Estimating Costs and Benefits of Congestion Pricing

Master's Thesis for the degree

Master of Economic Theory and Econometrics

Harald Støen Høyem

University of Oslo

Department of Economics

May 2015

Estimating Costs and Benefits of Congestion Pricing

Summary

Congestion pricing has long been viewed as a viable strategy to mitigate externalities generated in traffic (Walters, 1961),(Pigou, 1924), but is yet to be widely implemented. For several reasons, it is a politically contested policy area. Thus, having information on the economic consequences of introducing a congestion pricing scheme may be important for the ongoing debate.

This thesis seeks to estimate the economic consequences of introducing congestion pricing. To do this, I develop a pricing rule based on a paper by Johansson (1997) applied to multiple periods with heterogeneous users, and integrate it in a model where road users and the government interacts. Using the model formulation, I derive a system of equations which enables me to compute the optimal traffic flow and congestion toll, using data from a single road link in the Oslo cordon. The equation system is parameterized by estimating external cost functions and elasiticities with respect to tolls and generalized costs by econometric models. It is solved by an iterative algorithm called fixed-point iteration.

Combining the estimated models for external costs and parameters, I estimate the change in social surplus from introducing congestion pricing, as compared to the present situation with no congestion pricing. I use a road link on E18 Lysaker-Sandvika, direction towards Oslo, as an application/case study.

My results indicate a positive effect on the social surplus from introduction of congestion pricing. They further seem robust to some of the model assumptions, and the estimates are likely to be conservative. A caveat in the estimation is that I am not able to estimate optimal tolls during rush hours, due to data limitations and features of demand during that period. This does not, however, alter my main conclusion.

Preface

My supervisors, James Odeck at Norwegian Public Roads Administration (head supervisor) and Karine Nyborg at the University of Oslo (supplementary) have both been instrumental in the process of writing this thesis. I wish to thank them for their efforts and constructive feedback. Moreover, Karl Idar Gjerstad, Kjell Johansen and Jan Kristian Jensen at the NPRA helped me with getting most of the data used in this thesis; thanks! Gjerstad also helped me in developing a methodology for estimating environmental costs. Otherwise I would like to thank my fellow students at the Institute of Economics, Kristoffer Berg, Eyvind Søraa and Mikkel Myhre Walbækken in particular, for reading several manuscripts of this thesis and giving me feedback.

A part of the data used in this thesis is gathered from "Den Nasjonal Reisvaneundersøkelsen 2009"[1]. The survey is conducted by Transportøkonomisk Institutt[2]. The data is prepared and made available in anonymous form by Norsk samfunnsvitenskapelig datatjeneste AS (NSD)[3]. Neither Transportøkonomisk Institutt, nor NSD is responsible for the analysis of the data, or the interpretations conducted here.

I would also like thank the Norwegian Public Roads Administration for granting me financial support to complete this thesis.

Any remaining errors are my own.

Oslo, May 2015.

---

[1]English: Norwegian National Travel survey 2009.
[2]English: The Norwegian Centre for Transport Research.
[3]English: Norwegian Social Science Data Services.

Contents

List of Figures

List of Tables

1 Introduction

Congestion pricing has long been viewed as a viable strategy to mitigate externalities generated in traffic (Walters, 1961),(Pigou, 1924), but is yet to be widely implemented. Although some cities have introduced higher price levels during rush hours[4], there is still a large majority of urban centers that have yet to implement such a policy.

In Norway, two cities have introduced higher tolls during rush hours, that is, Kristiansand[5] and Trondheim[6]. Both cities have a cordon-based system in which tolls are higher in typical rush hours (07:00-09:00 and 15:00-17:00 for Trondheim, Kristiansand starting both periods half an hour earlier). Recently, the city of Bergen also decided to implement a similar pricing scheme to combat pollution associated with congestion (BT, 2015).

Estimates of the cost associated with congestion are generally very high. In Oslo, the Norwegian Automobile Union (NAF) estimated that the daily cost will rise to 20 million NOK in 2030 if efficient polices are not implemented (NRK, 2013). Thus, assessing different strategies for reducing such costs could be important.

However, congestion pricing is a politically contested area in Norway, and have generated considerable debate. Arguments focus on the positive effects from lower pollution and time costs, and the negative ones from some users being forced to pay a higher fee for road use, (and some) potentially being unable to drive by car. Another argument is one concerning distribution, that is, how congestion pricing impacts different income groups adversely. Consequently, it will be important to have reliable information on the economic consequences of introducing such a policy for further debate.

The main tool in the economic profession for evaluating the social desirability of policies, is cost-benefit analysis. That is, a unified framework of weighing the costs and benefits of a proposed policy, to see whether or not a policy adds to the total estimated social surplus.

The purpose of this thesis is to undertake a cost-benefit analysis of introducing congestion pricing in Oslo, using a tolled road as a case study. In doing so, more precise information on the attractiveness of congestion pricing as a strategy to mitigate externalities in traffic can be obtained.

Using a congestion price for heterogenous user groups developed by Johansson (1997) as a point of departure, I extend this framework to include multiple periods, and integrate it with the bi-level problem formulation on road pricing. From the problem's optimality condition, I derive a system of equations describing the optimal action of a government agent, road users, and one

---

[4]These include: Milan, London and Singapore.

[5]See http://www.nyekrsbom.no/takster.

[6]See http://www.trondelagbomveiselskap.no/.

equation describing the physical environment in which consumption (road usage) takes place. I use this system to model the trial-and-error procedure proposed by Li (2002) to find optimal traffic demand and congestion tolls.

The thesis is structured as follow: First, I give an overview of the theoretical foundations of congestion pricing, starting with the most simple example. Then I will be increasing complexity to capture relevant effects along the way. This ends in a derived equation system, where the solution characterizes the optimum. Second, I provide some background on the econometric techniques used in this thesis. Third, I parameterize the equation system. This is done by estimating different models for traffic demand elasticities with respect to tolls and motoring costs, and by finding appropriate parametric models to estimate the congestion toll. Fourth, I solve the equation system by a numerical method called fixed-point iteration. Using the estimated optimal congestion toll and level of demand, I perform a cost-benefit analysis of implementing the congestion pricing scheme as compared to today's situation.

All numerical operations were performed using $STATA\ 13^{TM}$.

Last, a comment on terminology: Throughout the thesis, I will use the expression *generalized cost* to signify any relevant motoring costs, except any road tolls paid. The reason being that tolls have special importance in this thesis, and to simplify the language, I decided to treat it separately from other transport costs. The reader is advised that this stands in contrast to the standard definition in the literature, see e.g, (Odeck and Bråthen, 2008).

2 Economic Theory

In this section, the economic theory underlying the empirical analysis in this thesis is outlined. The first part gives a summary of congestion externalities in traffic and how they arise. The second part is based on a paper by Johansson (1997) and develops the optimal congestion price. To establish this price theoretically is the aim of this section. I evaluate under what assumptions Johansson's framework is applicable when considering a single-link[7] and prices that changes over the course of day, as I intend to do, in this thesis. Moreover, the conditions for optimal flow is established, by means of a bi-level optimization problem. The specific problem formulation is chosen as to ease the empirical estimation of the tolls.

2.1 Externalities in traffic

The purpose of congestion pricing is to alleviate the external costs that often are created in congested traffic. There are many examples of such externalities. A non-comprehensive list,

---

[7]A link is a stretch of road that connects two destinations together.

according to Friedrich and Quinet (2011), includes: congestion, traffic accidents, air pollution (local and global), noise and effects on climate change. In this thesis, I will focus on the time cost, and environmental costs, that is, air pollution and effects on climate change, originating from increased congestion. In the following, a short description of how these externalities arise, follows.

In the market for road use, the supply is fixed at road capacity in the short run. This means there exists a certain number of vehicles that can physically pass through a given stretch of road during a given amount of time. Unless road improvements are undertaken, this capacity acts as a regulating factor in traffic demand. As more vehicles enter the road, average speed is lowered due to increased congestion. Travel time thus increases, and raises the average cost of travel, meaning that demand will fall as the generalized costs[8] increases. Moreover, increased demand also imposes an externality in the form of time lost in traffic, that could otherwise have been put to good use, e.g. performing work-tasks or enjoying leisure. That is, each additional vehicle entering the road, will lower the speed of all other vehicles as congestion increases, and not just its own. The effect is marginal for the private cost of the driver, but the aggregate effect for all vehicles may be substantial.

Emissions can also increase as demand increases, raising emissions both directly by increasing traffic volume and indirectly by altering the driving pattern. Lower average speeds often reduce the fuel-efficiency of engines, raising emissions.

These costs are externalities, as the individual driver does not internalize the effect his or her consumption has on other economic agents. They only consider their own time and driving costs, and not the increased travel time their road usage imposes on all other users as well. It can be illustrated graphically as in figure 1[9]. On the x-axis, the flow, defined as vehicles per minute, is seen. This could be regarded as a metric of demand. On the y-axis, the cost in NOK/km is depicted. MSC is the marginal social cost and MPC the private, marginal cost. The marginal private benefit MPB maps out the willingness-to-pay, and thus demand for different price levels.

The marginal private cost is equivalent to the generalized cost of a trip for a road user. As congestion increases, travel time increases and the cost rises as a convex function. The functional shape stems from the fact that marginally adding a vehicle at higher flows, tends to have a more severe effect on speed as compared to lower flows. Thus, as the flow increases (more vehicles entering the road), the curve becomes steeper, and the private cost goes up.

---

[8]Generalized costs is a concept in transport economics in which the price of a trip is regarded as the sum of motoring and time costs, in addition to any tolls or fees (e.g. tickets) that are paid. In this thesis I consider the tolls separately from other generalized costs.

[9]It is assumed here that the curve is not backward bending, concept that will be explored it the next section.

The marginal social cost curve includes the effect of increased demand on the generalized costs of *all* users on the road, in addition to the private cost of the marginal driver entering the road. The relationship between marginal social cost and demand will also be convex, but the former will be even steeper than the private cost curve. As more vehicles enters the road, the number of road users who are affected by the increase in generalized costs from the n-th vehicle entering the road rises. Thus, the marginal social cost curve becomes steeper the higher the flow is.

Moving on the analysis of the model equilibria, in the absence of government intervention, the equilibrium volume of traffic will be at C, where the private marginal cost curve intersects the demand curve. At this point, the willingness to pay for the last driver entering the road will exactly equal the generalized cost associated with the trip. However, as the drivers will only consider their private cost, and disregard the effect their consumption inflicts on other road users, social cost will be higher than the private cost. As the private cost equals the marginal willingness to pay for the last driver, social cost is higher than willingness to pay, and a deadweight loss arises. This is depicted in the figure as the area denoted "deadweight loss".

Figure 1: Traffic demand and social surplus. Adopted from (Button, 2004, p.7) with some minor modifications.
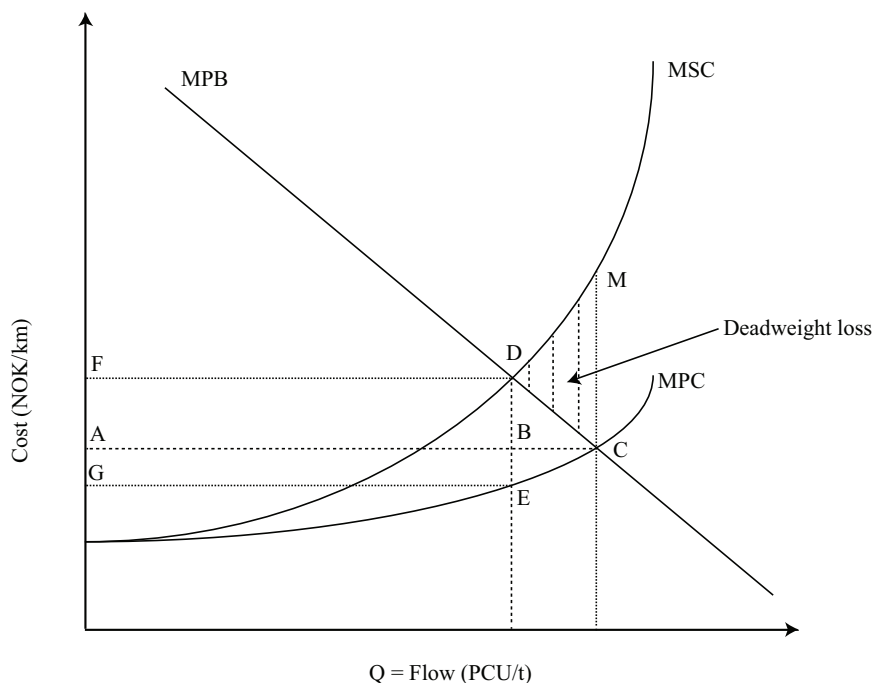
A proposed remedy addressing this problem is Pigouvian (Pigou, 1924) taxation of road consumption, that makes the agents internalize the externality they impose. By imposing an

externality-correcting tax on road use, traffic can be reduced to the level at which social marginal cost is equated to social marginal benefit. This occurs at point D. Traffic and congestion will then revert to point D in figure 1. Note that there will still be some congestion at this point (as $MSC > MPC$).

In this simple framework, it is assumed that traffic is homogeneous, meaning that all users have the same preferences and generate identical externalities, and there is a single time period. The optimal toll τ is then the difference between points F and G. Introducing this toll will generate a change in the consumers surplus equal to $-(FGED + DBC)$. The first, is the loss of welfare for remaining road users due to higher tolls, and the second, the welfare effect from users no longer traversing the link. Notice that the same price is paid by all road users, as the "supply curve", MPC, is not a supply curve in the traditional sense, but rather a curve describing the physical characteristics of the road. To find the social surplus, one must also take into account the increase in government revenue, dubbed "operator's surplus" and the reduction in externalities. The operators surplus is the equal to the increased payments for road users, with the opposite sign, $FGED$. The reduction in externalities, is the area DCM. To summarize, a table with the accounting procedure is given below.

Table 1:

Accounting table of changes to welfare from introducing congestion pricing. All quantities are measured in NOK. It shows how the loss in welfare from consumers being forced out of road usage due to higher tolls (CS), is weighed against the benefits of a reduction in social costs (ER) to produce the total change in social benefits. The increased revenue to government (OS) is netted out against the increase in trip price from those having a high enough marginal willingness to pay to still use the road (FGED), as this has no real effects on the agents behavior (they still use the road).

Consumer's surplus (CS)	Loss from increased tolls	-FGED
	Loss from diverted traffic	-DBC
	Sum	-(FGED+DBC)
Operator's surplus (OS)		FGED
Externality reduction (ER)		DCM
Social surplus		CS+OS+ER = DCM-DBC

As the increased government revenue from tolling is just a transfer from the users to the government, it is netted out in the social surplus. The transfer has no real effects, compared to the reduction in traffic, that is DBC. An implicit assumption here, is that the allocation in all other markets is unaffected by a change in the toll. Thus, users remaining on the road are assumed to not change their consumption of other goods, when the price of undertaking a trip is changed. This means that only the loss from users leaving the road, and the reduction in externalities constitutes the change in total social surplus. The social surplus estimation procedure is given in chapter 8.

2.2   The optimal congestion price with no bottleneck

The optimal congestion price is *the price that makes the users internalize the externalities they impose on others.* In reality, there are often many different types of road users, who may have different evaluations of time and generate quantitatively different external costs. The above analysis did not account for this heterogeneity, as traffic and users were assumed to be homogeneous. In this section, I derive the optimal congestion price using Johansson (1997) as a point of departure, where heterogeneous user groups are facing a congestion price determined by the size of the externalities they generate. The user groups are defined as light and heavy vehicles. Moreover, the exposition given here implicitly assumes there are no bottlenecks on the link. This means that the flow capacity[10] of the road is never reached, at any point along the road. The case with a bottleneck will be discussed in the next section.

2.2.1   The Johansson (1997) model

Johansson (1997) develops a model where there are n different types of road users, which differ according to time valuation, emissions, and congestion factors, that is how large externalities they generate. The model describes the problem of a social planner who seeks to maximize social welfare by adjusting tolls and traffic flow on a road network. A description of this model, with some minor modifications, will now follow.

A key variable in traffic analysis is "*flow*", defined as the number of vehicles that passes through a specific point per time unit. This will be used as a metric of traffic demand in this thesis. Let the flow of vehicles in group i be denoted as Q_i. The total flow of vehicles at the

---

[10]This is the maximum number of vehicles than can physically pass through the road link by a given amount of time.

road link[11] under consideration is given by the sum of flows of different user groups[12], $i = 1, ...n$,

$$Q = \sum_{i=1}^{n} Q_i \tag{1}$$

I will refer to this quantity as unadjusted flow. As flow is measured as the number of vehicles per time unit, it does not differentiate between light and heavy vehicle's impact on congestion, as heavier vehicles occupy more space than light ones. To account for this, several authors, among them Sen et al. (2010) and Mayeres et al. (1996) use what is commonly named *congestion factor*, which also Johansson (1997) does. This is a scalar which adjusts the flow variable according to its relative impact on congestion. Denote this as α_i for group i. A congestion factor of 1.2 thus means that the vehicles of group i generate an impact on congestion that is 20 % higher than the vehicle group generating the lowest impact[13]. The product between the congestion factor and unadjusted flow is referred to as *passenger car equivalent*. The flow used in the model is thus named the passenger car equivalent flow, defined as

$$Q_r = \sum_{i=1}^{n} \alpha_i Q_i \tag{2}$$

where the flow of group i is scaled by its relative impact on congestion to produce a flow variable where all vehicle counts are measured using the same unit, that is, passenger car equivalents. The r in the subscript is used to distinguish the passenger car equivalent flow, from the unadjusted flow, and is adopted from Johansson (1997).

Moreover, it is assumed that the government wishes to maximize welfare by maximizing the sum of net social benefits over n different user groups. Johansson (1997) does not explicitly declare any assumptions regarding the normative views of the government. The problem of the government is formulated by maximizing the following function

$$NB = \sum_{i=1}^{n} \int_0^{Q_i} \left[MB_i(Q_i) - MC_i^0 - MC_i^e(V(Q_r)) - MC_i^t(V(Q_r)) \right] dQ_r \tag{3}$$

$V(Q_r)$ is a function that relates overall link flow Q_r to speed. $MB_i(Q_i)$ is the marginal benefit per kilometer (willingness to pay for road use) at flow Q_i for user group i measured as NOK/km, MC_i^0 is a marginal cost for group i, that is fixed with respect to flow. $MC_i^e(V(Q_r)), MC_i^t(V(Q_r))$ are environmental and time costs per kilometer in NOK for user group i[14]. Both the environmental and time costs are functions of speed. For time costs, lowered speed means more time is needed to

---

[11]A link is a stretch of road that connects two destinations together.

[12]User groups are meant to represent road users having different purposes for their trip.

[13]In a later section, a functional relationship between congestion and flow (demand) is presented. The congestion factor effectively scales this function by α.

[14]Johansson also included fuel costs per kilometer, which I will omit, due to data constraints.

traverse the link, and for environmental costs, lower speed implies higher emissions per kilometer emitted by group i. Thus, both $MC_i^e(V(Q_r))$ and $MC_i^t(V(Q_r))$ are functions of the passenger car equivalent flow, which again determines speed. The reason costs are measured on NOK per kilometer rather than at flow, is the spatial dimension of consumption. That is, the number of vehicles does not describe consumption in isolation, the travel distance is also needed, as longer travels require more fuel, and produces higher emission, etc.. The conversion thus normalizes the demand (flow) at one unit of distance (km), such that the total effect is found by scaling costs and benefits to the appropriate travel distance for each group of users.

The relationship between the different variables and functions is depicted in the figure below. First, flow determines speed through congestion. Second, decreased speed increases travel time. Emission are raised both directly and indirectly, as more vehicles increases emissions directly as more vehicles enters the road. Moreover, emissions are raised indirectly as lower speeds decreases fuel efficiency.

Figure 2: Flow diagram of model in Johansson (1997).
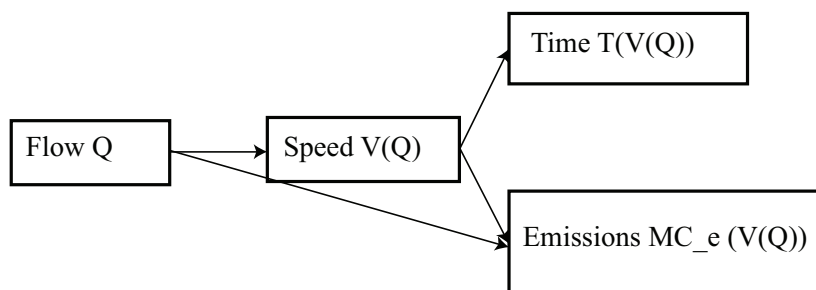
The marginal benefit per kilometer $MB_i(Q_i)$) in NOK/km for user group i, is only a function of the flow of vehicles in that specific group. The reason being that the marginal benefit is assumed to differ according to different user groups. It is important to distinguish here between the marginal benefit and net benefit to road users. The net benefit is dependent upon the flow in other user groups, as these users affect affect time costs. The marginal benefit (willingness to pay for road use) is however assumed to be independent of demand in other groups. The total benefit within user group i is found by integrating net benefits of the group over the traffic volume Q_i of that group. The total benefit over all groups is then just the sum over the individual group's net benefit.

Now, a small change in notation is made which differs from Johansson (1997). I assume the marginal value of time is constant, and further assume that the speed function $V(Q_r)$ is

invertible[15]. Travel time is then $T_i \equiv V(Q_r)^{-1}$, where $T_i(Q_r)$ is the time required to travel *one* kilometer, measured in hours. Letting VOT_i designate the value of time in NOK/hour for group i, the private marginal time cost for group i is equal to the valuation per hour, multiplied by the time needed to travel one kilometer, measured in hours:

$$MC_i^t(V(Q_r)) \equiv VOT_i \times T_i(Q_r) \tag{4}$$

The government planner's problem is to find the condition characterizing optimal toll for each user group. This is found by first finding the first-order conditions for each user group by taking the derivative of the objective function with respect to flow in group i. Then, finding the first order condition with respect to the same variable in the market solution without any government intervention. Lastly, one solves the toll for group i, τ_i that adjusts flow to the socially optimal level, using the difference between the two first-order conditions.

Proceeding to the first step, one first substitutes in for (4) into, (3). Then, taking the derivative of the government planner's problem with respect to unadjusted flow in group k, Q_k, to find the following first-order condition:

$$MB_k(Q_k) - MC_k^0 - MC_k^e - VOT_iT_i - \alpha_k \frac{\partial V}{\partial Q} \sum_{i=1}^{n} Q_i \left(\frac{\partial MC_i^e}{\partial V} + VOT_i \frac{\partial T_i}{\partial V} \right) = 0 \tag{5}$$

The intuition behind this condition is the following: A marginal increase in flow for user group k increases the benefit as the marginal driver in group k, enters the link. This is the first term. The second term is the constant cost component. The third is the direct marginal social environmental cost. The fourth shows the increase in travel time for the marginal driver that enters the road; as he enters, he spends T_i hours on the link, which is a loss to him of $VOT_i \times T_i$ NOK. The last term contains the externalities. The term inside the brackets are the marginal effect of environmental costs to society, and costs from increased time for user group i. This is scaled by the number of users in that particular group. This is summed over all groups to find the overall increase in the marginal cost for a marginal change in speed. To find the overall effect, one also needs to establish the effect of the marginal increase in flow on speed. This is comprised of the marginal change in speed as flow increases ($\frac{\partial V}{\partial Q}$), scaled by the congestion factor for user group k (α_k), which gives the total effect from a marginal increase in the flow on costs. The optimum corresponds to the point D in figure 1, if we interpret the figure as depicting the flow of group k.

Now, I will move on to the private user equilibrium, that is, the point C in figure 1. In this case, the externalities are not taken into account by the users; they only consider their own

---

[15]This assumption is crucial in being able to calculate a congestion price. If traffic is a "traffic jam", in an equilibrium state, this assumption may not hold, as will be pointed out later in the thesis.

private costs. This means that they do not solve the government planner's problem, but rather their own private optimization problem. This is comprised of the marginal benefit of undertaking the trip, and the marginal costs. Adding to the content of Johansson (1997), let T_i be the number of trips in user group i. Assuming trips equals flow, the private optimization problem of each user group problem can then be stated as the difference between the marginal willingness to pay for Q_i trips and the costs of those trips:

$$\max Q_i(MB_i(Q_i) - MC_i^0 - VOT_iT_i) \tag{6}$$

where $MC_i^e(V(Q_r))$ is left out, as drivers are assumed not to be concerned with the environmental externalities. Assuming the change in flow from increasing the number of trips by one unit is the same as the marginal increase in the flow $(dQ_i/dT_i = 1)$, the first-order condition with respect to trips Q_i, can be written as:

$$MB_i(Q_i) - MC_i^0 - VOT_iT_i - \alpha_k\frac{\partial V}{\partial Q}\left(\frac{\partial T_k}{\partial V}\right)Q = 0 \tag{7}$$

The difference between this condition and the government planner's problem, is that the sum over all user groups are not included (the term $\sum_{i=1}^{n} Q_i$ is excluded), and that the environmental cost term is excluded. This is because the users do not consider how their presence on the road affects all other users, or the impact on the environment. In general, the effect on speed from one extra vehicle on the road will be very small. Thus, assuming the marginal effect of one vehicle on the speed is close to zero ($\frac{\partial V}{\partial Q} \approx 0$), one can rewrite the condition as if the last term is zero. As from the perspective of the last road user adding to the flow of group i, the product between speed and marginal costs will be very small, then (7) becomes

$$MB_i(Q_i) - MC_0^i - VOT_iT_i = 0 \tag{8}$$

This is just the private cost, which is not equal to the first-order condition of the government planner's problem in general. The reason is that the marginal change in costs from a change in the speed, is assumed to be so small, that it will not be considered as relevant by a single road user.

