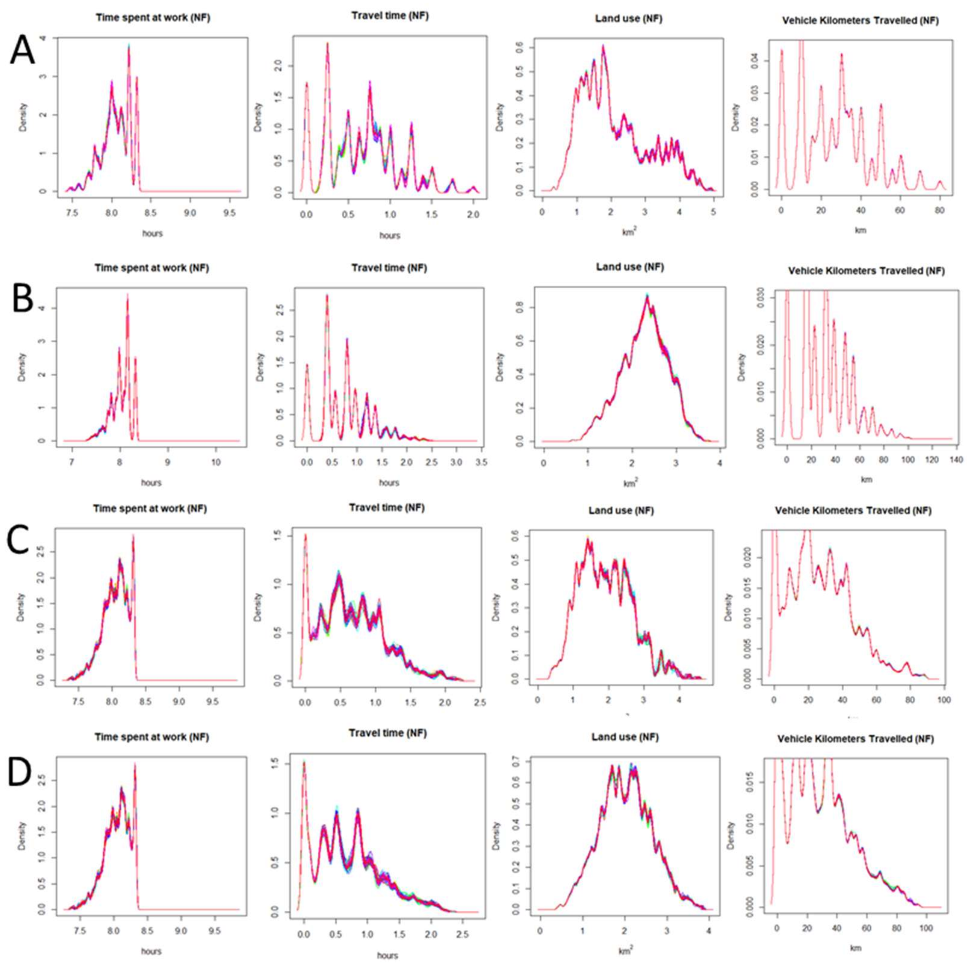

Figure 2. Distributions of working hours, travel time, residential densities and travel distances.

2.5 Keeping residential land use fixed

If travel times or travel costs change, the simulation model will calculate new choices of residence and workplace zones for all workers, and this will in result in new commuting patterns. The purpose of this paper is to compare this outcome, which we treat as the "true" outcome, to a "biased" forecasted outcome (and implied social benefits) when the residential location is kept fixed, since this is how transport forecasts and CBA:s are usually produced. Hence, we need to specify how this can be emulated using our simulation model.

Fixed residential land use in a transport model means not only that the number of residents in each zone is fixed, but also their socioeconomic characteristics are fixed. Workplace choices, on the other hand, are often allowed to change through the destination choice in the travel demand model, and this means that the implied number of workplaces per zone may change when the model is run[9]. To emulate this, we want to force the residential choice probability vector to stay constant for each class n, while allowing workplace choice probabilities to change.

