EVALUATING THE EFFECT OF ROAD PRICING STRATEGIES IN MATSIM, USING AGENT-SPECIFIC AND INCOME-DEPENDENT VALUES

W. du Plessis[1*] and J.W. Joubert[1]

[1]Department of Industrial and Systems Engineering

University of Pretoria, South Africa

Wilna.duplessis@up.ac.za

Johan.joubert@up.ac.za

ABSTRACT

Road pricing has become a very controversial subject in South Africa with the introduction of the Gauteng Freeway Improvement Project (GFIP). Transport authorities regard road pricing as an effective means to finance road infrastructure and reduce road congestion. On the other hand it has also been proven that policies which have a regressive effect on the welfare distribution of a society have very low levels of public acceptance, which has a direct effect on the project feasibility. Due to the extreme income inequality in South Africa it can be expected that the public acceptance will be very low, which will drastically impact the feasibility of GFIP. The state-of-practice models used for the feasibility studies of GFIP do not consider the effect of individual income and considers the same value of time for all road users. This paper will investigate means to consider road users' income in an attempt to determine the effect of using individual, income-dependent values of time in a micro-simulation approach towards policy evaluation. This will be done using the Multi-Agent Transport Simulation toolkit (MATSim) and applying the model to the Gauteng Freeway Improvement Project.

1 INTRODUCTION

1.1 Gauteng Freeway Improvement Project

The Gauteng Freeway Improvement Project (GFIP) has become a highly controversial subject in South Africa evoking a lot of public and political opposition even before it has been put into commission. Despite millions being invested in road infrastructure improvements, the demand for road space still exceeds road capacity and the supply of roads. This is especially the case in Gauteng, which, despite being the smallest province, account for 38% of the country's economic activity and the heaviest congested freeways in Africa. Since tax revenues cannot be the only source of funds for road infrastructure development, road user pricing through tolling is deemed a viable solution for funding, and this is the reasoning beyond the GFIP [1]. According to Motau [2], the aim of the GFIP is to improve the overall transportation of people, products, and services along Gauteng's major arterial freeways by improving 185 km of the busiest inter-city corridors in Gauteng through lane widening; median and road lighting; interchange upgrading; installation of directional ramps and traffic management systems; as well as the implementation of an electronic open toll system [3].

1.2 Document Structure

The remainder of this paper will focus on a comparison between the state-of-practice transport simulation models and the new agent-based approach of MATSim, followed by a more detailed description of MATSim. Subsequently the mathematical relationship between income and road users' willingness-to-pay will be investigated. The development of a synthetic population for Gauteng to be used in the MATSim simulation will be discussed, and ending with a conclusion.

2 STATE-OF-PRACTICE METHODS VS. AGENT-BASED SIMULATION

The success of the GFIP is directly dependent on the acceptance of the initiative by road users, and the feasibility studies performed need to consider the financial risk of the endeavour by considering the impact the road pricing will have on both the tolled and non-tolled routes.

The state-of-practice static assignment models used in the feasibility studies of the GFIP were however unable to predict the behavioural response of road users to the intended toll, and it was also unable to consider the differences in willingness-to-pay for the road users, and therefore assumed the same value for all road users in the respective vehicle classes. Although the static assignment models have proven to be an effective tool for urban transport development, there has been a recent shift towards dynamic agent-based modelling [4].

Gao [5] and Fourie [6] recently performed comparative studies to evaluate the performance of traditional state-of-practice models against the new popular agent-based micro simulation approach. For both studies EMME/2 and the Multi Agent-based Transport Simulation Toolkit (MATSim) were compared, two transport modelling approaches representing the current and future modelling directions.

2.1 EMME/2

EMME/2 is a traditional four step transport model that is extensively used as a transport modelling tool. Over the years it has proved to be a reliable long-term travel demand forecaster [5]. EMME/2 solves traffic assignment problems through a mathematical programming approach using the Franke-Wolfe algorithm to find a numerical solution which satisfies the objective function, which is the desired state of equilibrium where no additional improvements can be made [6]. In the traditional four step model, the trip generation step determines the number of outgoing and incoming trips for different zones. The second step, trip distribution, connects the origins and destinations from the different zones to form trips and produce an Origin-Destination matrix. In the modal split step, the different transport modes are considered, resulting in modal Origin-Destination matrices. During the final step, route assignment, routes are assigned to every trip that is simulated in the model.