The optimal congestion price is then *the price that makes the users internalize the externalities they impose on others*. This means that the first-order condition of the private and and government planner's problem should be equal, which is the point D in figure 1. The optimal tax can be found by taking the difference between the social and private optimum condition, that is[16]

$$\tau_k = MC_k^e + \alpha_k\frac{\partial V}{\partial Q_r}\sum_{i=1}^{n} Q_i\left(\frac{\partial MC_i^e}{\partial V} + VOT_i\frac{\partial T_i}{\partial V}\right) \tag{9}$$

---

[16]The original paper by Johansson (1997) contains an algebraic error, giving a minus in front of the α.

That is, the optimal tax consists of the direct effect on the environment through emissions, and the indirect effect on emissions through lower speed, and the total time lost in traffic for the user groups. The effect is also dependent on the congestion factor, α, that is, how the vehicles of group i affects the overall speed on the link in relation to the physical size of its vehicles. Relating the equation to figure 1, the first term corresponds to the direct arrow from flow to emissions. The terms inside the brackets relate to the arrow going from speed to environmental and time costs.

The optimal toll, could be interpreted as a toll per PCE when the direct effect on environment from increased traffic flow is left out:

$$\tau_k = \alpha_k \frac{\partial V}{\partial Q_r} \sum_{i=1}^{n} Q_i \left(\frac{\partial MC_i^e}{\partial V_k} + VOT_i \frac{\partial T_i}{\partial V} \right) \tag{10}$$

Setting $\alpha = 1$ gives the optimal toll for one extra vehicle entering the road of PCE class 1, that is, a passenger car. To find the toll for any other PCE unit, one just scales the toll by the appropriate congestion factor α. For heavy vehicles, the toll becomes

$$\tau_H = \alpha_H \tau_L \tag{11}$$

where τ_L is the congestion factor for passenger cars (with $\alpha = 1$), and α_H is the congestion factor for heavy vehicles.

The value of the flow needs to be Q^E, the optimal value that occurs at the point D in figure 1. Johansson (1997) does not prescribe a direct method to find this value, which will be done in the last subsection of this chapter. First, some more immediate limitations of the model will be highlighted.

2.2.2 Limitations of the model

The Johansson model does not account for the case of multiple periods. As it is an aggregate model, it does not consider the possibility of changes in the road network, or between different modes of transportation. In this section, I discuss under what assumptions the model is still applicable to a single link, as I intend to use it. I claim that under assumptions of independent demand across time, the pricing rule is still valid, period-by-period. This will be important since I intend to calculate a congestion price that varies over the course of the day. Moreover, I discuss the effect of only tolling one single link, and show that a similar assumption needs to apply. Lastly, I discuss some possible implications of changes in transportation mode. Derivations relating to the assumptions on the model can be found in the appendix.

Multiple time periods: If demand is not independent over time, the full effect of changing the price vector over time, $\boldsymbol{\tau} = \tau_1, ..., \tau_T$, might not be identified, as changing the tolls in one period affects the social surplus in other periods as demand is shifted between periods. For example, if raising the toll in one period shifts demand towards another period, the externalities will be reduced in one period, but possibly increased in another, meaning that the net effect on social surplus can be lower than what is believed when considering one period only. Moreover, if prices are changed in both periods, the net effect may be ambiguous. The reason is that traffic volume and size of externalities are connected by a convex relationship, opening the possibility that shifts from high to low demand periods can have an overall positive net effect on the social surplus. Thus, it is hard to produce any general statements, without conducting a more formal analysis, which out of scope, for this thesis. To use the specific results concerning the road price in this thesis, I will make the following simplifying assumption:

Assumption 1: Demand is independent over time

The assumption entails that changing the toll in one period, does not affect demand in other periods of the day. The appropriateness of this assumption may vary over the course of the day. For example, during the morning rush hours, demand is more peaked, which might suggest that users have less flexibility in the timing of their road use, whereas the afternoon traffic is more evenly spread out in time.

Multiple routes: A second caveat follows the same logic as above, but restates it. There might not just be possibilities for road users to substitute between different time periods, but also different routes. As such, raising the price at one geographical location (tolling station) might divert traffic to another route, increasing congestion and externalities, which leads to an overestimation of the correct congestion price, as the adverse effects on other routes are not accounted for. This thesis relies on the following simplifying assumption

Assumption 2: There are no externalities in the road network from altering tolls.

When the assumption holds, there will be no substitution from using the road link to other road links. If there is in fact some substitution from the tolled road to the untolled road, then as traffic increases on the other link, marginal social costs also rise, introducing an externality.

Multiple transport modes: Making use of the road by car is not the only viable option for transportation in many urban areas. Various forms of public transportation, along with walking and cycling, is also an option for movement. One may view the market for transportation as constituted by many different sub-markets. This means that not only is the route considered,

but also the mode of transportation mode. The analysis in Johansson (1997) did not take this into account.

There are two ways in which substitution between modes might affect the estimate of social benefits. First, through the welfare loss of users no longer using the road, and second, through the change in externalities. Concerning the first point, if users substitute towards another mode of transportation, it means that the consumer's surplus in the market for road use is diminished, whereas the consumer's surplus in the market other modes, may be increased. If one considers more than one market, the overall social benefit of congestion pricing may thus be higher than what is estimated when considering a single market. Second, according to Boadway (2006), evaluations of public policy projects should also take into account the changes, not just in the market studied, but also any change induced by the policy project in other distorted markets. Suppose that the increased tolls make some users substitute towards public transportation. Associated externalities in this mode, are, among others, discomfort from crowding (small space). If more users enter the market for public transport, without a proportional increase in capacity, an externality might be created or exacerbated. Not all types of modes might induce such an externality. Walking or cycling may or may not be prone to such crowding or comfort externalities, and can even be beneficial through increased health.

To make the analysis valid, there needs to be a further assumption that rules out such effects as described above. Either, there must not be externalities or welfare effects in general from in other markets, and/or, there must be no substitution between different modes.

Assumption 3: There is no substitution from road use to other modes of transportation and/or no external or welfare effects in other markets induced by shifting demand to road from other transportation markets, or the other way around.

I will proceed by working under the stated assumptions 1-3, holding by conclusions contingent on their validity.

2.3   The optimal congestion price with a bottleneck

The depiction in figure 1, is a simplified version of the congestion mechanics on a road. In reality, the time flow curve may be backward bending, implying that the function relating flow and travel time is no longer invertible. This was an important assumption in the model of Johansson (1997). I will now explain why this curve may become backward bending, and the implications for tolls. Even though I am not able to use this theory in practice due to data limitations, I nevertheless think it is worthwhile to explore possible points of improvements.

The time-flow curve can be seen in figure 3(a). The curve depicts how flow and travel time

per kilometer are related. As more vehicles enter the road, meaning higher flow, the time needed to travel the link increases up to the maximum flow. At this point, the curve becomes backward bending. Here, the effect on speed from adding another vehicle to the road, has a larger effect on the flow, than adding an extra vehicle. Flow drops, as traffic goes towards a jam-like state, as fewer vehicles are able to pass through the road link at each time point. As a consequence, speed is lowered, and the time needed to travel one kilometer increases. Thus, the relationship between flow and time is no longer invertible, as two different values of time corresponds to the same value of flow.

Figure 3: Two figures displaying the time flow relationship on a single road link. It shows how the needed time to travel one kilometer is related to the number of vehicles passing trough the road link per time unit. Panel (a) adopted from (Button, 2004, p.6) with some minor modifications. Panel (b) Adopted from (Yang and Huang, 1998) with some minor modifications.

(a) The time flow relationship.

(b) High demand $D2$ and low demand $D1$.
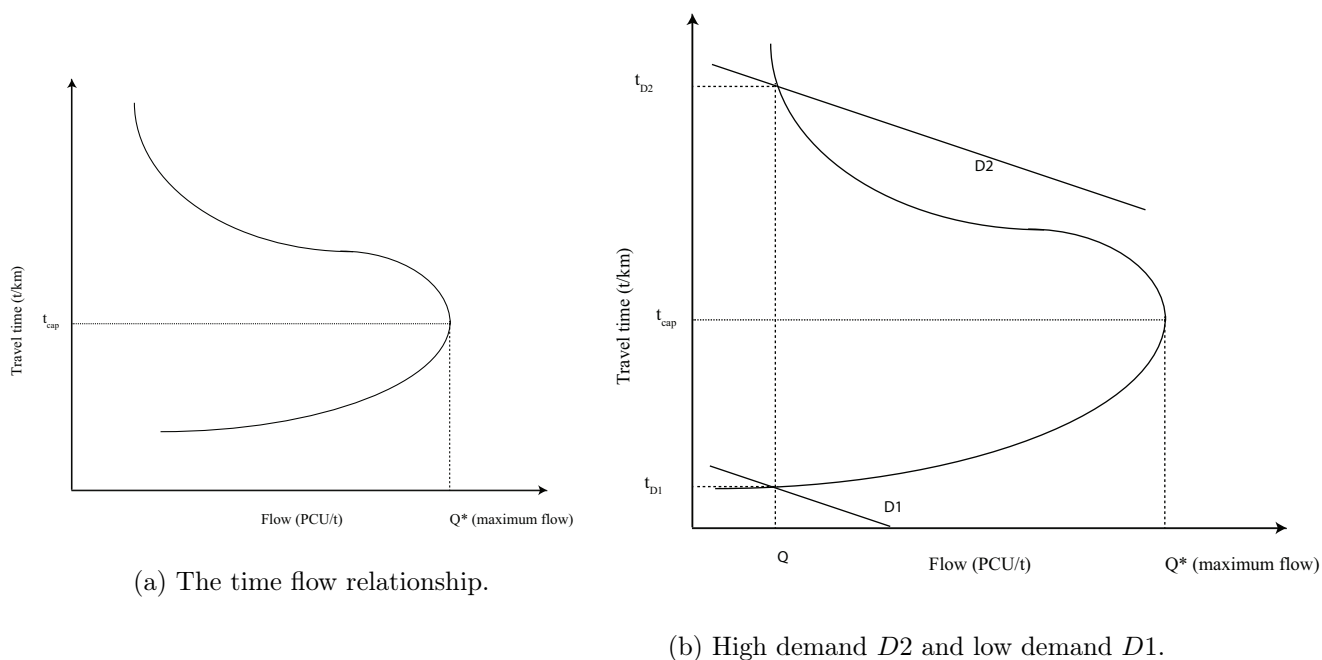
If there exists an equilibrium for some time period on the backward bending part of the curve, this will have implications for the congestion toll. The situation is illustrated in the figure 3(b), adapted from (Yang and Huang, 1998). Here there are two demand curves, $D1$ being demand in a non-rush hour period, and $D2$ demand in the rush hour. The y-axis measures the travel time, and thus, indirectly, time costs, and the x-axis, flow. Here, there are two equilibria for the same value of flow. One, where the rush hour demand curve intersects the high-end of the time-flow curve, and one where intersection is at the lower curve, $D1$. As the cost in the rush hour is significantly higher, in this simple figure, it implies that users are willing to stay in a traffic jam, since their marginal benefit of undertaking the trip is sufficiently high.

In the model of Johansson (1997), a main assumption is that the function relating time to flow $T(Q_r)$ is invertible, that is, a specific level of flow corresponds to one level of time. If traffic flow is in an equilibrium on the backward-bending part of the curve, this function is not invertible. The implication is that I am not able to use the current model to find the marginal external costs, if there exists an equilibrium on the backward bending slope of the flow-time curve.

The literature on congestion pricing has tackled the problem of a backward bending flow-time relationship from two separate angles, where some of the problems described above are dealt with, at the expense of introducing new ones.

The first, is represented by a series of papers where bottleneck congestion is explicitly modeled by a structural model as in (Ben-Akiva et al., 1984), (Arnott et al., 1990) and (Arnott et al., 1993). In these models, only peak period congestion pricing is considered, and the focus is mainly on how departure times during rush hours are affected by congestion pricing, and smoothing of demand over the rush period. Arnott et al. (1990) considers a model with inelastic demand, whereas Arnott et al. (1993) also treats demand as elastic. This approach is in my view, a more specialized approach devoted to the peak problem only, with attention mainly towards distributing demand in an optimal way.

The second approach, is represented by series of papers where the analysis is based upon the flow-time relationship, and does not explicitly model departure times, but rather extends the "traditional" approach. Mun (1994) provides, to the best of my knowledge, the first such analysis. His main result is that optimal peak load prices should be set according to the relative time increase a queue imposes on the other users of the road. A similar treatment of the problem is given in (Yang and Huang, 1998). They consider a road link where there is a bottleneck characterized by a lower capacity than the rest of the road link. Let Q_{CAP} be the capacity of the bottleneck, defined as the number of cars that can pass through its end during a given amount of time. Let Q_{MAX} be the capacity of the road section in total, that is outside the bottleneck. They show that the optimal congestion price should equal

$$\tau = \begin{cases} \tau_1(Q) & \text{if } 0 \leq Q \leq Q_{CAP} \\ \tau_1(Q) + \lambda & \text{if } Q = Q_{CAP} \end{cases} \tag{12}$$

where $\tau(Q) = L/S_1(Q)$ and

$$\lambda = L'[1/S_2(Q_{CAP}) - 1/S_1(Q_{CAP})] \tag{13}$$

Here $S_2(Q_{CAP})$ is the time flow-relationship on the backward-bending part of the speed-flow curve, $S_1(Q_{CAP})$ for the lower part, L is the length of the road and L' is the length of the queue. I will now briefly explain why I am not able to use these theoretical results in my thesis.

The first approach requires explicit assumptions on the rush hour queue 's start and end point to find an optimal congestion toll. Moreover, the model only treats the peak-period, leaving the other periods, possibly with congested traffic, out. A clear benefit is simple expressions for the congestion tolls and dependence on few parameters. This, however, also leaves out treating heterogeneity of road users, a point the authors stress themselves (Arnott et al., 1993), which is a main purpose of this thesis.

The second approach solves some of the problems of the first, and leaves some new ones behind. First, it extends the flow-time modeling approach. This means it shows how congestion prices should be set in traffic jam-like conditions, that is, in peak-load time intervals, on the basis of the approach I already have followed. In this setup, the time of queue initiation is determined endogenously as the time when flow capacity is reached. A problem arises in that the length of the queue , L' must be determined. According to Yang and Huang (1998), this should ideally be set in real time. Even though the problem of the backward-bending curve is solved principally, the practical problem of queue estimation remains. This means that I will not be able to calculate the optimal congestion tolls during rush hours if there is an equilibrium on the backward bending part, as I would need additional data on queue length.

2.4   The optimal flow

When estimating the congestion prices, one needs to find the equilibrium value of flow which is the relevant one for conducting an analysis of changes in social benefits. When using data, the estimated external costs at present are equal to the difference between M and C in figure 1. Using this as a price will not in general induce a demand equal to the optimal level of demand, associated with the equilibrium at D. The reason is that when $M - C$ is priced, traffic reverts to a level below the optimal flow. This will in turn reduce the externalities and hence tolls, in addition to generalized costs. This sets the system in a new state. So, one needs to find a level of demand where the induced level of toll and generalized cost would induce exactly that level of demand. The typical answer in the literature to this challenge, is to find an inverse demand function to solve for the toll at the equilibrium (De Palma and Lindsey, 2011). However, demand functions are very hard to observe in practice. Li (2002) proposed a trial-and-error practice in which the government sets a toll based on present externalities, observe a new demand, sets a new toll, and so on, until the optimal toll and flow is found. (ibid.) shows how the procedure converges to the desired flow. I do not have the opportunity of undertaking such a real-world experiment, and will instead try to model the process which Li (2002) proposed, by using a bi-level optimization problem.

In such a problem, the government maximizes an objective function, in which one of the variables is given as the solution to a lower-level problem, solved by the user groups. I solve such a problem, and model the trial-and-error procedure by using a system of equations, derived from the solution to the problem. In other words: this section integrates the bi-level approach with the results derived in the previous section, such that I am able to estimate a set of equilibrium flows and tolls by solving for the equilibrium through mimicking the process proposed by Li (2002).

2.4.1   A bi-level optimization problem

The literature on congestion pricing in traffic networks[17] has turned to analyzing the equilibrium problem by means of a bi-level optimization program, where both a planner and a user optimize their decisions simultaneously. In his doctoral thesis, Ekström (2008) considers such a problem for a whole network where some routes are left untolled. De Palma and Lindsey (2004) treats the problem in a *general equilibrium* setting with considerations of income taxation and equity preferences are taken into account. Yang and Bell (1997) uses a bi-level approach to study optimization of various criteria functions, among them, *consumer's surplus*. Ferrari (2002) optimizes *social surplus* on a tolled road network using a similar approach. Chen et al. (1999) seeks to minimize *total travel cost* on a network by using the bi-level formulation. Finally, Chen et al. (2004) uses the bi-level program for toll design with *stochastic route choice*. Thus, the bi-level formulation is widely used in the more technical literature on road pricing, and seems to be an accepted way of treating the problem of interaction between the decisions of a planning agency, and an optimizing agent.

Problem statement: I will now implement this approach by extending the Johansson model to incorporate the dependence of traffic volume on tolls. In general, a bi-level program is defined, according to Balwani and Singh (2009), as

$$\max_{\boldsymbol{\tau}, \mathbf{Q}} F(\boldsymbol{\tau}, \mathbf{Q}(\boldsymbol{\tau})) \tag{14}$$

subject to

$$G(\boldsymbol{\tau}, (\mathbf{Q}(\boldsymbol{\tau})) \leq 0 \tag{15}$$

where $\mathbf{Q}(\boldsymbol{\tau})$ is implicitly defined by

$$\min_{\mathbf{Q}} f(\boldsymbol{\tau}, \mathbf{Q}) \tag{16}$$

subject to

$$g(\boldsymbol{\tau}, \mathbf{Q}) \leq 0 \tag{17}$$

---

[17]More than one road link.

Q is vector of demand in T periods, and $\boldsymbol{\tau}$ is a vector of tolls. Here, one maximization problem depends on a variable which is implicitly defined by another maximization problem. The upper-level part of the problem is the function F. It corresponds to the government planner's optimization problem, where the congestion toll τ is determined. G is the constraint on the problem. f is the response function of users with respect to tolls, constrained by a set of conditions g. The solution to this problem implicitly defines flow, $\mathbf{Q}(\boldsymbol{\tau})$. This formulation is intended for a network, whereas I only study a single link, which reduces the complexity.

I now extend the Johansson model to include the equilibrium setting, by defining the different objective functions for period t, F_t and f_t in accordance with the general setup above. The equations become

$$\max_{\tau, Q} F_t = \sum_{i=1}^{n} \int_0^{Q_i} \left[MB_i(Q_i) - MC_i^0 - MC_i^e(V(Q_r)) - VOT_iT_i - \tau \right] dQ_r \tag{18}$$

and

$$\min_{Q} f_t = -Q_i(MB_i(Q_i) - MC_i^0 - MC_i^e(V(Q_r)) - VOT_iT_i - \tau), \forall i \tag{19}$$

the associated conditions being non-negative constraints on traffic flow Q_i, and τ being any other payments made to conduct a trip, such that the net benefit of a trip, is the marginal benefit, lest the generalized cost and other payments. τ is the still the toll, taken as constant by the road users (in the f-problem), and as a decision variable in the government's problem.

The first problem, or "upper-level problem" as it often is called, is the government's planner problem, where both the optimal flow and toll is determined simultaneously. The second, or "lower-level" problem, is the road users optimization problem, where the net benefit of undertaking a trip is maximized for n user groups, by minimizing its negative, to keep the formulation consistent with Balwani and Singh (2009). This means the planner decides a level of $\boldsymbol{\tau}$ in which the road users take into account when solving for their demand $\mathbf{Q}(\boldsymbol{\tau})$. Thus, the reaction of the road users to the tolls set by the planner, implicitly defines the demand, which is the essence of the simultaneity mentioned above.

It is further assumed that the government planner maximizes these function for T periods. Let the subscript t designate that the variable belongs to period t. I follow Ferrari (2002) in assuming that the different periods can be maximized separately. That is, the total maximization entails finding optimal Q and τ for all T periods, which is equal to maximizing the functions for each time interval, such that

$$\max \sum_{t=1}^{T} F_t(\tau_t, Q(\tau)_t) = \sum_{t=1}^{T} \max F_t(\tau_t, Q_t(\tau_t)) \tag{20}$$

where, again, $Q_t(\tau_t)$ is implicitly defined by τ_t. I will now show how this problem can be solved

for optimal flow and tolls on a single link. This assumption mirrors assumption 1, saying there is no substitution between different periods.

Problem solution: To find the solution to the above maximization problem, I start by looking at the upper-level problem. Taking the derivative of this function wrt. Q treating it as fixed with respect to tolls, gives me the first order condition from Johansson (1997). However, as we saw in the previous section, this condition implicitly defines the toll as a function of the flow as $\tau(Q)$. This condition was

$$\tau_k = \alpha_k \frac{\partial V}{\partial Q_r} \sum_{i=1}^{n} Q_i \left(\frac{\partial MC_i^e}{\partial Q_k} + VOT_i \frac{\partial T_i}{\partial Q_k} \right) \tag{21}$$

Thus, the planner simultaneously solves for the optimal price and toll differentiating with respect to the flow, since the toll is implicitly defined by the first order condition to flow. Moreover, flow is implicitly defined by the toll through the reaction function of the road users, f. The demand level Q is found as the solution to the lower-level problem with respect to Q. The first-order condition in this case becomes

$$MB_i(Q_i) - MC_i^0 - VOT_i T_i - \tau = 0 \tag{22}$$

assuming the same conditions apply to the private maximization as in the last section. Denote $GC(Q) = MC_i^0 + VOT_i T_i$ as the generalized costs. Now, letting τ depend on Q brings the reaction of the planner into the problem of the road user. At this point, the road users affects the decision of the planner by adjusting their demand in accordance with the toll set by the planner, and the planner adjusts the toll in accordance with the size of flow (demand) as determined by the road user. To proceed, rewrite the road user's first order condition (22) as

$$MB_i(Q) = GC(Q) + \tau(Q) \tag{23}$$

This is the standard economic condition saying the marginal benefit should equal the marginal social cost. This condition implicitly identifies the equilibrium value of flow as the one that makes this equality hold. This is also the optimal value of flow. More precisely, the optimal traffic flow can be defined as the one in which the marginal benefit of using the road link, is equated to the social cost of doing so[18]

$$Q^E = \{Q : MB_i(Q) = GC(Q) + \tau(Q)\} \tag{24}$$

Next, note that the marginal benefit is equal to the inverse demand function $MB_i(Q) = p^{-1}(Q)$. Let the demand function be denoted by $v(Q)$. The inverse demand function has the property

---

[18]The solution here is assumed to satisfy the second order conditions for a global maximum.

that $Q(p) = v(p^{-1}(Q)) = v(MB(Q)) = Q$. Thus, if $v(\tau(Q) + GC(Q)) = Q$, this implies that $MB_i(Q) = GC(Q) + \tau(Q)$, and the first-order condition for the lower-level problem is satisfied. This means the equilibrium condition, and thus optimal flow, could also characterized by

$$Q^E = \{Q : v(\tau(Q) + GC(Q)) = Q\} \tag{25}$$

When this condition is satisfied, with the formula for $\tau(Q)$ derived in the upper-level problem, both first-order conditions hold, and an optimal solution is found. I proceed by assuming that the problem satisfies the second order conditions for a local optimum, and further, that the optimum is unique.

2.5   The trial-and-error equation system

I will use the equilibrium condition in the estimation chapter to solve for the optimal toll and flow. I now define an equation system which solution gives me the estimated equilibrium of flow and tolls. I start by exploiting the equilibrium condition. It contains three equations, which together describe the equilibrium. That is, I have a system of three equations:

$$Q(\tau, GC) = v(\tau, GC), \tag{26}$$

$$\tau(Q) = g(Q), \tag{27}$$

$$GC(Q) = a(Q) \tag{28}$$

These relationships are summarized as follows: Q is the demand function, and assumed to be the solution to the lower level problem f for given τ and GC. τ is the toll, and is the solution to the government's optimization problem for given Q. GC are the generalized cost, that is, motoring and time costs for given flow Q. Now I have three equations describing 1) The optimal strategy for the government, given its objective 2) The optimal strategy of road users and 3) the behavior of the physical environment (the road link) in which consumption takes place, that is, the private cost function. Solving this system sequentially by an iterative procedure, is my substitute for finding the optimal flow and toll, by mimicking the trial-and-error process proposed by Li (2002).