Call the initial residence/workplace choice probabilities $\{P_{ij}^{n0}\}$ and the true choice probabilities after a scenario change $\{P_{ij}^{n1}\}$. A scenario change will also change land prices and wage rates from $\{p_i^0\}, \{w_j^{n0}\}$ to $\{p_i^1\}, \{w_j^{n1}\}$ to keep the equilibrium constraints (7) fulfilled. What we want to do is to define choice probabilities, land prices and wage rates that correspond to a case when residential land use is kept fixed. Call these $\{P_{ij}^{n2}\}$, $\{p_i^2\}$ and $\{w_j^{n2}\}$; this is what we will call the *biased* estimates of choice probabilities, land prices and so on. These are calculated by setting

$$P_{ij}^{n2} = P_{ij}^{n1} * \frac{\Sigma_j P_{ij}^{n0}}{\Sigma_j P_{ij}^{n1}} \quad \forall i, j, n \tag{17}$$

and calculating corresponding land prices and wage rates $\{p_i^2\}, \{w_j^{n2}\}$ such that the equilibrium constraints (7) are still fulfilled. This also gives a new set of working hours $\{W_{ij}^{n2}\}$ and land demands $\{L_{ij}^{n2}\}$. The rescaling of the choice probabilities ensures that the (expected) number of workers of each class n in each residential zone i is the same as in the initial situation. In other words, this ensures not only that the total number or residents per zone is fixed, but also that their characteristics – wage rate offers, taste parameters etc. – are fixed. All variables in the table of benefits with index "1" are then replaced with the corresponding index "2" variables. Note that this will not only affect travel demand: since we require equilibrium constraints to be fulfilled, it will also affect land owner revenues, tax revenues and wage rate changes.

Consider a set of travel time reductions Δt_{ij} which result in second-order changes in wage rates and working hours. With the notation above and using the rule-of-a-half approximation, we can write the difference between true benefits B and biased benefits B' as

$$B - B' = \Sigma_{nij} \frac{P_{ij}^{n1} - P_{ij}^{n2}}{2} \theta_{ij}^{n} \Delta t_{ij} + (1-\tau) \Sigma_{nij} \frac{P_{ij}^{n1} W_{ij}^{n1} - P_{ij}^{n2} W_{ij}^{n2}}{2} \left(w_{ij}^{n1} - w_{ij}^{n2}\right) + \tau\left(\Sigma_{nij} P_{ij}^{n1} W_{ij}^{n1} w_j^{n1} - \Sigma_{nij} P_{ij}^{n2} W_{ij}^{n2} w_j^{n2}\right) \tag{18}$$

This shows that the difference in benefits stems from three sources. The first term captures that changes in residential location will tend to attract more workers to those residential zones which enjoy the travel time savings. The second term captures wage

[9] Transport models are sometimes constrained such that the number of trip-ends per zone, and hence the number of implied workplaces, are kept constant between scenarios.

rate changes caused by the second-order effects on workplace choices, which in turn are caused by the changes in residential location. Note that if there had been no spillover mechanism, wage rates would have been constant in each workplace zone and this term would have vanished. The third term captures changes in tax revenues, partly due to differences in wage rates (as in the second term) and partly due to differences in working hours (since the choice of working hours depend on the choice of residential location).

The equation also shows that had there not been positive externalities in the model setup – i.e. no income tax and no positive spillover effects on wage rates – then the biased benefits would have always underestimated true benefits, since the first term is always positive. This is because constraining residential location to be fixed decreases the choice set in the "after" situation, which obviously decreases attainable total utility.

3 RESULTS

This section reports simulation results for the four cities described in section 2.3: a star-shaped city (A), a grid-shaped city (B) and two randomly generated geographies (C, D). As explained above, the benefits of a 10% travel time reduction are calculated for each link in the network in two scenarios: letting workers' residential location vary (called the *true* benefits), and keeping workers' residential location fixed (called the *biased* benefits). Figure 3 shows scattergrams of true vs. biased benefits for all investments in the four cities.