2.2 MATSim

MATSim is an agent-based transport modelling framework which solves traffic assignment problems through microsimulation [5]. In the simulation each agent aims to execute his daily activities on a physical road network in an attempt to maximise his utility. This is done by making decisions regarding route choices, transportation modes and activity sequence, in an attempt to improve past experiences.

2.3 Comparison Of Methods

From the descriptions it is evident that EMME/2 does not consider individual behaviour of population agents. To compare the performance of the two models under investigation, Fourie [6] modelled Gauteng's peak morning hour's private vehicle demand with both the models. The results were then compared to actual data obtained from traffic counts and the results from the National Household Transport Survey (2003).

From the results it was found that both models produced similar results for the prediction of traffic volumes. With regard to peak hour volumes on links it was found that MATSim performed significantly better, by producing lower volumes on high capacity links than EMME/2, since MATSim agents avoid congestion by utilising capacity on secondary links. MATSim also produced much more accurate predictions with regard to travel times, however, it took significantly longer to solve than the EMME/2 model for the same problem. This is however a comparison between approximately five minutes and two hours for the respective models. Except for the trade-off between computational time and simulation accuracy, MATSim yielded equal or superior results compared to the EMME/2 model. MATSim also executes the simulation considering the actual road network and road capacities, whereas the four step model only considers zones and major routes between zones. Another advantage of MATSim is that it is a free, open source software package, and the open source environment stimulates continuous development and improvement of MATSim [5, 6].

Based on these results it is expected that MATSim will also yield superior results to the current predictive modelling approach used for the GFIP.

3 SIMULATION APPROACH

The approach followed by MATSim assumes that every road user represented in the synthetic population can be modelled as an agent. To execute the simulation a network is required which represents the physical environment in which the activities will be executed. This network typically consist of nodes, where activities will be executed, and links connecting the nodes, characterised by their length, capacity, number of lanes, and maximum allowable speed. A comprehensive, open-source *OpenStreetMap* network of Gauteng is currently used.

This agent-based simulation approach used by MATSim, as indicated in Figure 1, will be described in the following sections according to Rieser [7].

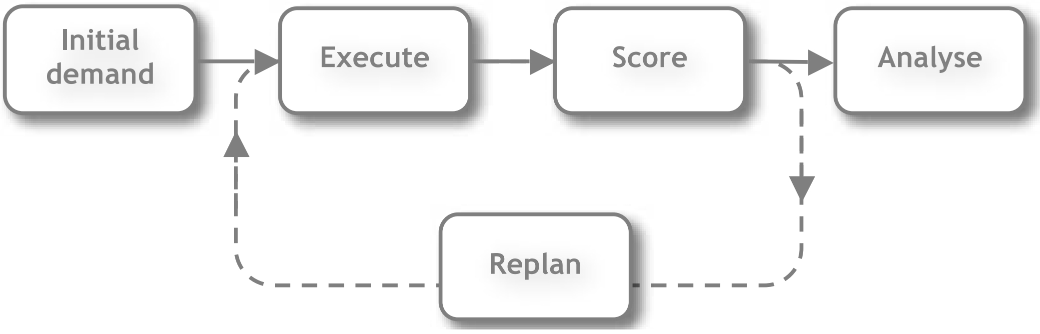

Figure 1: The MATSim Controller

3.1 Initial Demand

Prior to the simulation an initial plan is assigned to each agent which consists of the type, location, and distribution of activities to be performed. These plans are usually generated from available data sources such as travel surveys and census data.

3.2 Execution

In the mobility simulation all the agents' plans are executed simultaneously in the simulation of the physical environment. A queue simulation is favoured which models the network links as first-in-first-out queues which allows agents to only enter links with adequate capacity, or prevents agents from entering links with full capacity causing congestion [6].