The rest of the thesis is concerned with solving this system, and evaluating the implied consequences for social benefits. To accomplish this, the equations will need to be parameterized. The next chapter gives some background on the econometric procedures used for this purpose, and the section afterwards, gives the models and results used parameterize this system, before I use a numerical method in conjunction with the estimated parameters to solve for the optimal flow and tolls.

3 Econometric Theory

In this section, an overview of the different econometric techniques used in this thesis is given. This is to provide readers with some background on the methods employed that are not commonly used in econometric practice. The exposition in all of this section relies heavily on (Davidson and MacKinnon, 2004, ch. 6, 12.2), and partially on (Cameron and Trivedi, 2005, ch. 5.8) and (Wooldridge, 2010, ch. 12).

3.1 Nonlinear least squares (NLS)

Non-linear least squares is a convenient way to estimate parameters of non-linear functions (Davidson and MacKinnon, 2004, ch. 6). The method is chosen as it is easy to implement directly without performing any transformations on the functions that would otherwise be necessary for estimation by OLS. In the following, a brief summary of NLS estimation is given.

3.1.1 The Nonlinear Regression Model

Let y_i be the dependent variable and Ω_t represent an information set (e.g. a set of explanatory variables), then, nonlinear regression model is defined as

$$y_t = x_t(\boldsymbol{\beta}) + u_t, u_t \sim IID(0, \sigma^2), t = 1, ..., n \tag{29}$$

where $x_t(\boldsymbol{\beta})$ is a scalar non-linear regression functions of a parameter k-vector $\boldsymbol{\beta}$. The function determines the conditional expectation of y_t conditional on the information set containing the explanatory variables (Ω_t), as $E(y_t|\Omega_t)$, when this function is non-linear. The errors terms are assumed to be IID conditional on the information set, which entails that any two error terms are expected to follow the same distribution and be uncorrelated over time. For convenience, I rewrite the model in a more compact form as

$$\mathbf{y} = \mathbf{x}(\boldsymbol{\beta}) + \mathbf{u}, \mathbf{u} \sim IID(\mathbf{0}, \sigma^2\mathbf{I}) \tag{30}$$

Here, $\mathbf{y}$ and $\mathbf{u}$ is an n-vector with typical elements y_t and u_t, $\mathbf{I}$ being the identity matrix and $\mathbf{x}(\boldsymbol{\beta})$ an n-vector with typical element $x_t(\boldsymbol{\beta})$.

3.1.2 Estimation

Let $\mathbf{X}(\boldsymbol{\beta})$ be a $n \times h$ matrix of non-linear functions, where h is the number of functions. With this notation in place, the nonlinear least squares (NLS) estimator, can be defined as the parameter vector $\boldsymbol{\beta}$ that solves the following moment condition

$$\mathbf{X}^T(\boldsymbol{\beta})(\mathbf{y} - \mathbf{x}(\boldsymbol{\beta})) = \mathbf{0} \tag{31}$$

It can be shown that this condition is just the first-order condition with respect to the parameter vector from minimization of the sum-of-squared residuals (SSR) function, given as

$$SSR(\boldsymbol{\beta}) = \sum_{t=1}^{n}(y_t - x_t(\boldsymbol{\beta}))^2 = (\mathbf{y} - \mathbf{x}(\boldsymbol{\beta}))^T(\mathbf{y} - \mathbf{x}(\boldsymbol{\beta})) \tag{32}$$

This function closely resembles the one defining the OLS estimator. There are however, two important differences: 1) The matrix $\mathbf{X}(\boldsymbol{\beta})$ is a matrix of functions rather that explanatory variables and 2) the moment conditions are nonlinear functions of $\boldsymbol{\beta}$, as both $\mathbf{x}(\boldsymbol{\beta})$ and $\mathbf{X}(\boldsymbol{\beta})$ are also nonlinear functions of $\boldsymbol{\beta}$. The consequence of this is that there does not, in general, exist a closed-form solution[19], since changing $\boldsymbol{\beta}$ changes both $\mathbf{x}(\boldsymbol{\beta})$ and $\mathbf{X}(\boldsymbol{\beta})$. The usual solution to this problem, is to use numerical methods.

Another important concept in NLS-analysis is asymptotic identifiability. An estimator is said to be asymptotically identified[20] if the parameter vector $\boldsymbol{\beta}$ is the true parameter vector, the model is correctly specified and gives a unique solution to the minimization problem above. If this is the case, in addition to some more technical conditions[21], the NLS estimator will be both consistent and asymptotically normal (Davidson and MacKinnon, 2004, p. 225).

Define the covariance matrix of the estimators as Σ. Then, when the estimators are asymptotically identified, satisfying the assumptions above, and with IID errors, the estimate of this matrix becomes

$$\hat{\Sigma}(\boldsymbol{\beta}) = s^2(\mathbf{X}^T\mathbf{X}^T)^{-1} \tag{33}$$

where $\mathbf{X}_t$ is defined as above and s^2 is given by the sum of the estimated squared residuals:

$$s^2 = \frac{1}{n-k}\sum_{t=1}\hat{u}_t^2 = \frac{1}{n-k}\sum_{t=1}(y_t - x_t(\boldsymbol{\beta}))^2 \tag{34}$$

If $\hat{\boldsymbol{\beta}}$ is consistent, then u_t will converge towards the error terms as the sample size tends to infinity, and s^2 provides an approximately unbiased estimate of the true variance σ^2 (Davidson and MacKinnon, 2004, p. 227).

In the case with non-IID errors, such as heteroskedasticity, the true variances vary over t and the estimator above becomes inconsistent. There exists, in general, two ways of handling this problem (Cameron and Trivedi, 2005, p. 154-157). It can be done by using either a Feasible

---

[19]A closed-form solution is a solution to the minimization problem that is directly obtainable from the expression.

[20]Asymptotic identification requires, according to Davidson and MacKinnon (2004, p. 216) that "the estimation method provide a unique way to determine the parameter estimates in the limit as the sample size,n, tends to infinity".

[21]Good references on these details are is (Cameron and Trivedi, 2005), pages 152-155 for a general overview, and (Wooldridge, 2010) chapter 12 for a more rigorous treatment.

generalized least squares approach (FGNLS), or Weighted least squares (WNLS) (Cameron and Trivedi, 2005, p. 154). Further details on calculation of these estimates are omitted here due, but can be accessed in either Cameron and Trivedi (2005) pages 156-157 or Wooldridge (2010) pages 413-418, the latter providing the most detailed discussion.

3.2 Seemingly Unrelated Regressions (SUR)

The seemingly unrelated regressions estimator (Zellner, 1962), provides efficient estimates of a system of equations, that is *seemingly* unrelated in the sense that there is no simultaneity[22] in the system, but where the error terms are correlated across equations. Intuitively, the estimator assumes that the unobserved part of some processes, modeled as a system of equations, are correlated, and uses this information to obtain more efficient, that is, precise estimates of the parameters in the system. The exposition given here, will closely follow (Davidson and MacKinnon, 2004) page 501-507.

3.2.1 The SUR model

Suppose we have a system of g different equations, with k_i regressors, each observed n times, where i designates the ith of the g equations. Further, let $\mathbf{y}_i$ denote the $n \times 1$ vector of observations on the ith dependent variable, $\mathbf{X}_i$ the $n \times k_i$ matrix of k_i regressors on the same equation, $\boldsymbol{\beta}_i$ the corresponding parameter vector, and $\mathbf{u}_i$ the $n \times 1$ vector of errors terms. The ith equation is then given as

$$\mathbf{y}_i = \mathbf{X}_i \boldsymbol{\beta}_i + \mathbf{u}_i \tag{35}$$

where $\mathbf{I}_n$ is the identity matrix. with stochastic structure, given by

$$E(\mathbf{u}_i \mathbf{u}_i^T) = \sigma_{ii} \mathbf{I}_n \tag{36}$$

It is possible to obtain consistent estimates of this system by just applying OLS to each of the g equations. However, if it is the case that the errors term are correlated across equations, the assumptions above changes into the following assumptions on the stochastic structure of the model:

$$E(u_{si}u_{sj}) = \sigma_{ij}, \forall s, E(u_{si}u_{vj}) = 0, \forall s \neq v, \tag{37}$$

That is, the error terms are correlated across equations, but are homoskedastic and independent within any given equation. These covariances can be formed into an $g \times g$ matrix, often dubbed

---

[22]Simultaneity means, informally, that the value of all the dependent variables in the system are determined simultaneously, and that at least one of them is dependent on the value of one or more of the other endogenous variables in the system.

the contemporaneous covariance matrix. Arranging the error terms u_{sj} into an $n \times g$ matrix, $\mathbf{U}$, of which a typical element is the row $1 \times g$ vector $\mathbf{U}_s$, that is the sth observation of the g equations error terms, these matrices defines the contemporaneous covariance matrix, Σ, as

$$E(\mathbf{U}_s^T \mathbf{U}_s) = \Sigma \tag{38}$$

This matrix contains the variance of each equation's residual on its diagonal, and the coviariances between them on its off-diagonal elements, thus modeling the relatedness of the equation's unobserved parts by their error terms. Coinstraints on space does not permit a full exposition of this matrix written out, and the interested reader is referred to (Davidson and MacKinnon, 2004), pages 502-504. The model equation and the contemporaneous covariance matrix completes the description of the SUR model.

3.2.2   Estimation

Moving on to estimation of the system, there exists several possibilities, where the most common approach is the feasible generalized least squares estimator (FGLS), according to Davidson and MacKinnon (2004, p. 507), which I will shortly explain here. We first define the model in a more compact form by stacking each of the g equations on top of one another, thus forming the following matrices

$$\mathbf{y}_* = \begin{pmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \\ \vdots \\ \mathbf{y}_g \end{pmatrix}, \mathbf{X}_* = \begin{pmatrix} \mathbf{X}_1 & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{X}_2 & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{X}_g \end{pmatrix}, \mathbf{u}_* = \begin{pmatrix} \mathbf{u}_1 \\ \mathbf{u}_2 \\ \vdots \\ \mathbf{u}_g \end{pmatrix}, \boldsymbol{\beta}_* = \begin{pmatrix} \boldsymbol{\beta}_1 \\ \boldsymbol{\beta}_2 \\ \vdots \\ \boldsymbol{\beta}_g \end{pmatrix} \tag{39}$$

This enables one to write the model as a single equation using the matrices defined above

$$\mathbf{y}_* = \mathbf{X}_* \boldsymbol{\beta}_*^T + \mathbf{u}_* \tag{40}$$

where the parameter vector can be consistently, albeit ineffeciently, estimated by OLS. A more efficient estimator under our assumptions on the covariance matrix, and when this matrix is unknown, is the feasible generalized least squares estimator, which first estimates the covariance matrix, and uses this to produce more precise estimates as information of the dependence between the equations is utilized.

The estimation proceeds in two steps[23]. First, one estimates the system by OLS on each equation, which are used to produce estimates of the individual error terms, $\hat{u}_{ij}$. Then, the

---

[23]Some caveats on the discussion is warranted. First, when the form of the contemporaneous covariance matrix is known, it need not be estimated and, and one can use GLS (Davidson and MacKinnon, 2004) directly, see (Davidson and MacKinnon, 2004) page 503-504 for details. Moreover, two special cases of the SUR estimator

contemporaneous covariance matrix is estimated by its sample analouge, using the estimated OLS residuals

$$\frac{1}{n}(\hat{\mathbf{U}}_s^T \hat{\mathbf{U}}_s) = \hat{\Sigma} \tag{41}$$

When the OLS estimators on each equation are consistent, the residuals will be estimated consistently for the error terms, and so also the contemporaneous covariance matrix. Finally, the estimator is produced by minimizing a weighted sum of squares, which in matrix form can be written as the following minimization problem

$$\min_{\boldsymbol{\beta}}(\mathbf{y}_* - \mathbf{X}_*\boldsymbol{\beta}_*^T)^T(\hat{\Sigma}^{-1} \otimes \mathbf{I}_n)(\mathbf{y}_* - \mathbf{X}_*\boldsymbol{\beta}_*^T) \tag{42}$$

Taking derivatives with respect to the parameter vector, setting equal to zero, and solving for the parameter vector yields the following **generalized least squares estimator** (FGLS) of the system

$$\hat{\boldsymbol{\beta}}_{FGLS} = (\mathbf{X}_*^T(\hat{\Sigma}^{-1} \otimes \mathbf{I}_n)\mathbf{X}_*)^{-1}\mathbf{X}_*(\hat{\Sigma}^{-1} \otimes \mathbf{I}_n)\mathbf{y}_* \tag{43}$$

which is consistent and asymptotically normal under the stated assumptions[24]. Morover, the two-step procedure can be performed several times, by using the FGLS estimates to construct a new estimate of Σ, performing the procedure until the estimated parameter vector changes by little, i.e. until it converges. If it converges, the FGLS estimator converges to the maximum likelihood estimator ((Davidson and MacKinnon, 2004), pp. 511 & 516), a fact I will use later in this thesis.

It is customary to test whether or not the matrix Σ is diagonal; that is, whether the different equation's error terms are independent. This can be done by using the Lagrange multiplier test, proposed by (Breusch and Pagan, 1980):

$$Lm = N \sum_{i=2}^{m} \sum_{j=1}^{i-1} r_{ij}^2 \tag{44}$$

Here, N is the number of observations, and m is the number of equations. r_{ij}^2 is the correlation coefficient between equation i and j. This statistic will be distributed as χ^2, with $\frac{1}{2}m(m-1)$ degrees of freedom.

---

arises when either 1) The contemporaneous covariance matrix is diagonal, meaning that the equations are truly unrelated (independent) or 2) that each of the g regressor matrices are equal to one another. In both these cases, the SUR estimator (be it GLS or FGLS) collapses to the OLS estimator, as there are no gains to systems estimation in terms of efficiency. Moreover, the model can also be estimated by maximum likelihood, provided that the data is normally distributed.

[24]$\otimes$ denotes the Kronecker product.

4   Parameterizing the equation system

This part of the thesis deals with parameterization of the equation system developed from the condition for optimal flow and tolls. In the theoretical section, I developed the following equilibrium condition

$$Q^E = \{Q : v(\tau(Q) + GC(Q)) = Q\} \tag{45}$$

This condition says that at the equilibrium, the sum of tolls and generalized costs given by flow level Q, should induce a flow equal to that specific level. This equilibrium gives the flow where the marginal social benefit is equated to the marginal social cost, as given by the intersection at point D in figure 1. When this point is found, one can conduct a cost-benefit analysis of implementing congestion pricing by measuring the difference between the loss in consumer's surplus and the decrease in the external costs.

The equilibrium toll and flow is given as Q^E and $\tau(Q^E)$. To solve for these values, I need to find the equilibrium flow, to substitute into the toll equation. However, tolls and flow are dependent upon one another as explained in the theoretical section. Changing tolls changes flow, and generalized costs.

My solution to this problem is to model the process by parameterization of the demand equation. That is, I model the trial-and-error procedure by defining a three equation system. With one equation for the level of flow, $v(Q)$, one for tolls, $\tau(Q)$ as given in Johansson (1997) and generalized costs $g(Q)$. Together, this gives me the following system of equations

$$Q(\tau, GC) = v(\tau, GC), \tag{46}$$

$$\tau(Q) = g(Q), \tag{47}$$

$$GC(Q) = a(Q) \tag{48}$$

where $v(\tau, GC)$, $g(Q)$ and $a(Q)$ are parametrized equations. This formulation treats traffic volume as an endogenous variable who is to be determined by the optimal toll level equal to the external costs at demand Q, and the generalized cost (private cost) of a trip at the same level of traffic.

I now write the toll and generalized cost functions explicitly as given by the PCE flow: $\tau(Q_r)$ and $GC(Q_r)$. This means I also need to rewrite the demand function as dependent upon the PCE flow. The reason I do this, is the following: My data on flow does not differentiate between heavy and light vehicles. When setting a toll, it should be proportional to the congestion factor α of the vehicle. By using an adjustment factor outlined in appendix B, I convert the flow data into PCE flow data. Thus, I apply the toll equally to all counts of vehicle flow, as it is measured

in PCE to obtain a single toll. This means I will have to rewrite the demand function to be dependent upon PCE flow, and not the unadjusted flow.

The demand function, $v(.)$, should be able to predict changes to demand from changes to cost, as I use an iterative procedure trial-and-error procedure. This means I need a function in which demand at the next stage in the trial-and-error process depends on the demand in the last trial, the new toll, given an elasticity. Noting that one can insert the two bottom equations in the system into the first one to get a single equation enables me to formalize this,

$$Q_r = v(\tau(Q_r), GC(Q_r); \varepsilon_{PCE}) \tag{49}$$

Here, ε_{PCE} is the elasticity of demand measured as PCE flow with respect to generalized costs and tolls. That is, I define the system in terms of PCE units, and use an average elasticity of the PCE flow to find the equilibrium. Demand is now a function of the PCE flow, through $\tau(Q_r)$, $GC(Q_r)$ and the elasticity ε_{PCE}.

Now, I have one equation that is derived from the bi-level optimization problem from the theoretical section. Observe that solving this equation gives me a value in which the external costs are contained through the tolling equation, and the response of road users is incorporated through the demand equation. Solving for the Q_r that makes this equality hold, allows me to find the optimal flow, by taking into account the feedback effect of flow through changes in tolls and generalized costs. That is, as flow changes, so does tolls and generalized costs, which again affects flow.

The purpose of the next section is to parameterize the different equations of the system. That is, I want to find values parametric forms for $\tau(Q_r)$, $GC(Q_r)$, and ε_{PCE}. First, I show how the elasticities of unadjusted flow with respect to tolls and generalized costs are estimated. Second, I develop the parametric forms used to calculate the toll and generalized cost functions. Third, I show how I the PCE elasticity is estimated.

4.1   Traffic demand - ε

In this section, I will estimate elasticitets of annual average traffic volume at toll stations in Norway with respect to motoring costs and toll per trip for long and short vehicles. This means I estimate the percentage change in traffic volume, when the toll or generalized costs changes by one percent.

This will be important in determining the effect on traffic volume from introducing congestion pricing. That is, knowing the price sensitivity of different user groups. I formulate a model of traffic demand to obtain elasticities. First, I give a short survey on previous findings on

elasticities. The reason being that my data set has some important limitations, making it pertinent to have a solid comparison of my results with others.

4.1.1 Previous findings

Previous studies on the relationship between traffic levels generalized costs are numerous. Some of the most comprehensive studies are (Oum et al., 1992) and (Goodwin, 1992). The general conclusion from this literature promotes the fact that demand for travel by vehicle is relatively inelastic with regards to generalized costs in the short run, but increases in the long-run. Studies specifically investigating toll elasticities are fewer in number. A literature review by Odeck and Bråthen (2008) conclude that the typical value is -0.5 meaning that traffic volume decreases by -0.5 % when the generalized cost including tolls increases by 1%. Evidence from Spain put forth by Matas and Raymond (2003) estimated that short-run elasticities varied between -0.21 and -0.81, and varied also according to economic activity, fuel prices and travel conditions on parallel roads.

The evidence from Norway is generally concurring with the conclusions drawn in the international part of the literature. Odeck and Bråthen (2008), used a time-series model to estimate short-and long-run elasticities regarding income and generalized costs at specific toll projects, and calculated the elasticity for a number of toll projects to be either initiated or terminated. Their results indicate the following: a slightly higher elastic demand with an average value of the estimated elasticity, of -0.56 with respect to generalized costs from changes to tolls. Moreover, they found that elasticities vary according to the type of road, i.e rural roads, trunk roads and urban motorways, where the first proved the most elastic, and a higher variance on the latter type of roads. (Jones and Hervik, 1992) evaluated the effect of introducing road pricing in the Norwegian cities of Oslo and Ålesund, and found an estimated elasticity of -0.22 for the first and -0.45 for the second with respect to tolls.

The Norwegian Public Roads Administration have conducted a number of detailed studies on introduction and termination of toll projects. Odeck and Kjerkreit (2008) found an average short-run arc-elastcity between -0.67 and -0.52 with respect to generalized costs, including the toll. Moreover, their results indicated that elasticities are affected by the existence of alternative routes, lowering the elasticity as users can substitute one route for another. This contradicts my assumption regarding substitution between roads, made in the theoretical chapter. Other important findings, which are in line with Toftegaard (2007), are that the elasticity decreases with income and, that the purpose and length of travel also affects the estimates, where longer trips are generally viewed to be more elastic. Both studies found that leisure trips are the

most elastic, commuting trips covering the middle ground, while work-related ones are the least elastic. The reason might be that work-trips are in part, or fully, paid for by the employer, and further, that commuting trips are linked to labor supply, which is not easily adjustable in the short run. Toftegaard also noted that high-frequency travelers (such as commuting ones) had a lower elasticity, which supports this. Both studies reported that elasticties were lower when the trip was partly or fully financed by someone else. Lastly, the average elasticity reported by Toftegaard (2007) was -0.69 with respect to generalized costs including toll, very close to the estimate of Odeck and Kjerkreit (2008).

The general conclusion from this short survey, is that toll road elasticity is a complex phenomenon, which varies according to many different variables. As such, calculating exact elasticities will be difficult, as there are many latent variables that are not observed in my data set, which I do not observe.

4.1.2 Method

To investigate how demand for traffic changes, I will estimate elasticities with respect to tolls and generalized costs. These are defined as the percentage change in demand for traffic, after a 1 % change in either of the mentioned cost components.

As indicated in the literature review, elasticites typically vary with purpose of travel, income, whether or not the driver pays the toll himself, and so on. In my data set, there are many toll points from different areas, and such, I suspect there is some heterogeneity in my sample. Thus, a lack of a temporal dimension may bias my estimates. Local conditions I do not observe could be controlled for using a time-series set, where such variables plausibly might be kept constant. Moreover, as I have a cross-section set, I will only be able to estimate short-run elasticities.

In short, the data set at hand is not a ideal one for estimation of elasticities, but I hope my results will be indicative of the level. A main purpose of the estimation, is thus to find a function that based on observable characteristics of the toll points, enables me to find an estimated elasticity that reflects the local conditions to control for as much of the heterogeneity as possible.

I define a function, called the **elasticity function**, as the derivative of the log of daily traffic flow for vehicle group j at the toll point, with respect to the log of the total toll[25]

$$\epsilon(\mathbf{w}_{jV})_j = \frac{\partial \ln AADT_j(\mathbf{x}_{jV})}{\ln \partial \tau_j} \tag{50}$$

It shows how the elasticity of *daily* traffic volume changes according to a set of variables included in $\mathbf{w}_{jV}$. This set of variables is a subset of the ones in $\mathbf{x}_{jV}$, which enters the main equation for

---

[25]This definitions gives what is known as the arc-elasticity. In doing so, I follow (Odeck and Bråthen, 2008).

$\ln Q$. That is, the variables that affect the level of the elasticity, also affect the traffic volume directly, but not all the variables that affect the traffic volume directly, affects the elasticity. Estimation of the function is by first estimating total demand, and then taking its derivative.

As noted, estimating elasticities on my data will give a *daily* elasticity, whereas, I intend to calculate the change in demand over the day. As the composition of traffic changes during the day, with respect to the purpose of travel, traffic in different periods might have different elasticities. Thus, when trying to predict the change in demand, using daily data might introduce a bias, as the elasticity of daily traffic might not be the same as for any given period within the day. In fact, one can show that the elasticity for daily traffic is a weighted average of the period-wise elasticity with respect to tolls throughout the day[26]:

$$\varepsilon = \sum_{t=1}^{T} a_t \varepsilon_t \tag{51}$$

where t is one of the T periods during the day, a ε_t the elasticity with respect to tolls in that period, and $a_t \in [0, 1]$ the share of daily traffic that is present at the road in period t. This means that periods with a high proportion of demand, and a high daily elasticity, will influence the elasticity of daily traffic volume more than periods with low elasticity. Thus, I will not be able to identify if elasticities vary over the course of the day, only the average.