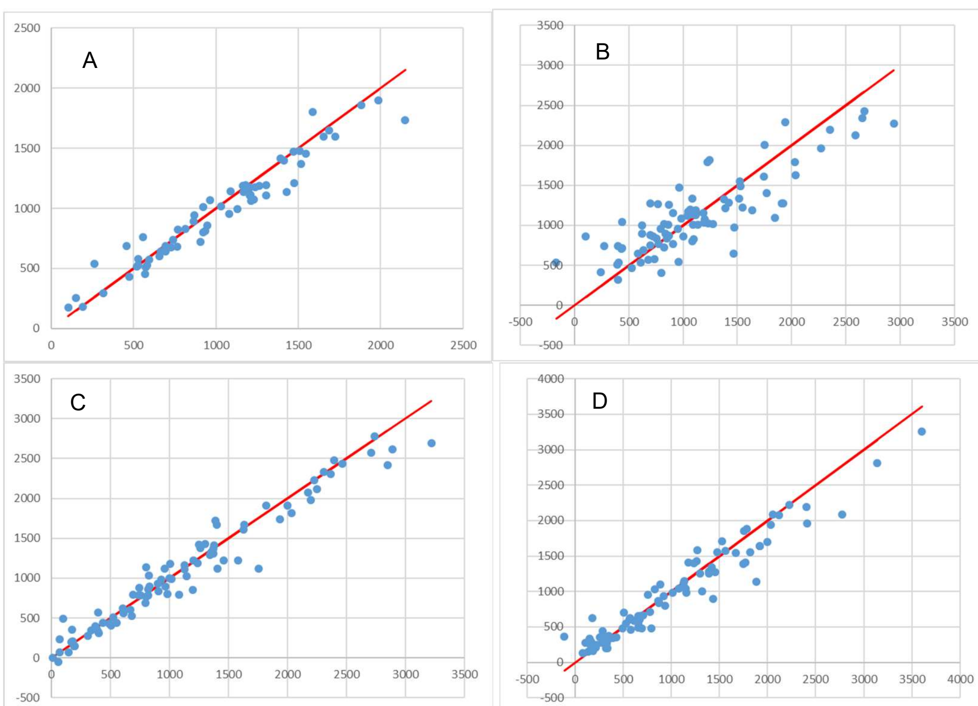

Figure 3. Scattergrams of true (x-axis) vs. biased (y-axis) estimates of benefits. From the top left: star (A), grid (B), random 1 (C), random 2 (D). The red lines are 45 degree lines.

The scattergrams show that the biased benefits may over- or underestimate the true benefits; as pointed out above, the biased benefits would always underestimate true benefits if there had not been positive externalities in the model.

The relative differences between true and biased benefits are on average 1% (star), 17% (grid), 5% (random 1) and -1% (random 2). The absolute relative differences between true and biased benefits are on average 12% (star), 36% (grid), 26% (random 1) and 27% (random 2). The differences in benefits are hence clearly appreciable, and the true benefits tend to be slightly higher than the biased estimates on average. The differences in the grid city are considerably higher than in the other three cities.

Next, consider a planner who wants to maximize welfare by selecting transport investments under a budget constraint (as described in section 2.1), but basing her selection on the biased estimates of the benefits rather than the true ones. Call this selection the biased selection, as opposed to the optimal selection based on the true benefits. The question is how much the two selections differ. Figure 4 gives a first indication for the star-shaped city. It shows the benefit-cost ratios (BCRs) for all investments, ranked from the best to the worst according to the true BCR. Apparently, the rankings of investments do not change much, since the biased benefits (the orange line) has a clearly decreasing overall trend, which means that investments are ranked in an order similar to the true one.

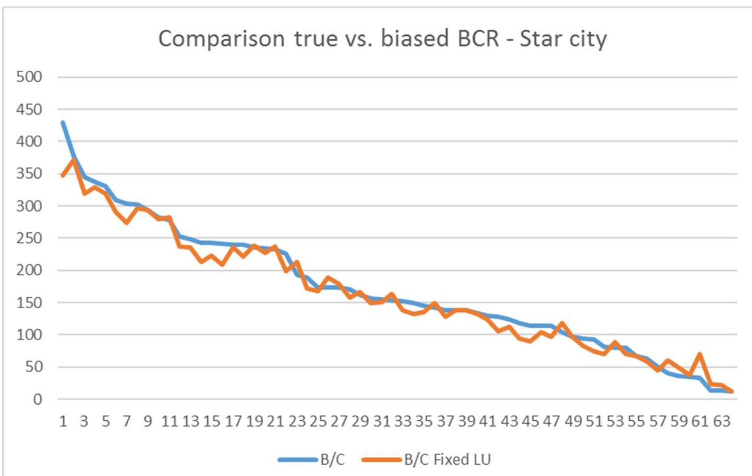

Figure 4. Comparison of benefit-cost ratios (BCR) between true and biased benefits, ranked according to true BCRs

This is a central result: it shows that the variation in benefits between projects is much larger than the errors in the benefit calculations, meaning that the overall project ranking (in terms of benefit-cost ratios) is fairly stable with respect to variations in the benefit calculations. On one hand, it is clear from the jagged line in Figure 4 (showing the biased BCR:s) that there is a lot of *smaller* changes in the project ranking, i.e. projects with similar BCR:s that change place with each other. On the other hand, it is also clear that there are only a few *large* changes in the ranking, where a project is ranked much higher or lower in the biased ranking compared to the true ranking.