3.3 Scoring

MATSim's agent-based approach allows agents to learn. The implementation of the simulation iterates between plans generation and the mobility simulation. The system remembers several of these plans for each agent, and scores the performance of each plan. The basic MATSim utility function calculates the total utility of an agent's plan U_{total} as the sum of the plan's individual contributions for performing, travelling, and being late, where *n* is the number of activities. The utility function is expressed as:

$$U_{total} = \sum_{i=1}^{n} U_{perform}^{i} + \sum_{i=1}^{n} U_{travel}^{i} + \sum_{i=1}^{n} U_{late}^{i} \tag{1}$$

3.4 Replanning

The learning ability of agents allows them to improve their future plans through replanning. This can either be done by agents choosing the plan with the highest utility (from the scoring phase) or obtaining new plans by copying and modifying existing plans.

Once a new plan is selected, the simulation returns to execution and the simulation is repeated until the desired relaxed state is achieved.

4 THE WILLINGNESS TO PAY IN ECONOMIC EVALUATION OF TRANSPORT PROJECTS

The economic viability of transport projects such as GFIP is directly dependent on public acceptance. Kickhöfer et al. [8] showed that tolling might have a disproportional regressive effect on the welfare of a population, affecting lower income groups worse than higher income groups. They have also proven that by considering income in transportation planning one could allow policy makers to anticipate implementation problems and public acceptance with higher certainty.

4.1 Income Distortion

It is already known that the socio-economic imbalances in South Africa play a large role in transport demand and travel patterns, especially with regard to origin-destination patterns, mode selection and route choice. This also raises the concern of the welfare effect of the GFIP and the accuracy of current GFIP estimations, considering that income inequality was not considered. According to the Gini-coefficient, the globally accepted measure of income distribution, South Africa currently ranks as the top country in the world with regard to income inequality with a Gini-coefficient of 0.7 [9]. The Gini-coefficient is defined mathematically by the Lorenz curve (Figure 2), that plots the proportion of households (x-axis) against the cumulative income of the population (y-axis). The Gini-coefficient is determined by calculating the ratio of the area between the Lorenz curve and the equality line (45 degrees) (A), and the total area beneath the equality line (A+B) [9]. This Lorenz curve and South Africa's corresponding high Gini-coefficient indicates the need to consider income in transport planning.

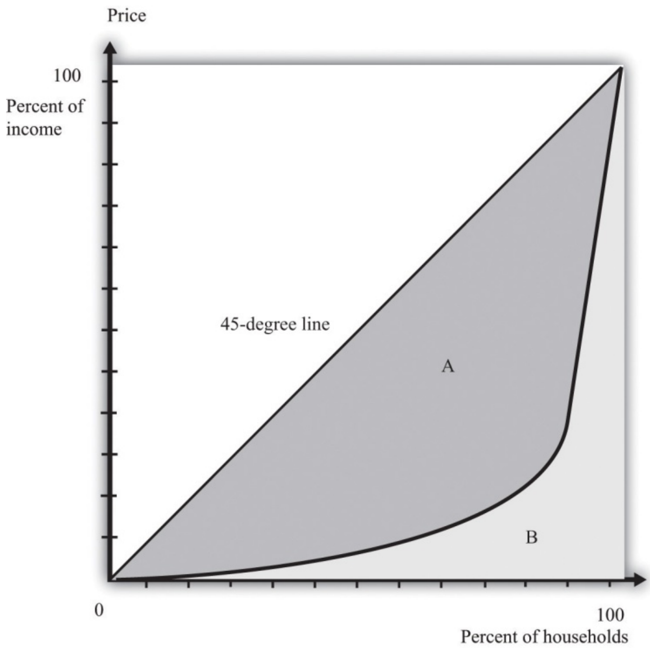

Figure 2: Lorenz Curve Depicting South Africa's Income Distribution

4.2 Value Of Travel Time Savings

The concept of willingness to pay or the valuation of travel time savings aims at maximising each individual's utility for performing activities, subject to:

- the individual's expenditure is limited by his income earned by using time to work,
- and, the total amount of time spent on work, leisure, and travel activities is limited by the number of hours in a day.

Agents are usually required to perform a trade-off between the extra consumption that work earns against the lost leisure opportunities it requires. Agents do however have the opportunity to extend their amount of work- or leisure time by spending money on tolling or alternative transport modes to save travel time, resulting in an agent's willingness to pay, or individual value of travel time savings [10].