Model: The statistical model is represented by the following system of equations, where $AADT_{LV}$ designates log annual average daily traffic for light vehicles, and $AADT_{HV}$ the corresponding variable for heavy vehicles.

$$\ln AADT_{LV} = \zeta_{0SV} + \boldsymbol{\pi}_{LV}\mathbf{x}_{LV} + \epsilon_{LV}, \tag{52}$$

$$\ln AADT_{HV} = \zeta_{0HV} + \boldsymbol{\pi}_{HV}\mathbf{x}_{HV} + \epsilon_{HV} \tag{53}$$

where $\mathbf{x}_{LV}$ is a vector regressors in the equation for light vehicles, and $\mathbf{x}_{HV}$ a corresponding vector for heavy vehicles, ζ is a constant. The vector of errors terms , $\boldsymbol{\epsilon} = (\epsilon_{LV}, \epsilon_{HV})$ of errors terms are assumed to be independent and normally distributed.

$$\boldsymbol{\epsilon} \sim NID(\mathbf{0}, \boldsymbol{\Sigma}) \tag{54}$$

The contemporaneous covariance matrix is equal to

$$\boldsymbol{\Sigma} = \begin{pmatrix} \sigma_{LL}^2 & \sigma_{HL} \\ \sigma_{LH} & \sigma_{HH}^2 \end{pmatrix} \tag{55}$$

where σ_{LL}^2 is the variance of the error term in the light-vehicle equation, σ_{HH}^2 the corresponding one for heavy vehicles, and $\sigma_{HL}^2 = \sigma_{LH}^2$ the covariance between the error terms of the equations for light and heavy vehicles.

---

[26]See appendix *A* for details.

Estimation   The main purpose of the traffic demand model is to obtain unbiased and consistent estimates of the toll elasticities. Given my limited sample size, this means I will have to give special attention to efficiency, that is, which estimator that exploits the maximum amount of information from my relatively small sample size of $N = 170$. I take two steps to achieve this. The first, lies in the choice of estimator, the second the choice of stochastic model for the residuals.

It is known that the maximum likelihood estimator is the most efficient estimator in small samples (Cameron and Trivedi, 2005, p. 143), under the assumption that the distribution of the data corresponds to the density used in the likelihood function. Doing a logarithmic transformation of AADT traffic volume, I tested, whether my data was not normally distributed by using the test for normality developed by (D'agostino et al., 1990), and could not reject, that my data is not normally distributed. As a result, I will use the maximum likelihood estimator. Details on the calculations are in the appendix.

My data contains observations on traffic volumes for heavy and light vehicles. As such, there might be unobserved factors that affect both types of traffic volumes simultaneously. Examples are quality of roads, differences in preferences and availability of alternative modes of transportation (for example, proximity to airports), etc. Even though these factors are not observed, they can be modeled by estimating a covariance matrix where the unobserved parts (the errors) of both traffic volumes are correlated. One way to implement this, and which have gained recent interest in the literature (Jou and Chen, 2014), is to use the SUR estimator (Zellner, 1962), which was introduced in section 3.2. This will extract more information in the sample, as the shared information in the two equations is exploited in addition to what is contained in each equation separately. Note that this does not control for unobserved factors' effect on my estimates, but rather their precision. This means that the unknown part of the processes that determine traffic volume are modeled simultaneously in the two equations. Moreover, I use a bootstrap procedure to account for possible heteroskedasticity, which can be done by using the bootstrap procedure implemented in $STATA13^{TM}$ command for SUR regression, according to Cameron and Trivedi (2010, p. 166). By clustering the bootstrap procedure at county level, I also obtain estimates robust to autocorrelation within any given county, as several toll points are located in the same one. Atkinson and Wilson (1992) investigated the performance of the SUR bootstrap estimator in small samples, and found ambiguous evidence related to its superiority as compared to the standard FGLS procedure. In the appendix, I provide regression results for the FGLS procedure of the most important models. They give smaller standard errors, such that using the bootstrap procedure should in any case give the most conservative estimate, which is also recommended

by Angrist and Pischke (2008, p. 307).

Details on the specification   Moving on to a discussion of the included regressors, the literature survey indicated that there are many factors influencing the sensitivity of demand with respect to price. My empirical specification of the demand function will try to incorporate as many of these as possible. Firstly, Odeck and Bråthen (2008) suggests that a log-log specification of the demand function, yielding a convex form is the most common in transport, which I will follow[27]. Moreover, using this form, the estimates will be directly interpretable as constant elasticities.

Starting with demand for light vehicles, several studies indicate that elasticities vary according to type of road, and the level of income of the drivers. To incorporate this into my model, I will interact the median, net income at county level with the level of tolls, and include a dummy for European Road (EV) and National Road (RV), which are interacted with the level of tolls[28]. For heavy vehicles, there is less empirical evidence to inform the specification. I will thus formulate a somewhat less extensive function, just incorporating the level of toll, interaction between toll and road type, and county-level income. The county-specific level of income and generalized cost for both traffic volumes are also added for both heavy and light vehicles. To account for the effect of alternative routes, enabling road users to substitute, I use a variable called *alternative routes*, which is defined as total road length in kilometers, divided by the square kilometers of area, classified as business area, at county level. As toll points are located in counties with widely different population levels, and more populous regions may have higher median income, it might bias the interaction term between income and toll. The reason begin the effect of population being included in the income variable, as regions with higher population could have higher traffic demand, irrespective of income. To adjust for this, I also include variable measuring the log of county population level. Further, I assume that generalized costs and tolls are treated differently by motorists. Adopting this approach follows Matas and Raymond (2003).

4.2   Congestion tolls - $\tau(Q_r)$

In this section, I parameterize the toll function to find the marginal external costs associated with a given level of flow. First, I develop a function for time costs second, environmental costs. The two costs are combined to provide the final estimate of total marginal cost. I use the current traffic flow to give a numerical illustration of the method, and highlight some pitfalls.

---

[27]I have also tested different specification, see appendix.

[28]I also tested whether or not main effects on EV and RV should be included, and found that they could be excluded. See appendix.

The method assumes traffic is in a normal flow state characterized by no traffic jams. All formulas are developed as functions of the flow, rather that current speed at the road link. The reason being that I will use the expressions when solving for the equilibrium value of tolls, which need to take into account the flow. This follows the approaches taken by TØI (2014), Mayeres et al. (1996) and Sen et al. (2010), estimated from data using a single link, rather than range of different links.

4.2.1 Method

In this part, I will give a description of the methodological approach I use to estimate the marginal external costs. The purpose is to estimate the optimal congestion charge from Johansson (1997), where I have left the term MC_k^e out. This is the direct effect of a vehicle using the road on the environment. It is assumed that this does not make an important difference to the results, and it will simplify the computational procedure significantly. This means I deviate somewhat from the theoretical model used earlier, as only the indirect effect on emissions through lowered speed of other vehicles is included. I will thus get a lower estimate of the external costs, and thus tolls. However, as mentioned earlier, there are other costs than environmental and time that could be included in the analysis. Excluding one more means I obtain a lower bound estimate on the overall change in social benefits, which is the main purpose of this thesis.

I will now develop the estimation of each of the terms.

Time costs: The aim is to estimate the term from Johansson (1997) (10), which are the costs associated with time lost in traffic. The term shows how the marginal social time costs of adding a vehicle to the road is related to the level of flow, and is given as:

$$\alpha_k \frac{\partial V}{\partial Q_r} \frac{\partial T_i}{\partial V} \sum_{i=1}^{n} Q_i VOT_i \tag{56}$$

where the term $\frac{\partial T_i}{\partial V_r}$ has been moved outside of the parenthesis. Here, α_k is the congestion factor for group k. $\frac{\partial V}{\partial Q_r}$ is the partial derivative of speed with respect to flow. $\frac{\partial T_i}{\partial V}$ is the partial derivative of time needed to travel one kilometer, with respect to speed. The last term gives the value of time as the sum over all user groups.

I will now discuss the practical implications of this expression, and the parametric forms chosen to represent the theoretical models. Starting with the last part of the expression, I have

two user groups, and thus define the following term[29]

$$\sum_{i=1}^{n} Q_i VOT_i = Q_r[\pi VOT_H + (1-\pi)VOT_L] \tag{57}$$

where π is the share of heavy vehicles at the road, implying that $1 - \pi$ is the share of light ones, as there is only two classes of vehicles.

Here, the total traffic value of time is the flow weighted average of time costs. This is based on the value of time for light vehicles, calculated as the average value of the different VOTs for different travel purposes. The numbers are taken from TØI (2010), giving $VOT_L = 187.5 NOK/hr$. VOT_H is set at $380 NOK/hr$, taken from the same reference.

For the first part of (57) ($\frac{\partial V}{\partial Q_r} \frac{\partial T_i}{\partial Q_k}$), I use a direct mapping between traffic flow and the time needed to traverse 1 km, such that the first factors are simplified into one function, given as

$$\frac{\partial V}{\partial Q_r} \frac{\partial T_i}{\partial V} = T'(Q_r) \tag{58}$$

I now need to specify the time-flow function's derivative, $T'(Q_r)$. To achieve this, I will specify the time-flow relationship on the road link. This is the approach I found to be most common in the literature. I follow the approach taken to by Mayeres et al. (1996) and Sen et al. (2010), and estimate an own functional relationship, albeit a different one from theirs.

The time-flow curve which I want to estimate can be seen in figure 7. The curve depicts how flow and the travel time per kilometer are related. As more vehicles enter the road, meaning higher flow, the time needed to travel the link increases with higher flow up to the maximum flow. At this point, the curve becomes backward bending. Here, the effect on speed from adding another vehicle to the road, has a larger effect on the flow, than adding an extra vehicle. Flow drops, as traffic goes towards a jam-like state, such that fewer vehicles are able to pass through the road link at each time point. As a consequence, speed is lowered, and the time needed to travel one kilometer increases. Thus, the relationship between flow and speed is no longer invertible, as two different values of time corresponds to the same value of flow.

To circumvent this problem, I only estimate the part of the time-flow curve which is not backward bending. As shown later, I will in either case not be able to estimate tolls for traffic on the backward-bending part due to data limitations.

To proceed in finding the lower part of the flow-time curve, I need to estimate the point at which the curve becomes backward bending to find the maximum flow, and thus, the lower part of the curve. This is the same as estimating the maximum flow of the road link, and then finding which time that is associated with it this value of flow. Maximum flow is the point t_{cap}

---

[29]I use Q_r as a substitute for Q_i in the sum to have simpler computations. It may slightly overestimate the costs time costs.

Figure 4: The time-flow relationship, adopted from Button (2004, p.6) with some minor alterations.
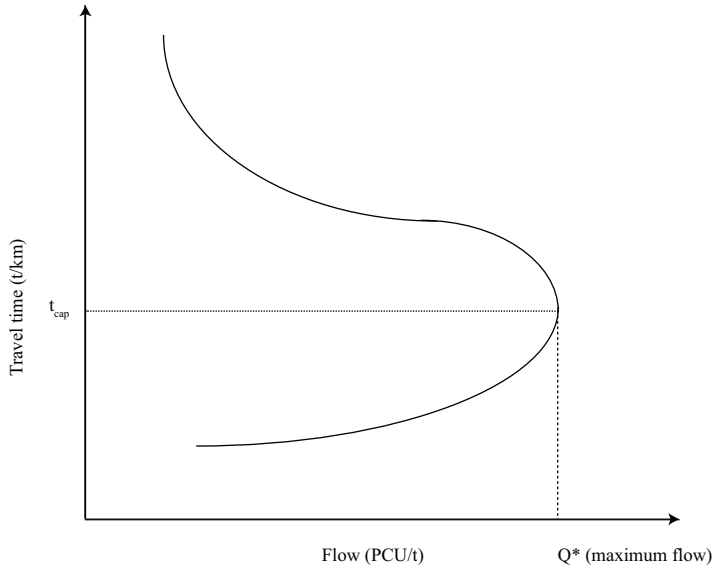
in figure 7. I estimated this point by inverting the parabola and finding the maximum of the inverse function where flow is given by time. That is, the maximum of $Q(T)$. Using this method, i estimated maximum flow to be $t_{cap} = 0.07$.

As mentioned earlier, I need to specify a function to describe the time flow relationship in (58). I assume that the function is convex, and that there exists a free flow time, defined as the travel time on the link when flow is zero. The convexity reflects that there is an increasing marginal social cost related to flow. Moreover, in (10), the toll is directly proportional to the congestion factor for vehicle group k. Thus, my function needs to be such that its derivative with respect to flow in group k, is directly proportional to the derivative with respect to the group with $\alpha = 1$, by a factor α_k. For example, marginal cost of time is given as $MC_i^t(V(Q_r))$ in Johansson (1997), where $V(Q_r)$ is speed as a function of passenger car equivalent flow. When differencing this, I should have $\alpha_k MC'(V)V'(Q_r)$ for the k-th group, where $MC'(V)V'(Q_r)$ is the marginal cost for vehicle group in which $\alpha = 1$.

I chose the following form of the function to reflect these criteria:

$$T_i = \frac{1}{s_t} = \beta_0(1 + e^{\beta_1 Q_r}) \tag{59}$$

where s_t is the speed recorded at time t at the road link, and Q_r is the passenger car equivalent flow. β_0 is there half the travel time at zero flow, and β_1 the congestive impact on time from one PCE unit. When the proportionality criteria is satisfied, the proposed congestion function would enable one to identify the tolls for different user groups by just scaling the function with

congestion factor, as the only difference between the congestive impact of group k and passenger cars is α_k. I omit the derivation here, and provide them in the appendix.

The implication of the scaling property is that the the marginal time cost of group k can be written as

$$MC^t = \alpha_k \beta_0 \beta_1 e^{\beta_1 Q_r} Q_r [\pi VOT_H + (1-\pi) VOT_L] \tag{60}$$

where $T'(Q_r)$ is substituted by the derivative of the congestion function with respect to PCE flow, and scaled by the congestion factor for the group the cost shall be calculated. This is the final expression used to estimate time costs on the lower part of the time-flow curve. The flow-time model is fitted by feasible non-linear least squares, as given in the econometric section. Regression results are provided in the appendix for reference. I also assumed there are on average 1.68 persons per passenger car. This means the final cost will be scaled by 1.68. The constant is taken froM (TØI, 2011a).

Environmental costs: The purpose of this part is to obtain a practical estimate of the following formula taken from the solution to the optimal toll in Johansson (1997) (10):

$$\alpha_k \sum_{i=1}^{n} Q_i \frac{\partial MC_i^e}{\partial V} \frac{\partial V}{\partial Q_r} \tag{61}$$

This shows the marginal social cost of increased traffic flow of group k on the road for pollutant j, as the sum of marginal costs for each user group, scaled by the congestion factor for vehicle group k. In my thesis, I include three different pollutants: CO_2, PM_{10} and NO_X, some of which have an effect globally, others locally. In Johansson (1997), there is not several pollutants included. The indexing is added by me. I will now parameterize each of the terms of this expression separately. I start with the term $\frac{\partial MC_i^e}{\partial V} \frac{\partial V}{\partial Q_r}$. This is the effect on marginal environmental social costs for one single vehicle with respect to flow. The first factor gives the relationship between speed and marginal costs, and the second factor the effect on speed from increased flow. Let $g(V(Q_r))$ the emission factor, defined as grams per kilometer of pollutant j, and $S(Q_r)$ the speed at the link as a function of passenger car flow, Q_r. I rewrite the expression, $\frac{\partial MC_i^e}{\partial V} \frac{\partial V}{\partial Q_r}$, as:

$$\frac{\partial MC_i^e}{\partial V} \frac{\partial V}{\partial Q_r} = \frac{d}{dQ_r} \Big[g_i(S(Q_r)) VOE_j \Big] = g_i'(S(Q_r)) S'(Q_r) VOE_j \tag{62}$$

where VOE_j is the social cost of pollutant j, measured as NOK/gram. Now, the sum in the (61) needs to be taken into consideration to find the total marginal cost. That is, the expression above, gives the increase in social cost of emission of a single vehicle who has its speed affected by an increase in the flow from one extra vehicle entering the road. But the extra vehicle on the road is assumed to affect all vehicles on the road, not just one, and all groups. This needs to be taken into consideration.

To achieve this, I assume that the proportion of PCE units on the road from light and heavy vehicles is a constant share of the total PCE flow, such that $Q_L = (1-\pi)Q_r$ and $Q_H = \pi Q_r$, where π is the share of the PCE flow coming from heavy vehicles. Inserting this into the expression and taking the sum yields

$$\alpha_k VOE_j[g'_H(S(Q_r))S'(Q_r)Q_r\pi + g'_L(S(Q_r))S'(Q_r)(1-\pi)Q_r] \tag{63}$$

The common terms, VOE_j, Q_r and $S'(Q_r)$, can now be taken outside, which gives

$$MC^e_{tj}(Q_t) = \alpha_k VOE_j S'(Q_r)Q_r\Big[g'_H(S(Q_r))\pi + g'_L(S(Q_r))(1-\pi)\Big] \tag{64}$$

Here, the marginal cost is given by the effect through a change in average speed on the emissions levels of all other vehicles. The term is measured in NOK/km. In this thesis, the vehicle's direct social environmental cost (MC^e_k) is assumed to be zero, that is the direct environmental social cost of flow. This will understate the true externalities, and consequently the congestion price. Thus, in estimation of changes to welfare, I will get a conservative estimate. Introducing the term increases the complexity of the numerical procedure. As I have already left several potential external costs out of the analysis (as mentioned at the beginning of section 2), I proceed by assuming leaving additional terms out leaves me with a lower bound on the estimated social benefit.

I will now specify the parametric form of the equations $g(V(Q_r))$ and $S(Q_r)$. Starting with the function relating speed to flow, $S(Q_r)$, I choose an exponential form to account for the concavity of the speed-flow relationship, that is, the marginal affect on the speed on of the road link, increases as flow increases. I have: $S(Q_r) = \nu_0[1 + \exp(\nu_1 Q)]$, where ν_0, ν_1 are parameters, the former being the free flow speed.

When specifying the emission factor function, $g(V(Q_r))$, it is important to account for the differential impact on emissions from different types of vehicles. I want to estimate two different functions, one for heavy vehicles, and one for light ones. Within these two groups, there are several sub-groups of vehicles called EURO-classes (TØI, 2011b). Each EURO class corresponds to a certain relationship between speed and emissions.

Using data from the NPRA on the relative share of the vehicle classes at the road link I am investigating, I find an average relationship between emissions and speed. The data relates the same speed in all EURO classes to different emission factors. I use a weighted average for the emission factors within each EURO class. This is done separately for light and heavy vehicles. The equation that is estimated is then the following one

$$g_{kj}(S) = \sum_{i=1}^{h} w_i g_i = \gamma_0 S^{\gamma_1} \tag{65}$$

where $w_i = PKM_i / \sum_{i=1}^{h} PKM_i$, and S is speed. PKM_i is the number of vehicles in any given EURO class multiplied by the average driving length per year for that class. The dependent variable is thus a weighted average of emissions based on the relative size of the average driven distance. The equation is estimated by non-linear least squares and results are given in the appendix.

Using the specified functions, the time and environmental costs are parametrized by appropriately deriving the empirical functions. That is, using $g_i'(S(Q_r)) = \gamma_1 \gamma_0 S^{\gamma_1 - 1}$ and $S'(Q_r) = \nu_1 \nu_0 \exp(\nu_1 Q)$, this is inserted into MC_t^e above to find

$$MC_{tj}^e(Q_t) = \alpha_k VOE_j \nu_1 \nu_0 \exp(\nu_1 Q_r) Q_r \left[\gamma_{1H,j} \gamma_{0H} S(Q_r)^{\gamma_{1H,j}-1} \pi + \gamma_{1L,j} \gamma_{0L,j} S(Q_r)^{\gamma_{1L,j}-1} (1-\pi) \right]$$

(66)

This expression is the final parametric form of external cost associated with pollutant j. The two functions are estimated by feasible non-linear least squares, and results are reported in the appendix.

4.2.2   PCE congestion toll

The total marginal cost per kilometer is found by summing over the J pollutants, and adding the marginal time costs, that is

$$MC_t(Q_{rt}) = MC_t^t + \sum_{j=1}^{J} MC_{tj}^e(Q_{rt})$$

(67)

This expression is then used to model $\tau(Q_r)$ from (Johansson, 1997), as the toll should be set equal to the marginal external costs. The time subscript t is introduced here to signal to the estimate belongs to specific time period.

4.3   Average PCE elasticity - ε_{PCE}

The estimated tolls are set according to one PCE unit, whereas my elasticities are set according to counts of each vehicle group. This creates the need for conversion of the estimated elasticities into a format that is applicable to the PCE data. In this section, I show how what I call the average PCE elasticity may be used for such a purpose. To start, let us first investigate how the traffic PCE flow is affected by the proposed pricing scheme. Let $Q_r = \alpha Q_H + Q_L$ designate the total PCE flow. Differencing this term with respect to tolls for light and heavy vehicles gives the total reaction to traffic flow

$$dQ_r = \frac{\partial Q_H}{\partial \tau_H} \times \alpha d\tau_L + \frac{\partial Q_L}{\partial \tau_L} \times d\tau_L$$

(68)

where $d\tau_L$ is the change in tolls for light vehicles corresponding to $\alpha = 1$. A finding in the last section was that when estimating tolls per PCE unit, the toll for vehicles having a PCE value larger than one is found by multiplying the toll for one PCE unit $\tau(Q_r)$, by the congestion factor α_k to get $\tau_k = \alpha_k \tau(Q_r)$ for group k. This means that the toll for heavy vehicles is α times larger than for light ones, implying that $d\tau_H = \alpha d\tau_L$. This shows that the pricing change is consistent with the expressions derived earlier. To proceed, move the common factor $d\tau_L$ outside a parenthesis, and multiply the term inside the parenthesis by $\frac{Q_k}{Q_k}\frac{\tau_k}{\tau_k}$, where $k \in (Light, Heavy)$ to get:

$$dQ_r = \left[\frac{\partial Q_H}{\partial \tau_H}\frac{Q_H}{Q_H}\frac{\tau_H}{\tau_H}\alpha + \frac{\partial Q_L}{\partial \tau_L}\frac{Q_L}{Q_L}\frac{\tau_L}{\tau_L}\right]d\tau_L \tag{69}$$

Notice this enables one to rewrite the equation in form of elasticities. Moreover, as $\tau_H = \alpha\tau_L$, the α cancels in the first term in the brackets. Taken together, the term becomes

$$dQ_r = \left[\alpha\varepsilon_H \frac{Q_H}{\alpha\tau_L} + \varepsilon_L \frac{Q_L}{\tau_L}\right]d\tau_L = \left[\varepsilon_H \frac{Q_H}{\tau_L} + \varepsilon_L \frac{Q_L}{\tau_L}\right]d\tau_L \tag{70}$$

Here, ε_k is the elasticity of *traffic flow*, not PCE flow, of demand in vehicle group k with respect to tolls. To obtain the elasticity of the PCE flow, I need to replace the unadjusted flow by the PCE flow.

To proceed, I first assume that the proportion of light and heavy vehicles of total unadjusted traffic flow Q is constant. Then, one could define: $Q_H = \pi Q$ $Q_L = (1 - \pi)Q$, where π is the share of total traffic volume comprised of heavy vehicles. Inserting these relations into (70) gives

$$dQ_r = \left[\varepsilon_H\pi\frac{Q}{\tau_L} + \varepsilon_L(1 - \pi)\frac{Q}{\tau_L}\right]d\tau_L \tag{71}$$

Next, the PCE flow is constructed by adjusting the observed flow by the following formula: $\frac{1-(1-\alpha)\pi}{\lambda}Q = Q_r$, where Q is the unadjusted flow as defined in section 6, and Q_r the PCE flow. This implies that the unadjusted flow is $Q = \frac{\lambda}{1-(1-\alpha)\pi}Q_r$. Denote $\frac{\lambda}{1-(1-\alpha)\pi}$ by Λ. Inserting $\Lambda Q_r = Q$ into (71), and moving the common term $\frac{Q_r}{\tau_L}$ gives the following expression:

$$dQ_r = \Lambda\frac{Q_r}{\tau_L}\left[\varepsilon_H\pi + \varepsilon_L(1 - \pi)\right]d\tau_L \tag{72}$$

Dividing by Q_r, $d\tau_L$ and multiplying τ_L on both sides gives the average PCE elasticity:

$$\frac{dQ_r}{d\tau_L}\frac{\tau_L}{Q_r} = \varepsilon = \Lambda\left[\varepsilon_H\pi + \varepsilon_L(1 - \pi)\right] \tag{73}$$

where ε is the average elasticity per PCE, π is the share of heavy vehicles, ε_H the elasticity of heavy vehicles, and ε_L the corresponding one for light vehicles. Λ converts the average elasticity of the undadjusted flow into the average elastictiy of the PCE flow. Λ is the inverse of the adjustment factor converting normal flow into PCE flow while simultaneously accounting for missing data counts. See appendix B for details. Notice that $d\tau_L$ is the toll per PCE, and is equal to the toll for light vehicles.

5 Solving the equation system

In this section I will solve the equation system proposed earlier. Solving a system of non-linear equations is in general, difficult. I will use a numerical method to solve for the equilibrium value, rather that solving the system algebraically. The ease of implementation and avoidance of algebraic errors are the main benefits of this approach. (De Palma and Lindsey, 2004) uses a numerical method similar to mine for computation of congestion prices in a general equilibrium setting, where budget constraints, public goods, income taxation and equity preference are taken into consideration. My solution will be much simpler, as I do not consider all of the mentioned aspects, but still similar intuitively.