Table 3 summarizes the results for all four cities. The optimal selection of investments has been calculated under four different budget constraints (100, 200, 300 and 500). The number of investments in each optimal selection is shown in the first column. The second column shows how many of these are also found in the biased selection. The last column shows S (see section 2.1), the share of projects in the optimal selection omitted in the biased selection. $S = 0\%$ means that the biased selection and the optimal selection are identical.

City	*Investment budget*	*Number of projects in optimal selection*	*Number of those projects included in biased selection*	*Share of optimal selection excluded from biased selection*
A) Star	100	20	18	10%
	200	40	40	0%
	300	49	49	0%
	500	64	64	0%
B) Grid	100	12	9	25%
	200	24	19	21%
	300	35	29	17%
	500	57	50	12%
C) Random 1	100	16	14	12%
	200	28	25	11%
	300	43	42	2%
	500	64	62	3%
D) Random 2	100	13	11	5%
	200	27	23	15%
	300	38	36	5%
	500	57	54	5%

Table 3. Biased selection compared to optimal selection for the four cities under four different budgets.

In several cases, the biased and optimal selection do not differ ($S = 0\%$), and in most cases $S < 15\%$. Not surprisingly, differences are largest in the grid city, where true and biased benefits differ the most, and we find the maximum $S = 25\%$.

It is illuminating to explore an example in some detail. Consider the grid-shaped city with an investment budget of 200, where $S = 21\%$. The optimal selection consists of 24 investments (which is 28% of all possible investments). Nineteen of these are included in the biased selection. The five omitted projects are ranked 17, 18, 20, 21, 23 in the optimal selection, but ranked 37, 26, 59, 33, 32 in the biased selection. The change in ranking is hence quite substantial in these cases. Looking at all investments in the grid city, the average absolute change in ranking is just over 11. The difference in ranking exhibits a substantial variation; many investments more or less keep their place in the ranking, while some change a lot. Overall, 39% of investments change their ranking less than 6 steps, while 22% of the investments change their ranking more than 15 steps.

In the other cities, however, differences are smaller. Considering a budget of 200, the average absolute change in ranking is 2 for the star city, 4 for random city 1 and 6 for random city 2. Consequently, the biased selections differ much less from the optimal in these cities, which is also evident from Table 3.

Finally, we turn to the most important comparison, namely how much total achieved benefits decreases due to the planner basing her selection on the biased benefit estimates. Arguably, it is of minor importance if some investments are replaced with others, as long as total achieved benefits do not decrease too much. Remember from section 2.1 that ΔB is the relative loss of aggregate net benefits due the planner's imperfect information, i.e. the total true benefits of the biased selection relative to the total true benefits of the optimal selection.

City	*Investment budget*	*Benefit of optimal selection*	*Benefit of biased selection*	*Relative loss of benefits*
A) Star	100	28 823	28 545	-1.0%
	200	45 129	45 129	0.0%
	300	56 494	56 341	-0.3%
	500	63 340	63 115	-0.4%
B) Grid	100	27 677	26 174	-5.4%
	200	45 061	42 769	-5.1%
	300	58 961	56 257	-4.6%
	500	70 213	65 802	-6.3%
C) Random 1	100	29 597	29 324	-0.9%
	200	49 705	48 963	-1.5%
	300	66 051	65 730	-0.5%
	500	78 321	77 574	-1.0%
D) Random 2	100	24 856	24 367	-2.0%
	200	42 489	41 579	-2.1%
	300	56 410	55 616	-1.4%
	500	67 033	65 878	-1.7%

Table 4. Benefits of optimal selection compared to benefits of biased selection.

The relative losses of achieved benefits are virtually negligible, at least compared to the magnitude of uncertainties always present in modeling and future scenario assumptions. This is a surprising result, given the often comparatively large differences between true benefits and the biased estimates, and the appreciable shifts in rankings of investments. However, the fairly large differences in these respects do not translate to large losses of benefits on an aggregate level. Intuitively, this is because if benefits vary substantially between investments, the ranking will be stable despite substantial biases in the benefit estimates. On the other hand, if several investments have similar benefits, their ranking is easier to affect; but on the other hand, that of course means that replacing one investment with another does not affect total benefits much.