4.3 MATSim and The Value Of Travel Time Savings

With the popularity of MATSim increasing, so does the complexity and diversity of its applications. In a recent study, conducted in Switzerland, it was again shown that MATSim can produce superior results to most existing transport modelling tools, but in this case for modelling road users' reaction to time-dependent tolls [8]. In a similar study, also performed in Switzerland, it was found that MATSim can be used to design better solutions to situations such as tolling, where policy implementations has low public acceptance [11]. The most significant difference between these applications and traditional transport modelling techniques, or even the standard MATSim applications, is the consideration of agent income.

4.4 Incorporating Income In MATSim

To incorporate income into MATSim, the standard MATSim utility function as shown in eq. (1) need to be adapted to accommodate agent-specific values. By decomposing eq. (1), the positive utility for performing an activity can be calculated by the logarithmic function

$$U_{perf,i}\left(t_{perf,i}\right) = \beta_{perf}.\, t_{*,i}.\, ln\left(\frac{t_{perf,i}}{t_{0,i}}\right) \quad \textbf{(2)}$$

where $t_{perf,i}$ is the time spent on activity $i, t_{*,i}$ is the typical duration of activity i, and β_{perf} is the marginal utility of an activity at its typical duration. $t_{0,i}$ is a scaling parameter related to both the minimum duration and importance of an activity, however, since dropping activities from plans are not allowed $t_{0,i}$ will have no effect.

The negative utility for being late is assumed by:

$$U_{late,i}\left(t_{late,i}\right) = \beta_{late}.\, t_{late,i} \quad \textbf{(3)}$$

where β_{late} is the marginal utility for being late (measured in h^{-1}) and $t_{late,i}$ is the number of hours late to activity i.

The above two components are not directly income dependent. As per the definition of willingness to pay, these two components will however be indirectly affected by income when an agent can use the money obtained from income to adjust his travel time to have more time for performing an activity or to prevent being late. This concept can be calculated by:

$$U_{mode,i,j} = \beta_{cost}.\ln\left(y_j\right) - \beta_{cost}\frac{c_{i,mode}}{y_j} + \beta_{tt,mode}.\, t_{i,mode} \quad \textbf{(4)}$$

where y_j is the daily agent income, β_{cost} is the marginal utility of money, $c_{i,mode}$ is the monetary cost for trip i, $\beta_{tt,mode}$ is the marginal disutility for travelling with mode i and $t_{i,mode}$ is the mode specific trip duration of trip i. From Eq. (4) it is evident that a higher agent income will result in a higher utility value [8, 11].

This section described the need to consider income in transportation studies, the concepts of income-inequality and income- and agent-dependent willingness-to pay, as well as MATSim's ability to incorporate income into the utility function. However, in order to incorporate income into MATSim it is critical that daily agent income is known, and more difficult, assigned to the specific agents in the population.

5 POPULATION GENERATION

One of the most important aspects, and also the initial step for any agent-based microsimulation is the definition of agents. Due to time and cost constraints it is usually impossible to collect comprehensive and detailed attribute data from census for a large population, and the data required for a simulation is not readily available. For the accuracy of the simulation it is however imperative that the synthetic population represents the actual population as close as possible.

To achieve a close-to-reality synthetic population a similar approach as followed by Müller and Axhausen [12,13] was followed. The approach to population synthesis is to fit available, up to date aggregate data with census data to achieve a population where the agents' attributes in the population are similar to those in the census, and the number of agents in each category matches the available aggregate data [12].

For South Africa, a 10% sample data set from the 2001 census is available providing very rich and useful data. This consists of a sample, representative of the total population, containing data for 948 592 households, and 3 725 655 corresponding individuals. This serves as the ideal source of aggregate data for the population synthesis, since households and individuals can be fitted simultaneously.

For the purpose of transport simulation the households, the corresponding individuals living in the respective households, as well as all the individuals' sex, age, race, living quarters or dwelling type, main mode of transportation and income were sourced from the available data. Serving as aggregate data, this data can now be used to estimate weights to generate a synthetic population either nationally, provincially, or for metropolitans, since group control data is available for each of these areas from census.