5.0.1 Fixed-point iteration

As pointed out above, I want to solve for a level Q^E such that $Q^E = v(\tau(Q^E), GC(Q^E))$. This is the same as saying Q^E is fixed point of the function $f(Q)$. A fixed point of a function f is defined as follows: c is a fixed point for the function f iff $f(c) = c$, according to (Wood, 1999, p. 145). Clearly, when $Q^E = v(Q^E)$ through $\tau(Q), GC(Q)$, Q^E is a fixed point on the demand function $v(Q)$.

Assuming such a fixed point exists, one could find it by using what is known as *fixed point iteration.* Having an initial condition x_0, first-order fixed point iteration is defined as generating a sequence $\{x_n\}$ such that (Wood, 1999, p.145):

1. x_{n+1} is obtained from x_n

2. $x_n \to \alpha$ as $n \to \infty$, where α is fixed point of f

To generate this sequence, one uses a recurrence relation $x_{n+1} = f(x_n)$. If the limiting value of this function is α, then α is fixed point of $f(x)$, according to (Wood, 1999, p. 144).

Translating this into the notation used in this thesis, I have a function $Q_{n+1} = v(\tau(Q_n), GC(Q_n))$, where substitution for tolls and generalized cost is made to generate a sequence Q_n to see if there is a limiting value of Q, corresponding to the equilibrium. This means I estimate the traffic flow equilibrium by the limiting value of such a sequence starting from an initial condition Q_0, that is

$$\hat{Q}^E = \lim_{n \to \infty} v(\tau(Q_n), GC(Q_n)) \tag{74}$$

This process is used to model the trial-and-error solution proposed in the literature by (Li, 2002).

5.0.2 The iteration process

The structure of the iteration process is set up such that at first a toll is estimated, which constitutes optimizing the F function in the bi-level problem. Then the reaction of flow from the introduction of the toll is predicted, which is the new solution to the f-problem. A new estimate of the generalized costs is predicted, subsequently measuring the change to demand. The new traffic level then requires a new toll to be estimated, to optimize F, and the process starts all over again until convergence is reached.

To find the solution, it remains to specify the functional forms to be used to model the equation system. The function that is used to model demand is the following one:

$$Q = v(Q) = Q[1 + h(Q)] \tag{75}$$

where the *correction function* $h(Q)$ is given as

$$h(Q) = \varepsilon_{t,\tau}\left(\frac{\tau_t(Q_k)}{\tau_t(Q_{k-1})} - 1\right) + \varepsilon_{t,GC}\left(\frac{GC_t(Q_k)}{GC_t(Q_{k-1})} - 1\right) \tag{76}$$

The function $h(Q)$ is used to correct towards the fixed point (equilibrium) for different starting values. My econometric models indicated that a constant-elasticity specification is the most suitable one, regarding the toll elasticities, see appendix F. Assuming this also applies to generalized costs, and for all periods of the day, the formulation of $h(Q)$ gives the percentage change in demand for the total percentage change in tolls and generalized costs, in line with the constant elasticity formulation. I assume the elasticities are different for tolls and other costs, as indicated by my traffic demand model. Moreover, I estimated the elasticity of motoring costs in my econometric section. Since I do not have a model of how motoring costs changes with different levels of flow, I will use time costs instead. This naturally places some limitations on my results, as they do not incorporate all the relevant elements to generalized costs. Further, I assume the coefficient for motoring and time costs are the same, such that I am able to use it in the equilibrium modeling.

The equation for τ and GC are collected from the simultaneous equation system developed in the theoretical section. For the generalized costs, I make a simplification, and only use the time costs. This gives me the following equations:

$$\tau(Q) = \tau = MC_t(Q) + \sum_{j=1}^{J} MC_j^e(Q), \tag{77}$$

$$GC(Q) = \beta_0(1 + e^{\beta_1 Q})VOT \tag{78}$$

The fixed-point iteration requires a starting value. This value determines where the iterative process starts from. From different starting values, one should obtain different convergence

points, as different starting values corresponds to a single point on each period's demand function, and each convergence point should correspond to different equilibrium points. To find this value, I start by defining the initial demand, Q_{0t} as

$$Q_{0t} = Q(GC, \tau; \theta_t) \tag{79}$$

This is the mean flow at present in a given time period t, as given by present marginal costs and tolls, given a set of preferences θ_t. The constant is estimated by taking the average of flow in period t using the flow-time data. Thus, I do not observe this demand function, only a *single* point on it; the one intersecting with the private cost curve.

The initial demand function represents a part of the *initial state* of the equation system. In this state, the toll is $\tau_0 = 31$ for all values of Q, as there are no congestion tolls. When the government introduces congestion pricing, they do so by introducing a toll equal to marginal cost of the initial demand, Q_0. The initial value of which the iteration starts, is then found by combining the initial toll and demand in $h(Q)$ to give

$$x_0 \equiv Q_0[1 + h(Q_0)] = Q_0 \left[1 + \varepsilon_{t,\tau} \left(\frac{\tau_t(Q_0)}{31} - 1 \right) \right] \tag{80}$$

which is just the percentage change in demand after a percentage change in price, given by the term inside the parenthesis. As Q_0 and the initial toll is determined from data, they are determined outside the model, and provide the exogenous starting point from which the iterative process commences. The time costs are assumed to be constant at the starting point.

Now, using standard fixed-point iteration, the following recurrence relation is exploited to find the point at which $h(Q) = 0$, such that $Q_{k+1} = Q_k$ and the fixed point (equilibrium) is found

$$x_{n+1} = f(x_n) \equiv Q_{k+1} = Q_k[1 + h(Q_k)] \tag{81}$$

where the equations in the correction function are given as

$$\tau(Q_{k+1}) \quad = \quad = MC_t^t(Q_k) + MC_{CO_2} + MC_{NO_X} + MC_{PM_{10}}, \tag{82}$$

$$GC(Q)_{k+1} \quad = \quad \beta_0(1 + e^{\beta_1 Q_k})VOT \tag{83}$$

To find the flow for multiple periods, I rely on the assumption that flows are independent between different periods (assumption 1), to solve for the optimal level of flow in each period $t = 1, ..., T$. This means I estimate a fixed point for each of the $T = 288$ different time periods.

6   Data

Here, I present the data used in this thesis. I use two different datasets. The first is a cross-sectional dataset that contains traffic Average Anual Daily Traffic Volumes (AADT) observed at different toll points within Norway[30], and the second, a time series of traffic volume and speed at a specific intersection in Oslo[31]. The first data set is to be used for estimation of elasticities, whereas the second for congestion externalities. A more detailed description of these data sets will now follow. First, I give a description of sources and construction of relevant variables, before proceeding to study some general charcteristics of the most important variables.

6.1   Cross-section data set

The purpose of this data set is to estimate elasticitets related to generalized costs for short and long vehicles. The main dependent variable is AADT collected at different toll points within Norway, and is an estimate of the average traffic volume per day on a specific road link. This is supplied with information on the level of toll for large and small vehicles, the share of large vehicles at the toll point, road classification[32], geographical location (longtitude and lattitude) and year of data registration. All of these variables are supplied by the Norwegian Public Roads Administration database Nasjonal Vegdatabank (National Road Data Bank). This is coupled with several county level variables from Statistics Norway. These are: the county population[33], median gross income and deduction corrected tax in NOK[34], the number of registered vehicles according to fuel type used[35], the area devoted to corporate buildings[36], the total kilometers of road[37] and data on fuel prices per liter[38].

From these data, the median net income at county level is calculated as gross income minus the equated tax. Generalized travel costs are constructed by multiplying the average trip length within a county by the average cost of travel per kilometre. The latter is constructed by taking the weighted average of driving costs by fuel and diesel driven cars. The weights are constructed by

---

[30]Data set obtained from the NPRA's database *Nasjonal Vegdatabank* (National Road Database)

[31]Data set obtained from NPRA's database*Reisetidsprosjektet* (Travel Time Project).

[32]In Norway, roads are classified according to: Kommunevei (county road) fylkesvei (District road), Riksvei (National road) and Europavei (European Network Road).

[33]Statistics Norway; Statistikkbanken; Table: 03027.

[34]Statistics Norway; Statistikkbanken; Table: 05671. Equated tax is defined as the total income and wealth tax paid to county, district and state, including pension payments. Gross income is defined as labor, business, capital and pension income

[35]Statistics Norway; Statistikkbanken; Table: 07849.

[36]Statistics Norway; Statistikkbanken: Table: 09594.

[37]Statistics Norway; Statistikkbanken; Table: 04694.

[38]Statistics Norway; Statistikkbanken: Table: 09654.

defining the relative proportion of diesel and petrol driven cars within a county. The generalized cost can the be expressed mathematically as

$$GC_j = TL_i \left(\sum c_{jk} + \alpha FP + (1 - \alpha)DP \right) \tag{84}$$

where GC_j designates generalized travel costs for vehicle type j (heavy and light), as average trip length within county i, multiplied by the sum of costs components c_k from SVV (2014). α is the share of petrol driven vehicles, FP the fuel price per kilometre, and DP the diesel price per kilometre. All the unit costs per kilometre are taken from (SVV, 2014). The average trip length is gathered from the (TØI, 2011a). The variable "alternative routes" is defined as the total square kilometers of corporate building mass, divided by the total square kilometers of road. Below, a table with summary statistics for the main variables is seen.

Table 2:

Summary statistics of cross-section data set.

Variable	Obs	Mean	Std. Dev.	Min	Max
Net income	170	284597.1	16715.89	244800	323100
Average trip length	170	17.32046	10.78572	4.736275	72.59545
Log toll Heavy vehicles	170	3.924981	.7051776	2.302585	6.49224
Log Motoring costs Heavy vehicles	170	12.89815	1.623705	9.458718	15.7422
Alternative routes	154	237.1647	470.8976	1.805869	4050
Log net income	170	12.55713	.0584409	12.4082	12.68572
Log population	170	10.92233	1.526188	6.887553	13.36053
Log Toll - Light vehicles	170	3.05578	.6263452	1.609438	5.010635
Log Motoring costs - Light vehicles	170	3.589747	.4838368	2.422715	5.149173

Net income after tax is somewhat compressed with a standard deviation of approximately 16 000 NOK. The average driving length within a given county has a mean value of about 17.3 kilometers. The tail of this distribution is quite thick, as one can see by the high maximum trip length, meaning that the typical trip is shorter (the median is 14.6 kilometers). Moreover, the number of inhabitants by county shows that toll points are situated in counties with widely different population levels. Lastly, one can also note that the toll for light vehicles is on average markedly lower than for heavy ones.

6.2   Volume-Delay data set

The second data set used in this thesis is for calculation of external costs. The data is gathered from the NPRA's[39] data base "Reisetidsprosjektet"[40] and contains 5-minute interval recordings of average speed and flow[41] for a number of toll projects in Norway. My data set is recorded at the road link between Sandvika and Lysaker Vest toll points in the Oslo Toll Ring between January the 1st and February the 27th, 2015. Using the data on speed and length of the link, I have calculated the average travel time within each of the 5-minute intervals by the following formula: $T_t = \frac{Length/1000}{Speed_t}$, where $length$ is the length of the road link in meters and $speed_t$ is the average speed in time interval t.

Flow is defined as the number of vehicles that passes a specific point during a specific amount of time. Following Maerivoet and De Moor (2005), let Q_t be the flow, then

$$Q_t = \frac{1}{T_{mp}} \sum_{i=1}^{L} N_{i,t} \tag{85}$$

where L is the number of lanes in the road link, and $N_{i,t}$ is the number of vehicles in lane i at time t, and T_{mp} the time interval used to record the traffic volume in hours. I use 5-minute intervals, and measure flow per fifth minute. The data is, however gathered for the lanes in direction towards Oslo, meaning I only study traffic going into the city. Heavy data management is the reason why I only consider one link, as considering both would be out of the scope of this thesis.

The variable was adjusted for two factors. First, the flow variable does not acknowledge that larger vehicles may have a stronger impact on congestion than smaller ones. That is, the flow is not measured in numbers of PCE units. Second, data collection is based upon cars having an electronic registration device. This means the actual number of PCE units is higher than what is estimated. Under the assumption of a fixed proportion of heavy vehicles of 15 % throughout the day[42], light and heavy vehicles having the same proportion of registration devices with $\lambda = .82$ taken from (Meld.St.17(2008-2009):), and an assumed congestion factor $\alpha = 1.5$, the following adjustment factor is used to calculate the PCE flow[43], :

$$Q_r = \frac{[1 - (1 - \alpha)\pi]Q_t^{OC}}{\lambda} \tag{86}$$

where Q_r is the flow in terms of passenger car equivalents, as defined in the theoretical chapter, Q_t^{OC} is the observed flow count from the raw data, λ is the share of cars having an electronic

---

[39]The Norwegian Public Roads Administration.

[40]English: "The Travel Time Project".

[41]See definition below.

[42]Taken from the cross-section data set.

[43]The factor is derived in appendix B.

registration device, and α being the congestion factor. The adjustment factor can be shown to equal inverse of the heavy vehicle adjustment factor (TRB, 1994) used in highway engineering, divided by λ. Moreover, α was chosen on the basis of the ranges of estimated congestion factors reported in (Benekohal and Zhao, 2000) and (Kimber et al., 1985). Using these formulas, the average travel time on the road link was 4.2 minutes, with a speed limit of 80 km/h, and a length of 4.6 kilometers. Average speed was 73.1 km/h, and the average flow of cars was 34.8 cars/5th-minute.

7   Results

In this section, the results of the various estimations are given. First, I present the results from the estimated traffic demand models. Second, I present the estimated present marginal external costs. Third, I estimate the optimal flow and tolls.

7.1   Elasticities - ε

The results of the estimations are given in the table below. A discussion on these results will now follow, where I highlight the most important findings[44].

---

[44]The number of observations are lower in some of the estimations as some of the variables have missing values. Moreover some observations are excluded, typically toll stations by by tunnels in areas with low population and on islands.

Table 3:
Estimation result from annual average daily traffic demand model. The first column shows
the base model without any effect on income. The three proceeding columns shows models
where income affects the elasticity, and various other variables are included. In column three,
the alternative routes variable is included, and in column four, the log of county population
is included. Model with AR, is the model where alternative routes are included as a regressor
without including county population level.

	(1) Model 1		(2) Model 2		(3) Model with AR		(4) Model 3	
Ln $AADT_L$								
Ln τ_L	-0.543***	(-3.60)	-40.17**	(-2.95)	-40.01**	(-2.97)	-33.24*	(-2.21)
Ln $\tau_L \times EV$	0.402***	(8.62)	0.388***	(8.71)	0.390***	(10.24)	0.384***	(8.87)
Ln $\tau_L \times RV$	0.220***	(3.57)	0.219***	(3.77)	0.239***	(3.90)	0.223***	(3.99)
Ln GC_L	-0.376***	(-5.63)	-0.338***	(-5.30)	-0.320***	(-5.53)	-0.316**	(-3.16)
Ln $Income$	4.367**	(3.09)	-5.636	(-1.48)	-6.395	(-1.81)	-6.506	(-1.79)
Ln $Income \times \tau_L$			3.155**	(2.92)	3.142**	(2.93)	2.605*	(2.18)
Alternative routes					-0.000507***	(-3.40)	-0.000401**	(-2.86)
Log $Population$							0.147*	(2.23)
Constant	-43.44*	(-2.43)	82.09	(1.71)	91.63*	(2.07)	91.31*	(2.00)
Ln $AADT_H$								
Ln τ_H	-0.375**	(-2.97)	-0.448***	(-3.35)	-0.451***	(-4.05)	-0.417***	(-3.76)
Ln $\tau_H \times RV$	0.221***	(3.80)	0.224***	(4.07)	0.230***	(4.12)	0.217***	(4.46)
Ln $\tau_H \times EV$	0.410***	(10.89)	0.404***	(10.76)	0.400***	(12.25)	0.394***	(11.11)
Ln GC_H	-0.0920*	(-2.28)	-0.0739	(-1.61)	-0.0571	(-1.37)	-0.0496	(-0.54)
Ln $Income$	3.904**	(2.69)	3.537**	(2.59)	2.610*	(2.19)	0.790	(0.56)
Alternative routes					-0.000495**	(-3.23)	-0.000388*	(-2.48)
Ln $Population$							0.138	(1.41)
Constant	-40.40*	(-2.21)	-35.73*	(-2.07)	-24.18	(-1.62)	-3.072	(-0.18)
N	156		156		141		141	

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

The most important feature of the models, are their ability to predict reasonable estimates
on elasticities, that is providing an unbiased and consistent estimate of how these quantities
vary at different toll points. Since the models are formulated such that elastiticies vary, to find
the estimated elasticities, they need to be constructed by taking the derivative of the demand
function with respect to tolls. As such, I will not get a single estimate for the elasticity, by one
that varies over type of road and county level income, as these variables were interacted with
the toll level.

Due to the interaction terms, I am able to define the elasticities as functions of dummies for
county and european road, along with income, by taking derivatives of the demand functions
with respect to the toll level, this gives me three different models of the elasticities

Model 1: This model only incorporates the simple relationship between type of road and
elasticities.

$$\epsilon_1(\mathbf{x}_{SV})_j = \alpha_0 + \beta_1 RV + \beta_2 EV \tag{87}$$

47

Model 2: This model incorporates the simple relationship between type of road and elasticities, and models dependence on county median net income on the elasticities, plus including the alternative routes variable in the main equation.

$$\epsilon_2(\mathbf{x}_{SV})_j = \alpha_0 + \beta_1 RV + \beta_2 EV + \beta_3 Income \tag{88}$$

Model 3: This model incorporates the simple relationship between type of road and elasticities, models dependence on county median net income on the elasticities, and the population size and alternative routes are included as regressors in the main equation.

$$\epsilon_3(\mathbf{x}_{SV})_j = \alpha_0 + \beta_1 RV + \beta_2 EV + \beta_3 \ln Income \tag{89}$$

In all models, EV is a dummy for european roads and RV for national roads and $\ln Income$ being the log median county level net income. I use these models to predict the elasticities for all toll points in my sample. The graphs below show kernel-density estimates[45], and histograms of the distribution of the predicted elasticities. It is important to stress that the even though standard errors of the individual coefficients used to model the elasticities indicates significance, it does not mean that the predicted elasticities are significantly different from zero. To investigate this, I would have to find the variance of the elasticity function $Var(\epsilon)$, which would involve heavy matrix computations, which is out of scope for this thesis. I proceed by assuming the predicted elasticities are significantly different from zero. I will now move on to discussing the model results.

The first panel, shows how the elasticities are distributed using model 1. For light vehicles, the estimates seem reasonable, and are all within the scope of the estimates found in the literature review on elasticities. For heavy vehicles, the model predicts positive elasticities for a part of the sample. This is a problem, since positive elasticities for seem at odds with the findings in the literature.

---

[45] A kernel-density estimate is similar to a continuous histogram.

(a) Light vehicles.   (b) Heavy vehicles.
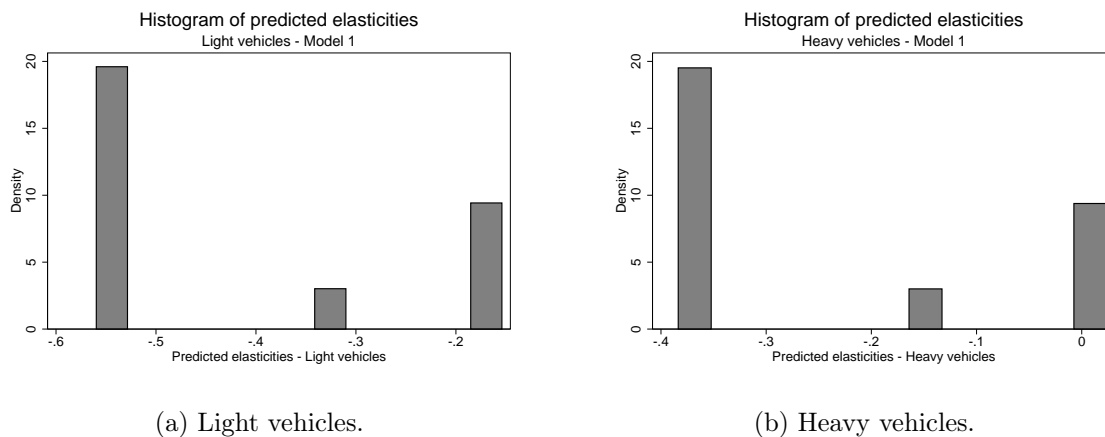
Figure 5: Model 1: Histogram of predicted elasticities.

(a) Light vehicles.   (b) Heavy vehicles.

Figure 6: Model 2: Kernel density estimate of predicted elasticities.

The second panel, shows the kernel-density estimate of the distribution of predicted elasticities for model 2. This model does not predict any positive elasticities for heavy vehicles, but does so for light ones. In particular for the specific road link I consider in this thesis, the predicted elasticity for light vehicles is .08 and -.013 for heavy vehicles. The link is situated just at the border between Oslo and Bærum county, the latter having a markedly higher income than the other counties in my sample. Thus, it might be that the model predicts badly at this point, since I have very few observations in the income range for Bærum.

(a) Light vehicles.   (b) Heavy vehicles.
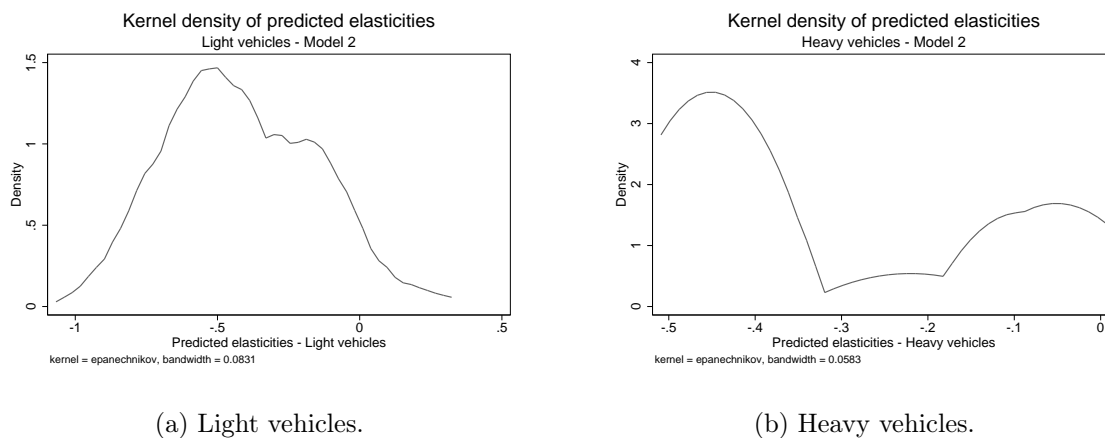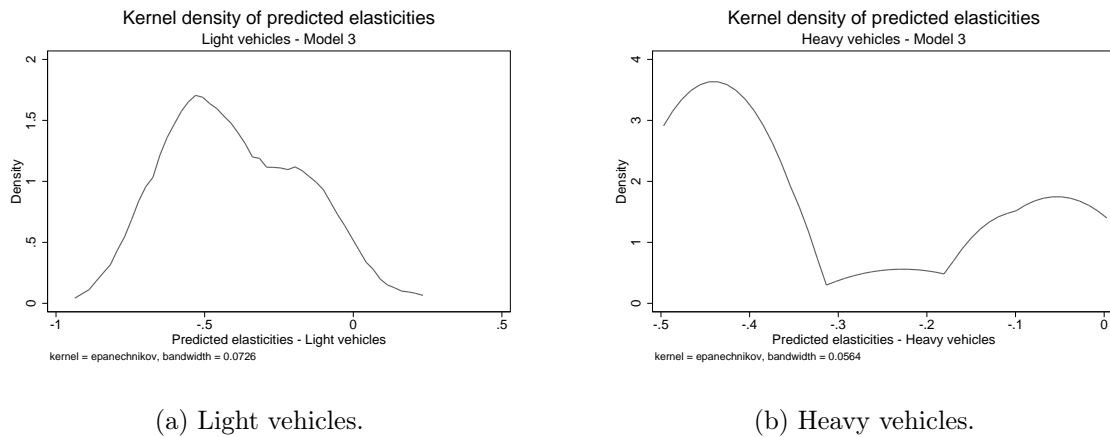
Figure 7: Model 3: Kernel density estimate of predicted elasticities.

The third panel, shows the kernel-density estimate of the distribution of predicted elasticities for model 3. The same pattern as seen in model 2, is evident here. The predicted elasticity for light vehicles is .03 and -.02 for heavy vehicles, which is very similar. Only 1 % of the predicted elasticities are positive, and these all correspond to the link I consider. However, toll points within Oslo county give more reasonable predictions, with an estimate of -0.051 for light vehicles and -0.051 for heavy ones, using model 2 and -0.0633 for light and -0.053 for heavy ones, using model 3. The nearest toll next toll point on my link which is situated on the same road, is 3 km.