4 CONCLUSIONS

Infrastructure investment plans are usually constructed, at least in principle, by choosing the best projects from a large pool of candidates, subject to an overall budget constraint. The objective – again, at least in principle – is to attain maximal total benefits given this budget constraint. To do this, policy makers use the output of transport models to assess the effects and benefits of the competing investments. However, transport models used in practice almost always neglect changes in residential location, and this is a source of error in the calculated benefits. Neglecting such changes means that the calculated benefits may either over- or underestimate the true benefits (as long as there are externalities in the setup). The central question of this paper is how large the error caused by this approximation is, and how much benefits are lost. The analysis where residential location is unchanged can also be interpreted as a medium-term equilibrium, if job locations and wage rates change quicker than residential location and lot sizes.

Of course, it is clearly preferable to take land use effects into account when evaluating investment benefits, if there is an opportunity to do so and the costs and efforts are not prohibitively high. Moreover, changes in residential location and land use are interesting and important in their own right, not just because they affect the benefits of an investment. But since integrated transport/land use models are still rare, and often costly to use in terms of data preparation and running times, an overwhelming majority of transport projects are evaluated assuming fixed residential land use. Hence, it is worthwhile to explore whether the errors this simplification induces are so large that it severely hampers the value of such evaluations.

It should be pointed out the focus of the present paper is on selecting between relatively modest improvements of an existing network. This situation probably differs from evaluating very large projects yielding extremely large changes in accessibility, and large relocation effects on a national scale, such as establishing the US interstate highway network after World War II, or the UK railway network in the 19th century. In such extreme cases, land use changes may be particularly important.

Broadly speaking, our results indicate that neglecting to account for changes in residential location (and hence changes in demand for land) can sometimes cause appreciable errors in the estimated benefits. True benefits tend to be slightly higher than the biased estimates on average, but this tendency is weak: benefits may be overestimated as well as underestimated. However, the errors only have a marginal impact on the attained total benefits of an infrastructure plan – at worst a loss of around 5% of the total attained benefits, and in most cases the losses are negligible. The difference in project selection can be larger: in our simulations, up to 25% of the projects that should have been included in the project selection were left outside. In most cases, however, the differences between the optimal selection (based on true benefits) and the biased selection are small. The intuition explaining these perhaps counterintuitive results is that the variation in project benefits is much larger than the errors in the benefit calculations. This results in the project ranking, in terms of benefit-cost ratios, is fairly stable. Even if errors in the benefit-cost calculations cause quite a lot of smaller changes in the project ranking, the number of large changes is relatively small, and hence only a few projects are replaced in the optimal selection set. Our model does not include endogenous congestion, so the stability of the project ranking is not due to the phenomenon where benefits of link improvements are partly cancelled out by latent demand or routing effects ("moving around congestion").

We conclude that the errors caused by ignoring residential relocation do not reduce the value of conventional cost-benefit analysis as an investment selection tool to any considerable extent. This is analogous to other results regarding the robustness of CBA with respect to uncertainties and errors in parameters, demand forecasts, cost estimates and scenario assumptions (Asplund & Eliasson, 2016; Börjesson, Eliasson, et al., 2014; Eliasson & Fosgerau, 2013). These studies, based on data on real infrastructure investments (as opposed to the simulated data in the present study), also concluded that the variation in benefits between projects was much larger than the errors and uncertainties in benefit calculations, which meant that the overall project ranking was relatively stable.

5 ACKNOWLEDGMENTS

We are grateful for helpful comments by Ismir Mulalic.

6 REFERENCES

Acheampong, R. A., & Silva, E. A. (2015). Land use–transport interaction modeling: A review of the literature and future research directions. *Journal of Transport and Land Use*, *8*(3), 11–38.