For the GFIP model, a Gauteng synthetic population was developed through an IPF algorithm written in Python. This yielded the necessary weights so that synthetic population derived from the aggregate data matched the household and individual totals for Gauteng with a high level of accuracy.

6 RELATIONSHIP BETWEEN COMPONENTS

The relevant data obtained from the population synthesis is incorporated into the MATSim plans and households file respectively, which are two of the inputs to the MATSim population.

```
<households>
        <household id="2">
                <members>
                        <personId refId="2"/>
                        <personId refId="4"/>
                        <personId refId="7"/>
                </members>
                <income period="year">
                                40000.0
                </income>
        </household>
</households>
```

Figure 3: Extract from Households.Xml File

Since members of a household do not act independently and resources such as income will usually be shared over the entire household, household income instead of agent income is considered, and equally distributed among agents within a household.

This household based agent income is then used when calculating an agent's utility as per Eq. (1-4).

The differences in utility resulting from varying income can then be used to interpret agents' willingness to pay and value of time.

7 CONCLUSION

This paper aimed at highlighting the need to consider individual road users' income in economic policy evaluation such as the Gauteng Freeway Improvement Project. It also highlights the capability of multi-agent microsimulation platforms such as MATSim to incorporate individual income in transport simulation and delivering accurate, close to reality results when evaluating the feasibility and public user acceptance of policy measures such as tolling, which might have regressive effects on the welfare of a population. Not only will this benefit government to make more accurate predictions regarding the acceptance and feasibility of high-cost and high-risk policy implementations, but the public will also benefit if the policies implemented has a smaller regressive impact on them.

8 REFERENCES

[1] **Malahleha, T. 2011**. An analysis of implementing open-road tolling through the Gauteng Freeway Improvemnt Project (GFIP), Master's thesis, University of Stellenbosch.

[2] **Motau, D. 2010**. Eighty percent of Freeway Improvement Project Completed, Engineering News, 25 June.

[3] **SANRAL Ltd. 2010**. Gauteng Freeway Improvement Project Overview. Retrieved October 2011 from http://www.sanral.co.za/gfip/index.php?q=con.23/

[4] **Van der Merwe, J. 2011**. Agent-based transport demand modelling for the South African commuter environment, Master's thesis, University of Pretoria.

[5] **Gao, W. 2009**. Comparisons between MATSim and EMME/2 on the greater Toronto and Hamilton area network, Master's thesis, University of Toronto.

[6] **Fourie, P. 2010**. Agent-based transport simulation versus equilibrium assignment for private vehicle traffic in Gauteng. 29th Annual Southern African Transport Conference, "Walk together", CSIR International Convention Centre, Pretoria, South Africa, 16-19 August 2010.

[7] **Rieser, M. 2010**. Adding transit to an agent-based transportation simulation, PhD dissertation, Berlin Institute of Technology.

[8] **Kickhofer, B., Zilske, M., and Nagel, K. 2010**. Income dependent economic evaluation and public acceptance of road user pricing, Technical report, Transport systems planning and transport telematics. Berlin Institute of Technology.

[9] **Worldbank. 2011**. Gini coefficient. Retrieved November 2011 from http://data.worldbank.org/indicator/SI.POV.GINI?order=wbapi$_d$ata$_v$alue$_2$006+wbapi$_d$ata$_v$alue+wbati$_d$ata$_v$alue-first&sort=asc/

[10] **Gwilliam, K.M. 2007**. The value of time in economic evaluation of transport projects.Worldbank. Transport Infrastructure Notes.

[11] **Kickhofer, B., Grether, D., and Nagel, K. 2011.** Income-contingent user preferences in policy evaluation:application and discussion based on multi-agent transport simulations, Transportation, Springer.

[12] **Müller, K. and Axhausen, K.W. 2010**. Population synthesis for microsimulation: State of the art, paper presented at the 90th Annual Meeting of the Transportation Research Board, Washinton, D.C.

[13] **Müller, K. and Axhausen, K.W. 2011.** Hierarchical IPF: Generating a synthetic population for Switzerland, ERSA.