The short distance between these two stations, and the special features of the income level in Bærum county, suggests that I might be able to get a more reasonable estimate of the elasticity by using the levels predicted for the rightmost Oslo stations. Moreover, both the income dependent models give very similar results for the other toll station. These estimates are still very low, as compared to previous findings both in Norway and internationally. One possible reason, is that I lack temporal variation in my data. This might bias my results. Thus, in the final analysis, I will compute the social surplus using a range of elasticities, with the estimates from my own models as a point of departure.

Model 1 predicted elasticities for short vehicles at -0.15, while the other models predict an elasticity at around -.05. Estimations conducted by Jones and Hervik (1992), indicated an elasticity around -0.22. Combining this evidence, suggest I should primarily use a range of -0.05 to -.22 when estimating the change in traffic volume. Moreover, even though model 2 and 3 indicated that the elasticity of light and heavy vehicles is almost identical, this might not be the case, as heavy vehicle might be used for work-related trips more frequently than light ones.

As the models have provided somewhat different answers to the question of the magnitude of the elasticities, I will conduct a sensitivity analysis with respect to elasticities when estimating the change in the social surplus. To get some bearing on the appropriate values to use, I

conduct a simulation experiment where the overall PCE elasticity is simulated on the basis of my estimates, and the findings in the literature. The simulated elasticities form the basis of the sensitivity analysis and can be viewed in appendix I,

7.2   Present marginal costs $\tau(Q_0)$

I will now present the estimated marginal social costs of traffic at the present. When choosing a statistic to describe the central tendency of the external costs, it is important to recognize that the average marginal external cost, is not the same as the external cost of the average traffic flow, due to Jensen's inequality[46]

Combining the two marginal costs, using the estimator above, the marginal costs (and thus the congestion price) is estimated for each five-minute interval, using a time series of 42 days with 288 observations on traffic volume per day.

The figures below show the estimated price per kilometer for different assumption regarding the congestion factor, α. That is, for different values of α, I have reestimated the marginal cost, MC_t^1, to assess how they affect the end price. As can be seen, the costs rise sharply during the morning hours, before falling again at approximately 8:00 AM, before rising once again. One can also note that the costs are somewhat dependent on the assumptions regarding the congestion factor. The right panel shows a plot with the estimated marginal cost for both light and heavy vehicles, with $\alpha = 1.5$. At all points, the costs for heavy vehicles lies above the light one's, and increases the higher the flow is.

---

[46]Jensen's inequality says: if f is convex, then $f\left(\sum_{\forall i} a_i x_i\right) \leq \sum_{\forall i} a_i f(x_i), \sum a_i = 1$. Note that the total marginal cost function, MC_{dt} is convex, implying that

$$MC_t^1 = \frac{1}{D} \sum_{\forall d} MC_{dt}(Q_{tT}^{OC}) \geq MC_t^2 = MC_{dt}\left(\frac{1}{D} \sum_{\forall d} Q_{tT}^{OC}\right) \tag{90}$$

where the predicted marginal cost for day d in period t is MC_{dt}, using a sample of D days. In this thesis, I have relied on the average marginal cost.

Note: Time of day indexed as; 1 = 00:00 - 288 = 23:55; Five-minute interval.
Assuming constant equal elasticities wrt. tolls in all periods
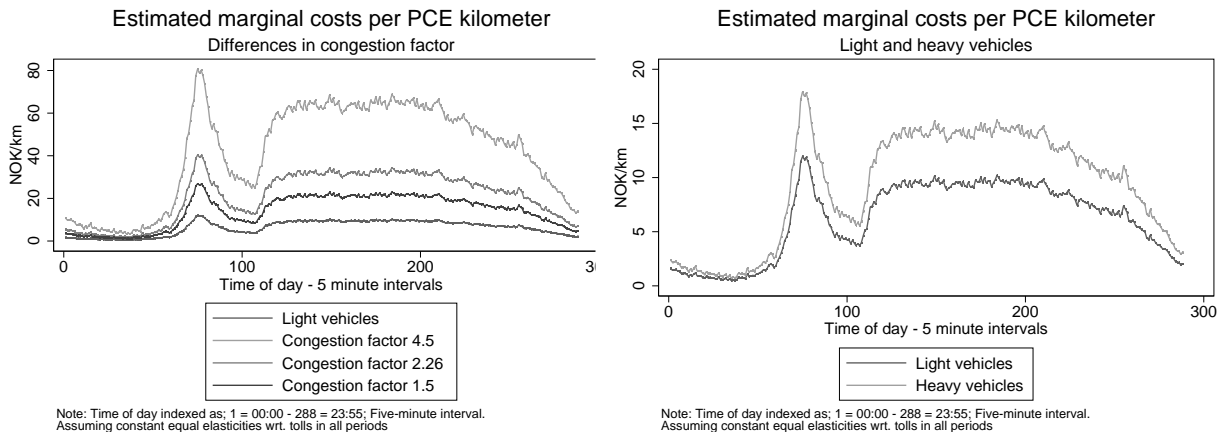
(a) Average marginal cost/km (NOK).   (b) Average marginal cost/km (NOK).

Figure 8: Average marginal cost (time and environmental) per PCE per kilometer for all T periods of the day. The drop around 8 AM suggests traffic is in a jam-like state, and underestimates the true costs.

The sharp fall in costs at around 8 AM, seems rather counterintuitive, as this is the period of the morning rush hour. One possible explanation is that there exists one or more equilibria in traffic jams, as illustrated in the theoretical section. There it was noted that an equilibrium on the backward bending curve where jams occur, makes the time-flow function $T(Q_r)$ non-invertible. The consequence being that costs are estimated as much lower than what is really the case. As the traffic is typically the highest during 8-9.30 AM, the above results might suggest there exists an equilibrium on the backward bending curve, such that the estimates of the rush-hour costs are too low. As pointed put in the theoretical section, this means I will be unable to estimate marginal costs or congestion tolls using the time flow relationship.

7.3 Iteration

The results of the equilibrium estimation on tolls and traffic volume is presented in this section. Starting with the equilibrium traffic level, the estimated equilibrium tolls is shown below, assuming an elasticity of -.091 for tolls and -.23 for generalized costs. The left panel shows the estimated toll per kilometer per PCE (henceforth called PCE kilometers), and the right panel the total toll. The latter is the toll that is actually levied on stations, assuming that drivers travel 12.6 km on average, to convert the per kilometer toll into a total one.

The same pattern emerges as for the present marginal costs. Tolls drop around 8 AM, which is the rush-hour. This suggests that only results not in this period can be considered valid. The current toll is 31 NOK for passenger car. Thus assuming a passenger car is one PCE, this is the

toll per PCE. Looking at the toll (panel (b), figure 9), one can see that almost all periods of the day have an estimated optimal toll higher than the current one, except during the night where the optimal toll is lower.
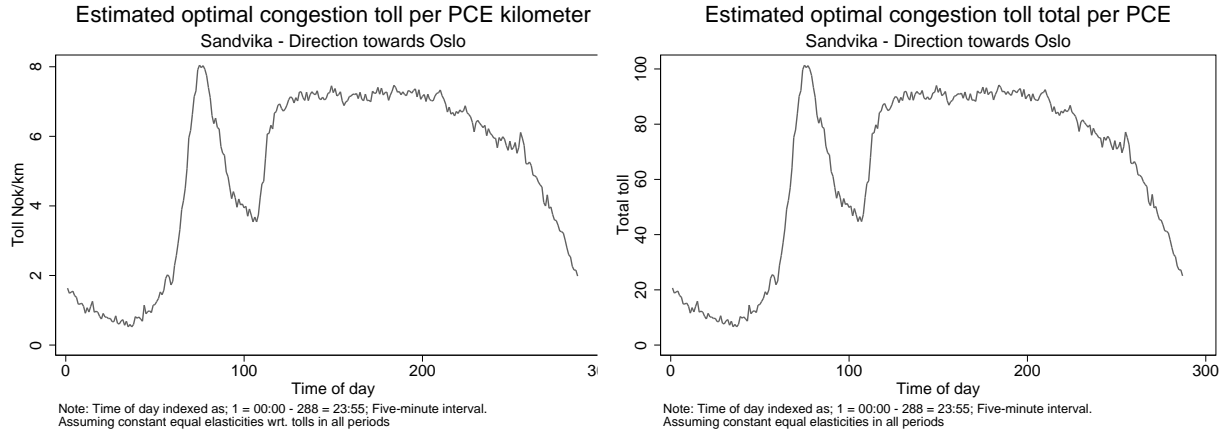
(a) Tolls per kilometer (NOK).

(b) Total tolls (NOK).

Figure 9: Estimated optimal tolls on link from Sandvika to Lysaker towards Oslo.

Moving on to the effect of flow, figure 10 (a) shows the difference between the present and optimal flow. Once again, the pattern of no change around the rush hour is seen. That is, the model suggests that the number of PCE is unchanged during rush-hours which seems counterintuitive. Once more, the existence of equilibria in traffic jams, is the likely cause of this phenomenon, and one should be cautious when interpreting these results.

Moreover, for the periods not affected by the existence of equilibria in traffic jams, one can see that optimal traffic flow is lower than the present one for most of the day. Only at night there is little difference. For periods with high flow, the relative change is higher as compared to periods with lower flow.

Figure 10 (b) shows the difference between the present marginal costs and the estimated optimal congestion tolls, the latter reflecting the marginal external costs in the optimum. The same pattern as regarding flow emerges, where marginal costs in the optimum are lower than at the present. Moreover, none of the periods have an estimated toll lower than zero. This is in line with the expected outcome as sketched in the theoretical section. There, it was pointed out that there will still be some congestion, and hence, social costs will be higher than just the private cost.
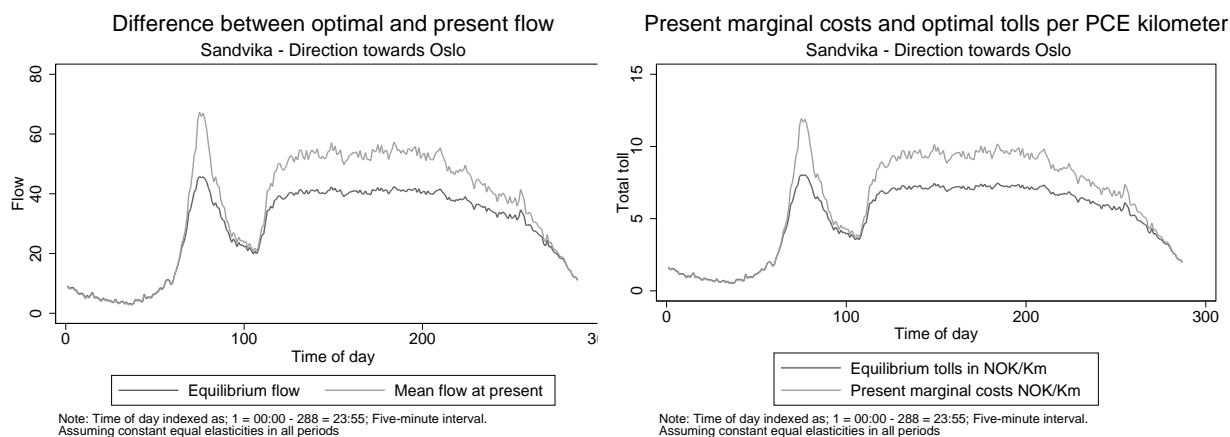
(a) Flow on link from Sandvika to Lysaker towards Oslo.

(b) Toll and present marginal costs on link from Sandvika to Lysaker towards Oslo.

Figure 10: Estimated optimal traffic flows and tolls on link from Sandvika to Lysaker towards Oslo.

8 Estimating the change in social surplus

In this section, I estimate the change in social surplus from introduction of congestion prices. Social surplus is assumed to be comprised of three components, consumer's surplus, social cost and operator surplus:

$$SS = CS + OS + SC \tag{91}$$

I start by giving an overview of the theoretical aspects and practical problems of estimating changes to consumer's surplus. Second, I explain how the consumer's and operators surplus is estimated along with social costs. I evaluate the change in each of these quantities separately before combining them to find the change to social surplus for a single period. I then show how I estimated the change to social surplus for all periods. Last, I present my results.

8.1 Discussion on the consumer's surplus

A standard measure of change in welfare is the consumer surplus, which measures the difference between the consumers' willingness to pay, and what they actually pay. This is defined formally as (Varian, 1992, p.163)

$$CS = \int_0^p Q(p)dp \tag{92}$$

where $Q(p)$ is the uncompensated demand curve. Even though it is highly popular in practice, there are some major problems associated with it. First, it is known from welfare theory that

it will be an accurate measure of welfare if, and only if, the agents utility functions are quasi-linear in income (Varian, 1992, p.163), with rational agents maximizing demand according to limited resources available (Slesnick, 1998). For a single price change, the quasi-linearity is vital as it ensures constant marginal utility of income, such that the consumer's surplus is an exact measure of welfare change (ibid.). If this is not the case, the marginal utility of income will change as prices change. Using the uncompensated demand curve will not take into account how the changes in the marginal utility of income affects overall welfare, as it is assumed to be constant. Moreover, when aggregating over many consumers, one must also assume that all consumers have the same marginal utility of income (Varian, 1992, p.169)

In the case of congestion pricing, there is not just a single price change, as prices are changed over different periods of the day. Assuming that some motorists travel more than once a day[47],there are even more restrictive conditions that need to be in place. First, income elasticities of different goods must be all be equal to one. Second, preferences need to be homothetic (Silberberg, 1972). As noted in (Slesnick, 1998), these requirements are particular challenges to applied welfare analysis. Should any of these requirements not be fulfilled, the consumer's surplus will not be an exact measure of welfare change (Varian, 1992, p.169). (Willig, 1976), however, shows that the consumers surplus can be regarded as an approximation to the true welfare change.However, this only applies to a single price change, and is not generalizable to several price changes (Slesnick, 1998). It is, however, the method used by the road authorities in Norway, in particular the National Public Roads Administration, see (SVV, 2014).

Moving the discussion towards the application of consumer's surplus in traffic analysis, some comments are in order. First, there are different ways of defining demand. Flow is one way, and the number of trips another. Generally, flow have been used by engineers to design road capacity at, and is such concerned with a single link. Trip counts is another way, and they are often paired with information on origin and destination, defined as, e.g. T_{ij} denoting the of trips from location i to j. Neuburger (1971) shows that the consumer's surplus should be calculated appropriately using such origin-destination pairs; that is, information on where a trip originates and ends, and the number for each pair of origin and destinations. Eliasson (2009) shows how using link flows, as I do, introduce an error, unless the optimal route choice for between an origin and destination is not affected by introduction of the congestion scheme. This is the same as saying there are no cross-elasticities between different routes.

I will make the simplifying assumption that consumer's surplus is a appropriate method for evaluating the welfare effects of introducing congestion pricing, as any other estimation method

---

[47]Such that they consume road travel in two separate periods, having different prices, that are changed differently compared to the situation with a uniform toll.

(for a survey, see (Slesnick, 1998)) is out of scope for this thesis.

8.2 Method

I will now explain how the social surplus is estimated, starting with the theoretical analysis of changes to its components, before moving on the the practical estimation. I will divide the affected groups into four, and analyze them separately. The analysis is undertaken at group level, where Q_r is the PCE flow, that is the number of PCE units that pass through the road link. The discussion is partly based on Immers and Stada (2004).

The standard textbook welfare analysis of congestion pricing, usually assumes there are no tolls levied prior to introduction of congestion pricing, see e.g. (Button, 2010, p. 22). In the Oslo toll ring there is already a toll of 31 NOK for passenger cars. Assuming a passenger car has a congestion factor of 1, the current toll per PCE is 31. Let p_0 denote total trip cost prior to the introduction of congestion prices, and p_1 the equivalent after introduction. Further, let τ_1 be the optimal toll and τ_0 the present toll, equal to 31 NOK. Last, let $GC(Q_1)$ be the generalized time costs in the optimal flow situation, and $GC(Q_0)$ in the current one. Now define the trip cost as $p_j = \tau_j + GC(Q_j), j = 0, 1$, Let us now investigate the welfare effects for four different groups: 1) Motorists remaining on the road 2) Motorists leaving the road 3) The operator (government) 4) Rest of society.

The motorists who stay on the road, enjoy a benefit of decreased time costs. This is the rectangle $CEFD$ in figure 11. However, they now pay τ_1 instead of τ_0, so they pay the rectangle $AEFB$. Their net increase in payment is what they pay minus what they gain, which is $-ACDB$. The motorists who leave the road, suffer a loss equal to the triangle BDG plus the rectangle DQ_1Q_0G. This is the groups total willingness to pay. At the same time, they are exempt from paying travel costs equal to DQ_1Q_0G, as they no longer use the road. Their net change to welfare is then the triangle $-BDG$. The government increases its revenue equal to the rectangle $ACDB$, as it already levied a toll τ_0. The rest of society accrues the reduction in deadweight loss, which equals the triangle BGF. In total, the net effect is $NET = -ACDB + ACDB + (-BDG) + BGF = BGF - BDG$. This is the net welfare loss of motorists leaving the road, and the reduction in deadweight loss. The motorists who remain on the road, do not change their behavior, so their reduction in consumer's surplus is cancelled by the increase in government revenue. This is just a transfer, and has no real effects.
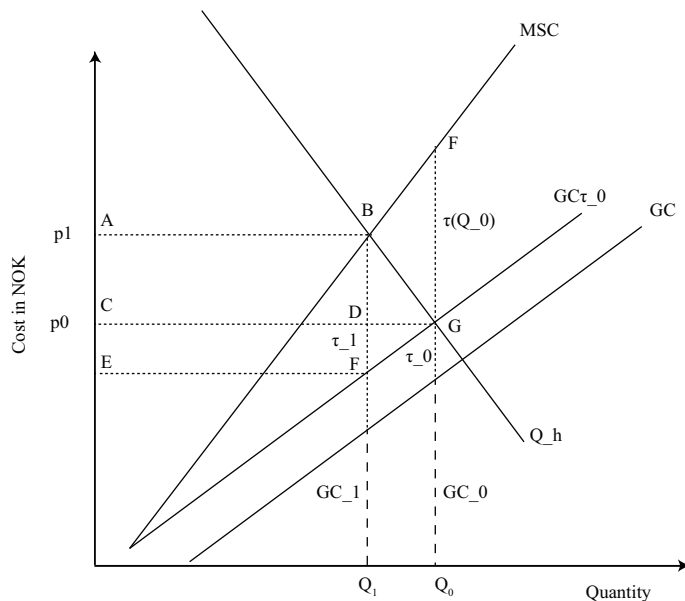
Figure 11: Effects of introducing congestion pricing in the Oslo cordon. Depiction of a single time period with demand Q_h. τ_0 is the current toll in the cordon ring, while τ_1 is the congestion toll. GC_1 is the generalized costs under the optimal flow, whereas GC_0 is the corresponding one under the current flow. GC_{τ_0} is the private cost curve (generalized cost of a trip) with the current toll, and GC without the toll.

I will now make a simplifying assumption. Although it is arguably incorrect, it will only make my estimation of welfare change more conservative, and reduces the need for calculations significantly. I will exclude the change in generalized costs when estimating the change in consumer's surplus, that is, $GC(Q_0) = GC(Q_1)$. This will give me a lower bound on the loss of motorists. Intuitively, by removing some benefits from the equation (saved time costs for users staying on the road), the loss of welfare for consumers will appear higher than what is really the case. I will now analyze the four groups separately. The assumptions stands contrary to the depiction in figure 11, where the line of the generalized cost curve (GC) is not horizontal.

The motorists who remain on the road enjoy the benefit of lower time costs, but pay a higher toll. In total their net change in welfare is $Q_1(p_1-p_0)$, as seen in figure 11. They enjoy a decrease in time costs, but the congestion toll is higher than the cost reduction, so their net cost increase is $Q_1(p_1 - p_0) = Q_1(\tau_1 - \tau_0 + GC(Q_1) - GC(Q_0))$. To simplify calculations, I will estimate a lower bound for the welfare effect. For this particular group, this means disregard the change in generalized cost, and assume it is constant before and after introducing congestion pricing. Under the assumption that time costs are unchanged, I have that $Q_1(p_1 - p_0) = Q_1(\tau_1 - \tau_0)$.

The motorists who leave the road, save travel costs equal to $p_0(Q_0-Q_1) = (\tau_0+GC(Q_0))(Q_1-Q_0)$, and incur a welfare loss equal to $-(\tau_0 + GC(Q_0))(Q_0 - Q_1) + (Q_0 - Q_1)(p_0 - p_1)\frac{1}{2}$, which

is the area under the demand curve Q_h between Q_1 and Q_0. To simplify calculations, I will estimate a lower bound for the welfare effect. For this particular group, this means disregard the change in generalized cost, and assume it is constant before and after introducing congestion pricing. Under the assumption of $GC(Q_1) = GC(Q_0)$, their net change in welfare is: $(\tau_0 + GC(Q_0))(Q_0 - Q_1) - (\tau_0 + GC(Q_0))(Q_1 - Q_0) + (Q_0 - Q_1)(p_0 - p_1)\frac{1}{2} = (Q_0 - Q_1)(\tau_0 - \tau_1)\frac{1}{2}$. This equals the triangle under the demand curve going from Q_1 to Q_0 in figure 11.

The government increases its revenue by $Q_1(\tau_1 - \tau_0)$, as the toll is raised by $\tau_1 - \tau_0$ for the Q_1 remaining motorists on the road link.

The rest of society enjoy a benefit equal to the triangle $\frac{1}{2}\tau(Q_0)(Q_0 - Q_1)$, which is the reduction in the deadweight loss due to flow being higher than what is socially optimal. This is the triangle BGF in figure 11. Time costs are a large part of the benefits to society, so I include them here, meaning that generalized costs are not assumed to be constant in this part of the estimation.

Adding all the different components above together enables me to find the social surplus. That is, adding the effect for motorists remaining on the road, motorists leaving the road, government, and society as a whole. The fact that the estimated optimal toll and marginal external costs are normalized per PCE kilometer, creates the need to assume something about the driving length for motorists. This is to make the estimated toll comparable to the present toll, which is given as 31 NOK in total. Based on data from (TØI, 2011a), I use an assumed driving length $K = 12.6$. Summing over the different components of social surplus it becomes[48]

$$\Delta SS = \frac{1}{2}(Q_0 - Q_1)\left[\tau_0 - K(\tau_1 - \tau(Q_0))\right] \tag{93}$$

To find the social surplus for all periods, I rely on the assumption that demand is independent over time, to write the surplus as the sum over all T periods.

$$\Delta SS_D = \frac{1}{2}\sum_{t=1}^{T}(Q_0^t - Q_1^t)[\tau_0^t - K(\tau_1^t - \tau(Q_0))] \tag{94}$$

where ΔSS_D is the daily change in consumer's surplus, Q_0^t mean flow at present, Q_1^t the estimated equilibrium flow, τ_0^t is the current total toll, τ_1^t the estimated equilibrium toll and $\tau_t(Q_0^t)$ the estimated average marginal external costs at the present. All variables are indexed by time to indicate that they belong to time period t. The total change is then found by aggregating over time, assuming demand is independent between each time period. I remind the reader that the expression is a lower bound on the change in welfare. Details on the computation can be found in the appendix.

---

[48]See appendix for derivation.

8.3   Results

In this section the results of the estimations are given.

Table 4 gives the estimated values of the reduction in external costs, consumer's surplus, daily change to social surplus, and yearly change under different assumptions on the elasticities. As can be seen, the estimated social benefits are quite sizable, keeping in mind that they only consider traffic on one specific link, and in only one direction. The lowest estimate of the yearly benefit is 3.8 million NOK, and the highest 29 million NOK. The estimation thus seems to support the theoretical arguments that congestion pricing will increase social surplus.

Table 4:

Estimated change to social surplus from introducing congestion pricing. Estimated with different values on the elasticity of PCE flow with respect to time costs and tolls. All costs are measured in NOK. CS = Consumer's surplus change, EC = External costs change (deadweight loss), SS = Daily social surplus change, $SStotal$ = Yearly change to social surplus, assuming 260 weekdays. PCE = Passenger car equivalent elasticity, GC = Elasticity of generalized costs. Elasticities gathered from the simulation in appendix I.

Elasticity	Mean	Min	Max
PCE	-.091	-.147	-.03
GC	-.23	-.25	-.21
CS	-53 206	-62 087	-25 200
EC	107 625	173 318	39 797
SS	54 418	111 230	14 597
$SS\ total$	14 148 857	28 919 872	3 795 291

Different assumptions regarding the elasticities does not seem to change the main conclusion, but rather underline it. More elastic demand indicates a higher change in the social surplus. The difference between the highest and lowest estimate is 25 million NOK, with only a difference of 0.117 percentage points between the elasticities. This indicates that the results are highly sensitive towards the assumptions regarding the elasticities, but not towards the main conclusion, that social benefit increases. Even for elasticities as low as -0.03, the estimated yearly benefit is still 3.8 million NOK. One also needs to keep in mind that for these estimates are likely to be low for several reasons. They only apply to a single link, in a single direction, and only during weekdays. Moreover, as I was not able to estimate the optimal congestion tolls, present external costs and flows during rush hours, these benefits are likely to be even higher still.