Asplund, D., & Eliasson, J. (2016). Does uncertainty make cost-benefit analyses pointless? *Transportation Research Part A: Policy and Practice*, *92*, 195–205. https://doi.org/10.1016/j.tra.2016.08.002

Bondemark, A., Sundbergh, P., Tornberg, P., & Brundell-Freij, K. (2020). Do impact assessments influence transport plans? The case of Sweden. *Transportation Research Part A: Policy and Practice*, *134*, 52–64. https://doi.org/10.1016/j.tra.2020.02.002

Börjesson, M., & Eliasson, J. (2019). Should values of time be differentiated? *Transport Reviews*, *39*(3), 357–375. https://doi.org/10.1080/01441647.2018.1480543

Börjesson, M., Eliasson, J., & Lundberg, M. (2014). Is CBA ranking of transport investments robust? *Journal of Transport Economics and Policy*, *48*(2), 189–204.

Börjesson, M., Isacsson, G., Andersson, M., & Anderstig, C. (2019). Agglomeration, productivity and the role of transport system improvements. *Economics of Transportation*, *18*, 27–39. https://doi.org/10.1016/j.ecotra.2018.12.002

Börjesson, M., Jonsson, R. D., Berglund, S., & Almström, P. (2014). Land-use impacts in transport appraisal. *Research in Transportation Economics*, *47*, 82–91. https://doi.org/10.1016/j.retrec.2014.09.021

Dagsvik, J. K., & Karlström, A. (2005). Compensating Variation and Hicksian Choice Probabilities in Random Utility Models that are Nonlinear in Income. *The Review of Economic Studies*, *72*(1), 57–76. https://doi.org/10.1111/0034-6527.00324

Eliasson, J. (2010). The influence of accessibility on residential location. Theory and estimation of an integrated land use-transportation model. In J. Preston, F.

Pagliara, & D. Simmonds (Eds.), *Residential Location Choice – Models and Applications*. Springer Verlag.

Eliasson, J., Börjesson, M., Odeck, J., & Welde, M. (2015). Does Benefit–Cost Efficiency Influence Transport Investment Decisions? *Journal of Transport Economics and Policy*, *49*(3), 377–396.

Eliasson, J., & Fosgerau, M. (2013). Cost overruns and benefit shortfalls—Deception or selection? *Transportation Research B*, *57*, 105–113. https://doi.org/10.1016/j.trb.2013.09.005

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. https://doi.org/10.1016/j.ecotra.2019.02.001

Eliasson, J., & Lundberg, M. (2012). Do Cost–Benefit Analyses Influence Transport Investment Decisions? Experiences from the Swedish Transport Investment Plan 2010–21. *Transport Reviews*, *32*(1), 29–48. https://doi.org/10.1080/01441647.2011.582541

Eliasson, J., & Martìnez, F. (2001). Interfaces between location, land-use and travel decisions. In *The leading edge in travel behaviour research*. Elsevier.

Eliasson, J., & Mattsson, L.-G. (2000). A model for integrated analysis of household location and travel choices. *Transportation Research Part A: Policy and Practice*, *34*(5), 375–394. https://doi.org/10.1016/S0965-8564(99)00038-5

Eliasson, J., & Mattsson, L.-G. (2001). Transport and Location Effects of Road Pricing: A Simulation Approach. *Journal of Transport Economics and Policy*, *35*(3), 417–456.

Fosgerau, M., McFadden, D., & Bierlaire, M. (2013). Choice probability generating functions. *Journal of Choice Modelling*, *8*, 1–18. https://doi.org/10.1016/j.jocm.2013.05.002

Franklin, J., Waddell, P., & Britting, J. (2002). *Sensitivity analysis approach for an integrated land development & travel demand modeling system*. Association of Collegiate Schools of Planning 44th Annual Conference.

Galvez, T. E., & Jara-Díaz, S. R. (1998). On the social valuation of travel time savings. *International Journal of Transport Economics*, *25*(1), 205–219.

Harberger, A. C. (1978). On the Use of Distributional Weights in Social Cost-Benefit Analysis. *Journal of Political Economy*, *86*(2), S87–S120. https://doi.org/10.2307/1829758

Iacono, M., Levinson, D., & El-Geneidy, A. (2008). Models of Transportation and Land Use Change: A Guide to the Territory. *Journal of Planning Literature*, *22*(4), 323–340. https://doi.org/10.1177/0885412207314010

Iniestra, J. G., & Gutiérrez, J. G. (2009). Multicriteria decisions on interdependent infrastructure transportation projects using an evolutionary-based framework. *Applied Soft Computing*, *9*(2), 512–526. https://doi.org/10.1016/j.asoc.2008.07.006

Jara-Díaz, S. R., & Guevara, C. A. (2003). Behind the Subjective Value of Travel Time Savings. *Journal of Transport Economics and Policy*, *37*, 29–46.