9   Conclusion

In this thesis, I have estimated the welfare effects of introducing congestion pricing, using a specific link in the Oslo cordon as an example. I extended a rule for congestion pricing based on Johansson (1997) to include multiple periods of the day. Moreover, I integrated this framework with the mathematical bi-level formulation found in the literature on road pricing. This was to provide a theoretical framework behind the model of a trial-and-error process of Li (2002) to find the optimal congestion toll. From the optimization problem, I derived a system of equations characterizing the equilibrium with congestion tolls. Using fixed-point iteration to solve this system after parameterization, I found an estimate of the optimal congestion tolls, and demand. Using these results, I estimated the change in social surplus of moving from today's pricing scheme towards one with congestion pricing.

The main result is that congestion pricing seems to provide an increase in social benefits, as promoted by theory. Moreover, my results indicate that these benefits are of some size, and still so under different assumptions regarding the price sensitivity of traffic. These benefits provide what is most likely a conservative estimate of the true benefits for several reasons. First, the formulas used to calculate changes in consumer's surplus and external costs, are only approximations which will provide lower bound estimates if some rather stringent criteria are not met. Second, there are several other external costs, associated with noise and accidents that are not included. Third, the direct environmental impact of cars is not taken into account due to time constraints. Fourth, as there is evidence of an equilibrium in traffic jam-like conditions during rush hours, I am not able to calculate the optimal toll during this period. Thus, I find no new level of equilibrium traffic, which precludes any estimation of social benefits in rush hours. However, the estimates that are provided are most likely too low, but do provide a conservative estimate, and do not bias the results towards reaching the opposite conclusion.

The results however, hinges on elasticities estimated for daily average traffic on cross-section data. As there is evidence that the purpose of travel changes over the course of day, and travel purpose affects price sensitivity, the results should be cautiously interpreted. The effect of social benefits might be higher in some periods, and lower in other due to differences in elasticities, and the net effect is ambiguous. The estimated equilibria, and thus social benefits, also rest on the assumption that I have modeled the adjustment system for the three simultaneous equations correctly.

A natural extension of this thesis is to relax some of the assumptions made in the theoretical section. First, relaxing the assumptions of independent demand with respect to other routes on the network is an obvious one. Second, the assumption of no substitution between modes could

be relaxed. Estimating queue length is also a possible extension, as I was not able to estimate a toll during rush hours because of insufficient data to determine queue length.

In my view, this thesis has highlighted some practical difficulties in estimating a vector of congestion tolls using data on speed and flow. Lastly, it seems to provide some evidence in favor of the theoretical result that congestion pricing increases social benefits, as I have at least not found any results suggesting otherwise.

References

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, 19(6):716–723.

Angrist, J. D. and Pischke, J.-S. (2008). *Mostly harmless econometrics: An empiricist's companion*. Princeton, New Jersey: Princeton university press.

Arnott, R., De Palma, A., and Lindsey, R. (1990). Economics of a bottleneck. *Journal of urban Economics*, 27(1):111–130.

Arnott, R., De Palma, A., and Lindsey, R. (1993). A structural model of peak-period congestion: A traffic bottleneck with elastic demand. *The American Economic Review*, 83(1):161–179.

Atkinson, S. E. and Wilson, P. W. (1992). The bias of bootstrapped versus conventional standard errors in the general linear and sur models. *Econometric Theory*, 8(2):258–275.

Balwani, A. and Singh, S. (2009). Network impacts of distance-based road user charging. *NETNOMICS: Economic Research and Electronic Networking*, 10(1):53–75.

Ben-Akiva, M., Cyna, M., and De Palma, A. (1984). Dynamic model of peak period congestion. *Transportation Research Part B: Methodological*, 18(4):339–355.

Benekohal, R. F. and Zhao, W. (2000). Delay-based passenger car equivalents for trucks at signalized intersections. *Transportation Research Part A: Policy and Practice*, 34(6):437–457.

Boadway, R. (2006). Principles of cost-benefit analysis. *Public Policy Review*, 2(1):1–44.

Breusch, T. S. and Pagan, A. R. (1980). The lagrange multiplier test and its applications to model specification in econometrics. *The Review of Economic Studies*, 47(1):239–253.

BT (2015). Bergens tidende: Nå er køprising vedtatt. http://www.bt.no/nyheter/lokalt/Na-er-koprisingen-vedtatt-3305245.html. Accessed 15.05.2015.

Button, K. (2004). The rationale for road pricing: Standard theory and latest advances. In Santos, G., editor, *Road Pricing: Theory and Evidence*. Amsterdam: Elsevier.

Button, K. (2010). *Transport economics*. London: Edward Elgar Publishing.

Cameron, A. C. and Trivedi, P. K. (2005). *Microeconometrics: methods and applications*. Cambridge: Cambridge university press.

Cameron, A. C. and Trivedi, P. K. (2010). *Microeconometrics using stata*. College Station, Texas: Stata Press.

Chen, M., Bernstein, D. H., Chien, S. I., and Mouskos, K. C. (1999). Simplified formulation of the toll design problem. *Transportation Research Record: Journal of the Transportation Research Board*, 1667(1):88–95.

Chen, M., Bernstein, D. H., and Spasovic, L. N. (2004). Toll-design problem with stochastic route choice. *Environment and Planning B*, 31(5):731–742.

D'agostino, R. B., Belanger, A., and D'Agostino Jr, R. B. (1990). A suggestion for using powerful and informative tests of normality. *The American Statistician*, 44(4):316–321.

Davidson, R. and MacKinnon, J. G. (2004). *Econometric theory and methods*, volume 5. Cary, North Carolina: Oxford University Press New York.

De Palma, A. and Lindsey, R. (2004). Congestion pricing with heterogeneous travelers: A general-equilibrium welfare analysis. *Networks and spatial economics*, 4(2):135–160.

De Palma, A. and Lindsey, R. (2011). Traffic congestion pricing methodologies and technologies. *Transportation Research Part C: Emerging Technologies*, 19(6):1377–1399.

Ekström, J. (2008). *Designing Urban Road Congestion Charging Systems, Phd Thesis, Linköping University*. Department of Science and Technology, Linköping University.

Eliasson, J. (2009). A cost–benefit analysis of the stockholm congestion charging system. *Transportation Research Part A: Policy and Practice*, 43(4):468–480.

Ferrari, P. (2002). Road network toll pricing and social welfare. *Transportation Research Part B: Methodological*, 36(5):471–483.

Friedrich, R. and Quinet, E. (2011). External costs of transport in europe. In De Palma, A., Lindsey, R., Quinet, E., and Vickerman, R., editors, *A handbook of transport economics*. London: Edward Elgar Publishing.

Goodwin, P. B. (1992). A review of new demand elasticities with special reference to short and long run effects of price changes. *Journal of transport economics and policy*, 26(2):155–169.

Immers, L. and Stada, J. (2004). Lecture notes: Basics of transport economics. *Faculty of Engineering, Katholike Universiteit Leuven*, http://www.kuleuven.be/traffic/dwn/H111part4.pdf. Accessed 15.05.2015.

Johansson, O. (1997). Optimal road-pricing: simultaneous treatment of time losses, increased fuel consumption, and emissions. *Transportation Research Part D: Transport and Environment*, 2(2):77–87.

Jones, P. and Hervik, A. (1992). Restraining car traffic in european cities: an emerging role for road pricing. *Transportation Research Part A: Policy and Practice*, 26(2):133–145.

Jou, R.-C. and Chen, T.-Y. (2014). Factors affecting public transportation, car, and motorcycle usage. *Transportation research part A: policy and practice*, 61:186–198.

Kimber, R., McDonald, M., and Hounsell, N. (1985). Passenger car units in saturation flows: concept, definition, derivation. *Transportation Research Part B: Methodological*, 19(1):39–61.

Li, M. Z. (2002). The role of speed–flow relationship in congestion pricing implementation with an application to singapore. *Transportation Research Part B: Methodological*, 36(8):731–754.

Maerivoet, S. and De Moor, B. (2005). Traffic flow theory. *arXiv preprint physics/0507126*, http://arxiv.org/pdf/physics/0507126v1.pdf. Accessed 15.05.2015.

Matas, A. and Raymond, J.-L. (2003). Demand elasticity on tolled motorways. *Journal of Transportation and Statistics*, 6(2):91–108.

Mayeres, I., Ochelen, S., and Proost, S. (1996). The marginal external costs of urban transport. *Transportation Research Part D: Transport and Environment*, 1(2):111–130.

Meld.St.17(2008-2009):. *Om Oslopakke 3 trinn 2 (On the Oslo Package 3 step 2)*. Oslo: Det Kongelige Samferdsselsdepartement.

Mun, S.-I. (1994). Traffic jams and the congestion toll. *Transportation Research Part B: Methodological*, 28(5):365–375.

Neuburger, H. (1971). User benefit in the evaluation of transport and land use plans. *Journal of Transport Economics and Policy*, 5(1):52–75.

NRK (2013). Kø koster 76 millioner kroner per dag i 2030. http://www.nrk.no/valg2013/ko-koster-76-mill.-per-dag-i-2030-1.11201678. Accessed 15.05.2015.

Odeck, J. and Bråthen, S. (2008). Travel demand elasticities and users attitudes: A case study of norwegian toll projects. *Transportation Research Part A: Policy and Practice*, 42(1):77–94.

Odeck, J. and Kjerkreit, A. (2008). Priselastisiteter og bompengeinnkreving - innføring av bompengeinnkreving på e18 Østfold (price elasticities and road tolling - introduction of tolls at at e18 in Østfold). Technical report, Statens vegvesen - Utbyggingsavdelingen - Transportanalyseksjonen.

Oum, T. H., Waters, W. G., and Yong, J.-S. (1992). Concepts of price elasticities of transport demand and recent empirical estimates: an interpretative survey. *Journal of Transport Economics and policy*, 26(2):139–154.

Pigou, A. C. (1924). *The economics of welfare*. Piscataway, New Jersey: Transaction Publishers.

Schwarz, G. (1978). Estimating the dimension of a model. *The annals of statistics*, 6(2):461–464.

Sen, A. K., Tiwari, G., and Upadhyay, V. (2010). Estimating marginal external costs of transport in delhi. *Transport Policy*, 17(1):27–37.

Silberberg, E. (1972). Duality and the many consumer's surpluses. *The American Economic Review*, 62(5):942–952.

Slesnick, D. T. (1998). Empirical approaches to the measurement of welfare. *Journal of Economic Literature*, XXXVI:2108–2165.

SVV (2014). *Håndbok V712 Konsekvensanalyser (Handbook V712 Impact Analysis)*. Oslo: Statens Vegvesen, Vegdirektoratet.

Toftegaard, H. (2007). Priselastisiteter og bompengeinnkreving - innføring av bommen på sykkylvsbrua i møre og romsdal (price elasticities and road tolling - introduction of tolls at sykkylvsbrua in møre and romsdal). Technical report, Statens vegvesen - Utbyggingsavdelingen - Transportanalyseksjonen.

TRB (1994). Trb special report 209: Highway capacity manual, 3rd edition. *National Research Council, Washington, DC*.

TØI (2010). Den norske verdsettingsstudien - tid (value of time, safety and environment in passenger transport – time). *TØI rapport 1053b*, Oslo: Transportøkonomisk Institutt.

TØI (2011a). Den nasjonale reisevaneundersøkelsen 2009: nøkkelrapport (norwegian national travel survey - executive summary report). *TØI rapport 1130*, Oslo: Transportøkonomisk institutt.

TØI (2011b). No2-utslipp fra kjøretøyparken i norske storbyer (no2 emission from the vehicle fleet in major norwegian cities). *TØI rapport 1168*, Oslo: Transportøkonomisk Institutt.

TØI (2014). Marginale eksterne kostnader ved vegtrafikk (external marginal costs of road traffic). *TØI rapport 1307*, Oslo: Transportøkonomisk Institutt.

Van Aerde, M. and Yagar, S. (1984). Capacity, speed and platooning vehicle equivalents for two-lane rural highways. (971):58–67.

Varian, H. R. . (1992). *Microeconomic analysis. Third Edition.* New York: WW Norton.

Walters, A. A. (1961). The theory and measurement of private and social cost of highway congestion. *Econometrica: Journal of the Econometric Society*, 29(4):676–699.

Wardrop, J. G. (1952). Some theoretical aspects of road traffic research. In *ICE Proceedings: Engineering Divisions*, volume 1, pages 325–362. London: Thomas Telford.

Willig, R. D. (1976). Consumer's surplus without apology. *The American Economic Review*, 66(4):589–597.

Wood, A. (1999). *Introduction to numerical analysis.* Essex: Addison Wesley.

Wooldridge, J. M. (2010). *Econometric analysis of cross section and panel data.* Cambridge Massachusetts: The MIT press.

Yang, H. and Bell, M. G. (1997). Traffic restraint, road pricing and network equilibrium. *Transportation Research Part B: Methodological*, 31(4):303–314.

Yang, H. and Huang, H.-J. (1998). Principle of marginal-cost pricing: how does it work in a general road network? *Transportation Research Part A: Policy and Practice*, 32(1):45–54.

Zellner, A. (1962). An efficient method of estimating seemingly unrelated regressions and tests for aggregation bias. *Journal of the American statistical Association*, 57(298):348–368.

A   Decomposition of AADT elasticity.

First, define the daily traffic volume Q, as the sum of of traffic volumes over the T periods of the day, Q_t; $Q = \sum_{i=1}^{T} Q_i$. Now, suppose the toll τ, is increased by the same amount all through the day. The change to daily demand is the comprised of the change to demand in each period t:

$$\frac{\partial Q}{\partial \tau} = \sum_{t=1}^{T} \frac{\partial Q_t}{\partial \tau_t} \tag{95}$$

Multiplying with $\frac{Q_t}{\tau_t} \frac{\tau_t}{Q_t}$ on the left hand side, enables one to rewrite this as the sum of demand in each period, weighted by the elasticity for each particular period ε_t:

$$\frac{\partial Q}{\partial \tau_t} = \sum_{t=1}^{T} \frac{Q_t}{\tau_t} \varepsilon_t \tag{96}$$

Assuming the toll is the same in every period (no congestion pricing; $\tau = \tau_t$), it can be moved outside the sum

$$\frac{\partial Q}{\partial \tau} = \frac{1}{\tau} \sum_{t=1}^{T} Q_t \varepsilon_t \tag{97}$$

Multiplying this by $\frac{\tau}{Q}, Q = \sum_{i=1}^{T} Q_i$ on both sides, gives the elasticity of daily demand on the right hand side, and a weighted average of the period-wise elasticities on the left hand side:

$$\varepsilon = \frac{\tau}{Q} \frac{\partial Q}{\partial \tau} = \frac{\tau}{Q} \frac{1}{\tau} \sum_{t=1}^{T} Q_t \varepsilon_t = \sum_{t=1}^{T} \underbrace{\frac{Q_t}{Q}}_{=a_t} \varepsilon_t \Rightarrow \varepsilon = \sum_{t=1}^{T} a_t \varepsilon_t \tag{98}$$

where a_t is the share of daily demand in period t; $a_t = \frac{Q_t}{\sum_{i=1}^{T} Q_i}$. The expression shows that the elasticity of daily traffic volume, will be influenced more by periods where demand is high, all else equal, and the elasticity is high as well, all else equal.

B   Calculating passenger equivalent flow

There are two principle difficulties in calculating the PCE flow. The first being that only a count variable of each vehicle is observed, and not the congestion factor. This means that the differential impact on congestion from small and large vehicles is not identified. The congestion factor is essentially what drives the difference between light and heavy vehicles. To account for this, I adjust the traffic flow towards the passenger equivalent flow, by using the inverse of the heavy vehicle adjustment factor developed in (TRB, 1994).

The inverse can be derived in the following way: Let Q_r^{TC} be the true, observed count of passenger car units. This will be the sum of the true count of heavy vehicles Q_H^{TC}, scaled by the congestion factor α, and the count for light vehicles, Q_L^{TC}.

$$Q_r = \alpha Q_H^{TC} + Q_L^{TC} \tag{99}$$

Now, assuming heavy and light vehicles constitute a fixed proportion of traffic, where π is the share of heavy ones, and $1 - \pi$ the corresponding for light, I have:

$$Q_H^{TC} = \pi Q_T^{TC}, Q_L^{TC} = (1 - \pi)Q_T^{TC} \tag{100}$$

Here, Q_H^{TC} is the true count of heavy vehicles, and Q_L^{SC} the corresponding ones light light vehicles. These constitute a fixed proportion of the true, total count on the link.

Substituting these into the first equation gives me the following expression for the true count of passenger car equivalents

$$Q_r = \alpha \pi Q_T^{TC} + (1 - \pi)Q_T^{TC} = (1 - (1 - \alpha)\pi)Q_T^{TC} \tag{101}$$

As one can see, having the true count of vehicles on the link, enables one to obtain an estimate of the passenger car equivalent.

$$Q_r = [1 - (1 - \alpha)\pi]Q_T^{TC} \tag{102}$$

The second challenge, is that only cars that have an electronic registration device are counted. This means that the actual flow of vehicles will be higher than what is observed in my data. Assuming there is only one lane for simplicity, my observed count (flow) value, Q_t^{OC} will be bounded by the true count value Q_T^{TC}, by the share of users with an registration device λ_t at time t: $Q_t^{OC} = \lambda_t Q_t^{TC}$, $\lambda_t \in (0, 1)$. As one can see, this will potentially have implications for the measurement of my variable. To account for this, one could start by assuming that the share of users having a device is constant throughout during the day ($\lambda_t = \lambda, \forall t$), and scale the traffic volume to find the true volume, that is: $Q_t^{TC} = Q_t^{OC}/\lambda$.

Resolving these two problems is done by adjusting traffic volume accordingly, by using the adjustment formula above, substituting in $Q_t^{TC} = Q_t^{OC}/\lambda$, into it. Then I get

$$Q_r = \frac{[1 - (1 - \alpha)\pi]Q_t^{OC}}{\lambda} \tag{103}$$

Using this adjustment factor, I simultaneously account for missing counts, and difference in congestion factors.

C Proportionality criteria for the congestion function

To see that the proportionality criteria for the congestion function is satisfied, I will show that

$$\alpha_k = \frac{C_k}{C_1} \tag{104}$$

where C_k is the speed reduction coefficient for vehicle type k, and C_1 the coefficient for passenger cars. This is the definition of passenger car units given by (Van Aerde and Yagar, 1984). I

assume that the speed reduction coefficients are the inverse of the derivative of the time function, $C_k = T_i'(Q^{TC})^{-1} = s'(Q_r)$. To see that (104) holds, replace Q_r with the passenger car flow variables of light and heavy vehicles as done in appendix B. Now, one can rewrite the congestion function as a function of the individuals PCE counts of heavy and light vehicles:

$$T_i = \frac{1}{s_t} = \beta_0[1 + e^{\beta_1(\alpha Q_L^{TC} + Q_S^{TC})}] \tag{105}$$

Then, let $\Gamma = \beta_0[1 + e^{\beta_1(\alpha Q_L^{TC} + Q_S^{TC})}]$ and $\Gamma' = \beta_0 e^{\beta_1(\alpha Q_L^{TC} + Q_S^{TC})}$, then stacking the inverse of the time derivatives on top of another gives

$$\frac{C_k}{C_1} = \frac{T_i'(Q_L^{TC})^{-1}}{T_i'(Q_S^{TC})^{-1}} = \frac{-\alpha_k \Gamma'/\Gamma^2}{-\Gamma'/\Gamma^2} = \frac{\alpha_k \Gamma'}{\Gamma^2}\frac{\Gamma^2}{\Gamma'} = \alpha_k \tag{106}$$

which shows that the proposed congestion function would enable one to identify the congestion factor by just scaling the function, as the only difference between the congestive impact of group k and passenger cars is α_k.

C.1   Details on estimation of the change in social surplus

Adding all the different components given in section 8 of the main text together, enables me to find the social surplus. That is, adding the effect for motorists remaining on the road, motorists leaving the road, government, and society as a whole. Taken in that order, the sum becomes

$$
\begin{aligned}
\Delta SS ={} & \Delta CS + \Delta EC + \Delta OS \\
={} & Q_1(\tau_0 - \tau_1) + \frac{1}{2}(Q_0 - Q_1)(\tau_0 - \tau_1) + Q_1(\tau_1 - \tau_0) + \frac{1}{2}\tau(Q_0)(Q_0 - Q_1) \\
={} & Q_1(\tau_0 - \tau_1) - Q_1(\tau_1 - \tau_0) + \frac{1}{2}(Q_0 - Q_1)(\tau_0 - \tau_1) + \frac{1}{2}\tau(Q_0)(Q_0 - Q_1) \\
={} & \frac{1}{2}(Q_0 - Q_1)\left[\tau_0 - \tau_1 + \tau(Q_0)\right]
\end{aligned} \tag{107}
$$

The fact that the estimated optimal toll and marginal external costs are normalized per PCE kilometer, creates the need to assume something about the driving length for motorists. This is to make the estimated toll comparable to the present toll, which is given as 31 NOK in total. Based on data from (TØI, 2011a), I use an assumed driving length $K = 12.6$. Summing over the different components of social surplus it becomes[49]

$$\Delta SS = \frac{1}{2}(Q_0 - Q_1)\left[\tau_0 - K(\tau_1 - \tau(Q_0))\right] \tag{108}$$

To find the social surplus for all periods, I rely on the assumption that demand is independent over time, to write the surplus as the sum over all T periods.

$$\Delta SS_D = \frac{1}{2}\sum_{t=1}^{T}(Q_0^t - Q_1^t)[\tau_0^t - K(\tau_1^t - \tau(Q_0))] \tag{109}$$

---

[49]See appendix for derivation.

where ΔSS_D is the daily change in consumer's surplus, Q_1^t the estimated equilibrium flow, Q_0^t the mean flow at present, τ_0^t is the current total toll, τ_1^t the estimated equilibrium toll and $\tau_t(Q_0^t)$ the estimated average marginal external costs at the present. All variables are indexed by time to indicate that they belong to time period t. The total change is then found by aggregating over time, assuming demand is independent between each time period. I remind the reader that the expression is a lower bound on the change in welfare.

The aggregation process is the following: The iteration scheme produced a set of $T \times 1$ vectors containing the needed quantities to estimate the change in consumer's surplus. Let $\mathbf{Q}_1$ be the vector of optimal PCE flows, $\mathbf{Q}_0$ the vector of present mean flow at the road link, $\boldsymbol{\tau}(\mathbf{Q}_0)$ the present marginal costs and $\boldsymbol{\tau}(\mathbf{Q}_1)$ the optimal congestion tolls per PCE. Finally, let $\mathbf{p}_0 = 31\mathbf{k}$ where $\mathbf{k}$ is an $T \times 1$ matrix containing 1 in each entry, be the vector of current toll per PCE. The estimated vectors containing the tolls and flow over all periods then look like this:

$$\mathbf{Q}_1 = \begin{pmatrix} Q_{1,1} \\ Q_{1,2} \\ \vdots \\ Q_{1,T} \end{pmatrix}, \mathbf{Q}_0 = \begin{pmatrix} Q_{0,1} \\ Q_{0,2} \\ \vdots \\ Q_{0,T} \end{pmatrix}, \boldsymbol{\tau}(\mathbf{Q}_0) = \begin{pmatrix} \tau_{0,1} \\ \tau_{0,2} \\ \vdots \\ \tau_{0,T} \end{pmatrix}, \boldsymbol{\tau}(\mathbf{Q}_1) = \begin{pmatrix} \tau_{1,1} \\ \tau_{1,2} \\ \vdots \\ \tau_{1,T} \end{pmatrix} \tag{110}$$

Using these matrices, the sum over all periods can be found as:

$$\Delta SS_{TOTAL} = \frac{1}{2}(\mathbf{Q}_0 - \mathbf{Q}_1)'(\mathbf{p}_0 - K[\boldsymbol{\tau}(\mathbf{Q}_1) - \boldsymbol{\tau}(\mathbf{Q}_0)]) \tag{111}$$

where $'$ denotes the transpose. This expression equals the sum in (109), and provides an estimate of the social surplus change.

D   Details on the numerical algorithm

To estimate the equilibrium, I use fixed-point iteration. Some details on the technical aspects, and implementation are given in this appendix.

The numerical algorithm is given below. First, initial toll, demand and generalized costs are determined. $\tau_0 = 31$, Q_0 is estimated by data, and GC_0 is estimated by using Q_0 in the equation for generalized costs. From $h(Q_0, \tau_0)$, the starting value is determined as $Q(Q_0, \tau_0) = Q_0[1 + h(Q_0, \tau_0)]$. Then, a loop starts. Tolls and generalized costs are determined iteratively, and at each step, convergence of Q is evaluated. If the absolute value of difference between Q_{k+1} and Q_k falls below a stopping criterion $\eta = 0.000000001$, the iteration stops, and the results are reported.