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. https://doi.org/10.1111/j.1435-5957.2012.00462.x

Moeckel, R., Garcia, C. L., Chou, A. T. M., & Okrah, M. B. (2018). Trends in integrated land use/transport modeling: An evaluation of the state of the art. *Journal of Transport and Land Use*, *11*(1). https://doi.org/10.5198/jtlu.2018.1205

Nellthorp, J., & Mackie, P. (2000). The UK Roads Review—A hedonic model of decision making. *Transport Policy*, *7*(2), 127–138. https://doi.org/10.1016/S0967-070X(00)00002-0

Pearce, D. W., & Nash, C. A. (1981). *The Social Appraisal of Projects – A Text in Cost-Benefit Analysis*. Macmillan.

Sugden, R. (1999). *Developing a consistent cost-benefit framework for multi-modal transport appraisal* [Report to the Department for Transport]. University of East Anglia.

Szimba, E., & Rothengatter, W. (2012). Spending Scarce Funds More Efficiently—Including the Pattern of Interdependence in Cost-Benefit Analysis. *Journal of Infrastructure Systems*, *18*(4), 242–251. https://doi.org/10.1061/(ASCE)IS.1943-555X.0000102

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.

Waddell, P. (2011). Integrated Land Use and Transportation Planning and Modelling: Addressing Challenges in Research and Practice. *Transport Reviews*, *31*(2), 209–229. https://doi.org/10.1080/01441647.2010.525671

Waddell, P., Ulfarsson, G. F., Franklin, J. P., & Lobb, J. (2007). Incorporating land use in metropolitan transportation planning. *Transportation Research Part A: Policy and Practice*, *41*(5), 382–410. https://doi.org/10.1016/j.tra.2006.09.008

7 APPENDIX: RANDOM UTILITY WELFARE MEASURES

A natural monetary measure of the welfare of a change is the equivalent variation (EV), which is the amount of money that keeps utility after and before the change constant. In a random utility setting, there are two alternative ways to define the EV: either as the amount of money that keeps expected utility constant before and after the change, or as the expected amount of money that is needed to keep actual utility constant, keeping the random components in the utility function fixed. The interpretation in the former case is that the random components of the utility function are redrawn after the change, and then the choice maker makes new choices conditional on the new random components. The interpretation in the latter case is that random components are kept constant after the change, so the choice maker makes new choices given the same random components as before the change. The former definition is often more natural to use for long-run evaluations, and the latter for short-term evaluations. The latter definition leads to the welfare measures developed in Dagsvik and Karlström (2005). Since our setting models the long run (all locations and lot sizes are allowed change without memory of the past), we use the former definition.

Let $v_k(Y) + \varepsilon_k$ be the indirect utility of alternative k conditional on income Y, where $v_k(Y)$ is a deterministic component and ε_k is a random component. Let $V = E\left(\max_k v_k(Y) + \varepsilon_k\right)$ be the expected utility for an individual facing the choice between the alternatives. The equivalent variation dEV of a vector of infinitesimal changes $\{dv_k\}$ in the deterministic part of the indirect utility function can then be defined as the amount of money that gives the same expected utility as this vector of utility changes, i.e. dEV is defined by this condition:

$$V\big(v_k(Y + dEV)\big) = V(v_k(Y) + dv_k) \qquad (19)$$

Taylor expanding both sides gives

$$\sum_k \frac{\partial V}{\partial v_k} \frac{\partial v_k}{\partial Y} dEV = \sum_k \frac{\partial V}{\partial v_k} dv_k \qquad (20)$$

Using $\frac{\partial V}{\partial v_k} = P_k$ where P_k is the choice probability, and denoting the marginal utility of income conditional on choosing alternative k $\frac{\partial v_k}{\partial Y} = \lambda_k$, we get

$$dEV = \frac{\sum_k P_k dv_k}{\sum_k P_k \lambda_k} \qquad (21)$$

The numerator is the difference in expected utilities before/after the change, so this implies that the monetary value of the expected utility V is simply $\frac{V}{\lambda}$, where $\lambda = \sum_k P_k \lambda_k$. If the random components are Gumbel distributed, the difference in expected utility is the difference in logsums before and after, which can be calculated exactly.