Algorithm 1: Algorithm to solve for the traffic equilibrium

    Input: Input

    Initial τ_0, GC_0 and Q_0;

    Initialization: Compute $Q(\tau_0)$, set $n = 1$;

    while $n > 1$ **do**

        Compute $\tau_n = g(Q_n)$;

        Compute $GC_n = h(Q_n)$;

        Compute $Q_{n+1} = v(\tau_n, GC_n)$;

1   Compute $a = Q_{n+1} - Q_n$;

        if $a < \eta$ **then**

         | **Output**: τ_n, Q_n

        else

         | n = n+1;

        end

    end

D.1 The complete equation system

The toll, τ, is given as the sum of the marginal external costs associated with time and emissions. Each pollutant has its own equation. Written out in its entirety, the system looks like this.

$$v(Q)_t = Q\left[1 + \varepsilon_{t,\tau}\left(\frac{\tau_t(Q_k)}{\tau_t(Q_{k-1})} - 1\right) + \varepsilon_{t,GC}\left(\frac{GC_t(Q_k)}{GC_t(Q_{k-1})} - 1\right)\right]$$

$$\tau(Q) = \alpha_k \beta_0 \beta_1 e^{\beta_1 Q_r} Q_r [\pi VOT_H + (1-\pi)VOT_L]$$

$$+ \quad \alpha_k \nu_1 \nu_0 \exp(\nu_1 Q_r) Q_r \sum_{j \in \mathbf{C}} VOE_j \left[\gamma_{1H,j}\gamma_{0H} S^{\gamma_{1H,j}-1}\pi + \gamma_{1L,j}\gamma_{0L,j} S^{\gamma_{1L,j}-1}(1-\pi)\right]$$

$$GC(Q)_{k+1} = \beta_0(1 + e^{\beta_1 Q_k})VOT$$

$$(112)$$

where $\mathbf{C} = (NO_X, PM_{10}, CO_2)$.

E Details on the assumptions of the congestion price model

E.0.1 Limitations of the model

The Johansson model did not account for the case of multiple periods. As it is an aggregate model, it does not consider the possibility of changes on the road network, or between different modes of transportation. In this section, I discuss under what assumptions the model is still applicable to a single link, as I intend to use it. I show that under assumptions of independent

demand across time, the pricing rule is still valid, period-by-period. This will be important since I intend to calculate a congestion price that varies over the course of the day. Moreover, I discuss the effect of only tolling one single link, and show that a similar assumption needs to apply. Lastly, I discuss some possible implications of changes in transportation mode.

Multiple time periods: Let $\mathbf{Q}_r = (Q_{r1}, ..., Q_{rT})$ be a vector of total demand in periods $t = 1, ..., T$, defined as the flow in the corresponding period. Then, total social surplus for all T periods can be written as the sum over T periods, which can be written as a sum of the different social benefits for the different periods, that is

$$NB(\mathbf{Q}_r) = \sum_{t=1}^{T} NB_t(Q_r) \tag{113}$$

Now, letting demand in period t, be a function of tolls in all periods, one can redefine the vector of demand functions as: $\mathbf{Q}_r = (Q_{r1}(\mathbf{p}), ..., Q_{rT}(\mathbf{p}))$, where $\mathbf{p} = (p_1, ..., p_T)$ is a vector of tolls that vary over T periods. To study the effect of increasing the toll in one period, take the derivative of $NB(\mathbf{Q}_r(\mathbf{p}))$ with respect to the toll in period k, which is the sum of the derivative of all social benefits over T:

$$\frac{\partial NB}{\partial p_{t_k}} = \sum_{t=1}^{T} \frac{\partial NB_t(Q_{rt})}{\partial Q_{rt}} \frac{\partial Q_{rt}}{\partial p_{t_k}} \tag{114}$$

That is, the effect on social benefits is comprised of the product of the derivative of the net benefit in period t, multiplied by the change in demand in that period, with respect to the price in period k. Dividing and multiplying with p_{t_k} and Q_{rt}, this can be rewritten as:

$$\frac{\partial NB}{\partial p_{t_k}} = \sum_{t=1}^{T} \frac{\partial NB_t(Q_{rt})}{\partial Q_{rt}} \frac{\partial Q_{rt}}{\partial p_{t_k}} \frac{p_{t_k}}{Q_{rt}} \frac{Q_{rt}}{p_{t_k}} \tag{115}$$

Noting that $\frac{\partial Q_{rt}}{\partial p_{t_k}} \frac{p_{t_k}}{Q_{rt}} = \epsilon_{t,t_k}$, that is, the elasticity of demand in period t with respect to the price in period t_k, this is the same as

$$\frac{\partial NB}{\partial p_{t_k}} = \sum_{t=1}^{T} \frac{\partial NB_t(Q_{rt})}{\partial Q_{rt}} \frac{Q_{rt}}{p_{t_k}} \epsilon_{t,t_k} \tag{116}$$

If the price elasticity is zero for all periods other than the one the price is changed at, meaning that: $\epsilon_{t,t_k} = 0, \forall t \neq t_k$, the expression reduces to:

$$\frac{\partial NB}{\partial p_{t_k}} = \frac{\partial NB_t(Q_{rt})}{\partial Q_{rt}} \frac{\partial Q_{rt}}{\partial p_{t_k}} \tag{117}$$

which is just the change in the social surplus in the period where the price is changed. If demand is not independent, the full effect of changing the price vector $\mathbf{p}$ might not be identified, as changing the tolls in one period, affects the social surplus in other periods as demand is shifted between periods. For example, if raising the toll in one period shifts demand towards another

period, the externalities will be reduced in one period, but possibly increased in another, meaning that the net effect on social surplus can be lower than what is believed when considering one period only. Thus, it is hard to produce any general statements, without conducting a more formal analysis, which out of scope, for this thesis. To use the specific results concerning the road price in this thesis, I will make the following assumption

Assumption 1: Demand is independent over time; $\epsilon_{t,t_k} = 0, \forall t \neq t_k$

The assumption entails that changing the toll in one period, does not affect demand in other periods of the day. The appropriateness of this assumption may vary over the course of the day. For example, during the morning rush hours, demand is more peaked, which might suggest that users have less flexibility in the timing of their road use, whereas the afternoon traffic is more evenly spread out in time.

Multiple routes: A second caveat, follows the same logic as above, but restates it. There might not just be possibilities for road users to substitute between different time periods, but also different routes. As such, raising the price at one point, might divert traffic to another route, increasing congestion and externalities, which leads to an overestimation of the correct congestion price, as the adverse effects on other routes are not accounted for. I chose the particular road link, since it seems reasonable that it is not as prone to be easily substituted as a route choice, than, say, a smaller county road. This thesis relies on the following assumption

Assumption 2: There are no externalities in the road network from altering tolls.

When the assumption holds, there will be no substitution from using the road link to other road links. Suppose first that this does there is in fact some substitution from the tolled road to the untolled road. As traffic increases on the other link, marginal social costs also rise, introducing an externality. Thus, this assumptions concerns $-FABD$ in the accounting table.

To illustrate this point, I consider a very simple network model. Suppose there are two different routes a that can be chosen from an origin (O) to a destination (D). Let the total flow from O to D be denoted Q, which is set equal to 1 for simplicity. Further, let the flow on route $i = 1, 2$ be denoted as Q_i, where $Q_1 + Q_2 = Q = 1$. Further, let the cost per flow unit be given as the sum of all marginal costs, as given in the Johansson model; $c_{r_i} = \sum_j MC_j(Q_{r_1})$, where j denotes the jth cost component. The government planner would seek to minimize the social cost of channeling traffic flow from O to D, that is

$$\min_{Q_{r_1}, Q_{r_2}} Q_{r_1} c_{r_1} + Q_{r_2} c_{r_2} \ \ s.t. \ \ Q_{r_1} + Q_{r_2} = 1 \tag{118}$$

Let λ denote the lagrange multiplier of the constraint. The first-order conditions for the problem

are given by

$$c_{r_1} + Q_{r_1} \sum_j MC_j'(Q_{r_1}) = c_{r_2} + Q_{r_2} \sum_j MC_j'(Q_{r_2}) = \lambda \tag{119}$$

The condition states that at the minimum, the social cost of using route 1 should equal the social cost of using route 2. The two summation terms contain the externalities, and the toll on each road correspond to these. The private user equilibrium characterized by the fact that externalities are not taken into account. That is, if the minimization problem is viewed as from a representative road user's point of view, the effect on cost c_j is from an increase in the flow, is not taken into consideration. The first-order condition is then just

$$c_{r_2} = c_{r_1} \tag{120}$$

which states that the private cost of the routes should be equal. This is just the Wardrop's user equilibrium in its simplest form (Wardrop, 1952). Now, suppose that just one of the routes are tolled. Let τ_i denote the toll on route j. Assume further that the network was in user equilibrium prior to introduction of the road price. Now if the toll on route 1 is set to equal the marginal external costs, condition (120) becomes

$$c_{r_1} + \tau_1 > c_{r_2} \tag{121}$$

Assuming that the constraint is binding at the minimum, Q_1 can be rewritten as $1 - Q_2$. Using this and noting that one can substitute in for c_{r_1}, c_{r_2} τ_1, the condition becomes

$$\sum_j MC_j(1 - Q_{r_2}) + \tau_1 > \sum_j MC_j(Q_{r_2}) \tag{122}$$

The convexity of MC_j described above, implies that the new equilibrium when the right hand side, equals the left hand side, is found by lowering Q_{r_1}, and increasing Q_{r_2}, that is, traffic is diverted from one route to another. As a consequence, the externalities become larger on the untolled route. It effectively means that the estimated reduction in total deadweight loss across the two links would be higher than what is the case, as a decrease in loss for one route, entails an increase in the loss for another route, when only one route is tolled.

F   Specification test

Here I conduct various specification checks on the equation used for estimation of externalities and elasticities.

F.1   Test of traffic demand model 2 and 3

To evaluate traffic model 2 and 3, I use an F-test to test whether or not the population variable should be included. I will also perform the likelihood ratio test, which is asymptotic, to assess the whether results differ across testing method.

I used a likelihood ratio test to assess whether or not the increase in likelihood value by dropping the variable of log population could, be seen as statistically significant. Let θ be the coefficient on log population. I then have $H_0 : \theta = 0$ vs $H_1 : \theta \neq 0$. The likelihood ratio test is chi-squared distributed with k degrees of freedom, where k is the number of restrictions.

The observed log-likelihood for model 3 was -184.70636, and for model 2 -181.6654. I obtained a p-value of .0032, suggesting that the reduction in the likelihood value is statistically significant at any conventional level. This provides evidence in favor of including the log of population in the specification.

Using an F-test, I tested whether or not log of population was significantly different from zero in at least one of the equations. My observed statistic was 4.86 with 2 restrictions and 261 degrees of freedom, corresponding to a p-value of 0.0085. This provides evidence in favor of including the log of population in the specification. I proceed by basing my simulation analysis on model 3.

F.2   Lagrange multiplier test of diagonal contemporaneous covariance matrix

Using the lagrange multiplier test outlined in the econometric chapter, I calculated the test that the contemporaneous covariance matrix is diagonal, meaning that residuals in the two equations are uncorrelated. In this case, simple OLS estimation on each equation is equal to the SUR-estimator. I have $H_0 : \Sigma \; diagonal$ vs $H_1 : \Sigma \; not \; diagonal$

For model 3, the observed test-statistic was 117.858, which is chi squared with two degrees of freedom, giving a p-value ≈ 0. For model 2, the observed test-statistic was 118.807, which is chi squared with two degrees of freedom, giving a p-value ≈ 0.

There is thus strong evidence in the data to suggest that the contemporaneous covariance matrix is not diagonal, and that the error terms in the two equations are correlated.

F.3   Testing log-log-specification

To assess whether the correct specification is the one I have chosen, I estimated model 3 using four different specifications: Log-log, lin-log, log-lin, and lin-lin. The estimated results are given in table 5. To assess model suitability, I calculated the Bayesian (Schwarz, 1978) and Aikaike (Akaike, 1974) information criterion for all models. The criteria are developed to aid model

selection, and the model with the lowest estimated value, is the preferred one. Based on the results, both the BIC and AIC, suggests the log-log specification is the most suitable.

F.4   FGLS estimation output

Below, the estimation results from estimation of the different elasticity models as in table 3, estimated by FGLS instead of the bootstrap procedure is given:

Table 6:
Estimation results from traffic demand model estimated by FGLS. Model numbers corresponds to models in section 7.1. See table 3 for details.

	(1) Model 1		(2) Model with AR		(3) Model 2		(4) Model 3	
Ln $AADT_L$								
Ln τ_L	-0.543***	(-6.36)	-40.17***	(-3.33)	-40.01***	(-3.46)	-33.24**	(-2.73)
Ln $\tau_L \times EV$	0.402***	(8.65)	0.388***	(8.52)	0.390***	(8.77)	0.384***	(8.84)
Ln $\tau_L \times RV$	0.220***	(3.45)	0.219***	(3.52)	0.239***	(3.93)	0.223***	(3.73)
Ln GC_L	-0.376***	(-6.84)	-0.338***	(-6.12)	-0.320***	(-5.96)	-0.316***	(-4.29)
Ln $Income$	4.367***	(4.10)	-5.636	(-1.75)	-6.395*	(-2.08)	-6.506*	(-2.07)
Ln $Income \times \tau_s$			3.155**	(3.28)	3.142***	(3.41)	2.605**	(2.69)
Alternative routes					-0.000507***	(-4.01)	-0.000401**	(-3.13)
Ln $Population$							0.147**	(2.99)
Constant	-43.44**	(-3.24)	82.09*	(2.03)	91.63*	(2.38)	91.31*	(2.32)
Ln $AADT_H$								
Ln τ_H	-0.375***	(-4.78)	-0.448***	(-5.56)	-0.451***	(-5.59)	-0.417***	(-5.16)
Ln $\tau_H \times RV$	0.221***	(4.12)	0.224***	(4.22)	0.230***	(4.31)	0.217***	(4.12)
Ln $\tau_H \times EV$	0.410***	(10.58)	0.404***	(10.54)	0.400***	(10.36)	0.394***	(10.32)
Ln GC_H	-0.0920***	(-4.42)	-0.0739***	(-3.44)	-0.0571*	(-2.51)	-0.0496	(-0.73)
Ln $Income$	3.904**	(3.30)	3.537**	(3.01)	2.610*	(2.31)	0.790	(0.63)
Alternative routes					-0.000495***	(-3.50)	-0.000388**	(-2.71)
Ln $Population$							0.138	(1.62)
Constant	-40.40**	(-2.73)	-35.73*	(-2.43)	-24.18	(-1.71)	-3.072	(-0.20)
N	156		156		141		141	

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

F.5   Testing main effects of road classification type

In the estimated demand equations, the main effects of road type, European and national road, was excluded. I conducted both an F-test and lagrange multiplier test to assess the suitability of this restriction(s).

The F-tests assesses whether or not at least one of the main effects coefficient is statistically significantly different from zero in any of the equations. I have $H_0 : \theta = 0$ vs $H_1 : \theta \neq 0$, where θ are the main effect's coefficients. The test has four restrictions and 261 degrees of freedom (from 2 equations). The observed value of the statistic is .7, with a p-value of 0.5906. This indicates

Table 5:

Specification tests of demand equation - choosing functional form.

	(1) Log-Log		(2) Log-Lin		(3) Lin-Lin		(4) Lin-log	
Log $AADT_S$								
Ln τ_L	-33.24**	(-2.73)					142091.0	(0.84)
Ln $\tau_L \times EV$	0.384***	(8.84)					4087.0***	(6.79)
Ln $\tau_L \times RV$	0.223***	(3.73)					1959.8*	(2.31)
Ln GC_L	-0.316***	(-4.29)	-0.333***	(-4.30)	-3658.6***	(-3.47)	-3290.7**	(-3.24)
Log $Income$	-6.506*	(-2.07)	1.387	(1.15)	8638.8	(0.54)	41688.8	(0.95)
Ln $Income \times \tau_s$	2.605**	(2.69)	-0.00260	(-0.32)	-21.52	(-0.20)	-11414.0	(-0.84)
Alternative routes	-0.000401**	(-3.13)	-0.000313*	(-2.25)	-0.863	(-0.46)	-1.419	(-0.78)
Log $Population$	0.147**	(2.99)	0.194***	(3.74)	2622.7***	(3.87)	2647.5***	(3.90)
τ_L			-0.0131***	(-3.61)	-39.88	(-0.79)		
$\tau_L \times RV$			0.00873*	(2.20)	84.77	(1.56)		
$\tau_L \times EV$			0.0331***	(6.10)	389.2***	(5.32)		
Constant	91.31*	(2.32)	-9.266	(-0.63)	-113816.6	(-0.58)	-530070.4	(-0.96)
Ln $AADT_H$								
Ln τ_H	-0.417***	(-5.16)						
$Ln\tau_H \times RV$	0.217***	(4.12)						
$Ln\tau_H \times RV$	0.394***	(10.32)						
Ln GC_H	-0.0496	(-0.73)	0.00919	(0.13)				
Ln $Income$	0.790	(0.63)	0.710	(0.49)				
Alternative routes	-0.000388**	(-2.71)	-0.000341*	(-2.09)				
Log $Population$	0.138	(1.62)	0.0689	(0.73)				
τ_H			-0.00320***	(-5.12)				
$\tau_H \times RV$			0.00240	(1.77)				
$\tau_H \times EV$			0.0171***	(6.83)				
Constant	-3.072	(-0.20)	-3.228	(-0.18)				
$AADT_H$								
τ_H					-1.525	(-1.79)		
$\tau_H \times RV$					2.327	(1.30)		
$\tau_H \times EV$					24.09***	(7.30)		
Ln GC_H					-82.79	(-0.84)	-91.29	(-0.95)
Ln $Income$					117.1	(0.06)	-145.1	(-0.08)
Alternative routes					-0.126	(-0.59)	-0.213	(-1.04)
Log $Population$					260.2*	(2.10)	300.8*	(2.51)
$Ln\tau_H$							-13.73	(-0.12)
$Ln\tau_H \times RV$							166.2*	(2.20)
$Ln\tau_H \times EV$							469.7***	(8.80)
Constant					-2323.7	(-0.10)	377.4	(0.02)
N	141		141		149		149	
R^2	0.594		0.524		0.332		0.364	
AIC	394.7		440.1		5452.0		5433.0	
BIC	444.9		490.2		5503.1		5484.1	

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

there is not enough evidence in the data to conclude that the main effects are different from zero in any of the two equations.

The likelihood ratio tests assesses whether the increase in the observed value of likelihood is statistically different from zero when including the main effects. The observed value of the statistic is 2.8, with 4 degrees of freedom, corresponding to a p-value of .594. This indicates there is not enough evidence in the data to conclude that the main effects are different from zero in any of the two equations.

F.6   Normality test - Log of AADT

The test developed by D'agostino et al. (1990), tests normality by combining tests for skewness and kurtosis. H_0 is that the skewness and kurtosis is equal to the values for a normal distribution, and H_1, that any of them are statistically signifcantly different from the values characterizing the normal distribution. The test output is given below. Both tests do not provide strong enough evidence to reject normality at any conventional significance level.

```
TEST OF NORMALITY LOG OF AADT HEAVY VEHICLES


                 Skewness/Kurtosis tests for Normality
                                                 ------- joint ------
    Variable |   Obs   Pr(Skewness)  Pr(Kurtosis)  adj chi2(2)   Prob>chi2
-------------+---------------------------------------------------------------
  lnaadtlv_h |   156      0.3971        0.5484         1.09         0.5792
```

```
TEST OF NORMALITY LOG OF AADT LIGHT VEHICLES


                 Skewness/Kurtosis tests for Normality
                                                 ------- joint ------
    Variable |   Obs   Pr(Skewness)  Pr(Kurtosis)  adj chi2(2)   Prob>chi2
-------------+---------------------------------------------------------------
    lnaadtlv |   165      0.1609        0.5462         2.36         0.3068
```

G   Details on speed and emissions models, including regression tables.

The results of estimations based on the emissions, speed and time models are given in table 3 below. To recap, these equations are

Emission factors:

$$g_{kj}(S) = \gamma_0 S^{\gamma_1}, j \in (NO_X, PM_{10}, CO_2), k \in (Light, heavy) \tag{123}$$

Road link speed as function of current PCE flow:

$$S(Q_r) = \nu_0[1 + \exp(\nu_1 Q_r)] \tag{124}$$

Required time to travel 1 kilometer as a function of current PCE flow:

$$T = \beta_0(1 + e^{\beta_1 Q_r}) \tag{125}$$

The estimation results are given below.

Table 7:

Regression table containing results for emissions, speed and time models. Emissions are grouped into three pollutants (CO_2, PM_{10}, NO_X), estimated separately for light and heavy vehicles. Speed and time models are estimated on the basis of PCU flow.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	$NO_{X,L}$	$NO_{2,L}$	$PM_{10,L}$	$CO_{2,L}$	$NO_{X,H}$	$NO_{2,H}$	$PM_{10,H}$	$CO_{2,H}$	T	S
Parameter										
$\nu_0/\gamma_0/\beta_0$	2.878	0.737	0.135	1115.0	74.59*	8.934**	1.095*	4056.6	0.0246***	43.24***
	(0.084)	(0.077)	(0.135)	(0.112)	(0.041)	(0.009)	(0.036)	(0.087)	(0.000)	(0.000)
Parameter										
$\nu_1/\gamma_1/\beta_1$	-0.489	-0.475	-0.582	-0.499	-0.662*	-0.683**	-0.645*	-0.448	0.00438***	-0.00510***
	(0.064)	(0.060)	(0.089)	(0.084)	(0.024)	(0.005)	(0.021)	(0.072)	(0.000)	(0.000)
N	3	3	3	3	3	3	3	3	9827	10571
R^2	0.999	0.999	0.998	0.998	1.000	1.000	1.000	0.999	0.998	0.996

p-values in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Please not that there are only 3 data points for the emissions factors. However, these are constructed by the Norwegian Institute for Air Research and the Norwegian Centre for Transport Research on the basis of many more data points (TØI, 2011b).

The number of observations are different for T and S. I only estimated part of the curve by a estimating the maximum flow at the road link. The point at which the curve becomes backward bending was declared on the basis of the time needed to travel the link. I used this as a qualifier for including observations in both the regression for T and S. Thus, there might be some difference in the points included.

Moreover, it is out of the scope of this thesis to evaluate the asymptotic identification of the functions fitted by non-linear least squares. That is, I assume there exists a global minimum of the objective function that is minimized. (see econometric theory section).

H   Simulating possible ranges of PCE elasticities

As my models have provided different answers to the question of elasticities with respect to tolls, I will need to introduce some further assumptions in order to provide some estimate of the average elasticity. I construct a simulation experiment to model a possible range of estimates, based on the evidence I have gathered so far. The experiment is not intended to find an answer to the question of the magnitude of the elasticities, but rather indicate some possible ranges that it may lie within, given the evidence I have gathered so far. Moreover, the features of the experiment's outcome are intended to provide some guideance when performing a sensitivity analysis with regards to elasticities.

To achieve this, I construct a model in which the elasticity of vehicle group $j = Light, Heavy$ is given by the following equation:

$$\varepsilon(U)_j = a_j + b_j U_j \tag{126}$$

Inserting the expression into the formula derived for average PCE elasticity gives:

$$\varepsilon(U) = \Lambda \left[\varepsilon(U)_H \pi + \varepsilon(U)_L (1 - \pi) \right] \tag{127}$$

where U is uniformly distributed in $U \in [0, 1]$. This assumption reflects that I have no a-priori knowledge of the distribution of the estimates. To provide a conservative suggestion, I thus use a uniform distribution where all outcomes are equally likely. The equation is calibrated for each user group such that $\varepsilon(1)_j$ is equal to a maximum value and $\varepsilon(0)_j$ to a minimum.

I will now use the elasticity with respect to tolls as an example. For light vehicles, I choose $\varepsilon(1)_L = -.22$ which is the estimate from Jones and Hervik (1992), and $\varepsilon(0)_L = -.05$, which is my lowest estimate. For heavy vehicles, model 1 predicted a positive elasticity. As this seems highly unreasonable, I choose $\varepsilon(0)_H = 0$ which is my lowest and sensible estimate. Further, I let $\varepsilon(1)_L = -.051$ which is my highest estimate. Using this information, simple algebra on a two-equation system with two unknowns gives me the following equations:

$$\varepsilon_H = -.051 U_H \tag{128}$$

$$\varepsilon_L = -.05 - .17 U_L \tag{129}$$

These two equations are combined in a simulation of ε, where 1000 draws from the uniform distribution is conducted for both U_H and U_L. Under the assumption of 15 % heavy vehicles,

the results are given in the histogram below, for both types of elasticities. I also provide a table with some summary statistics from the simulations.
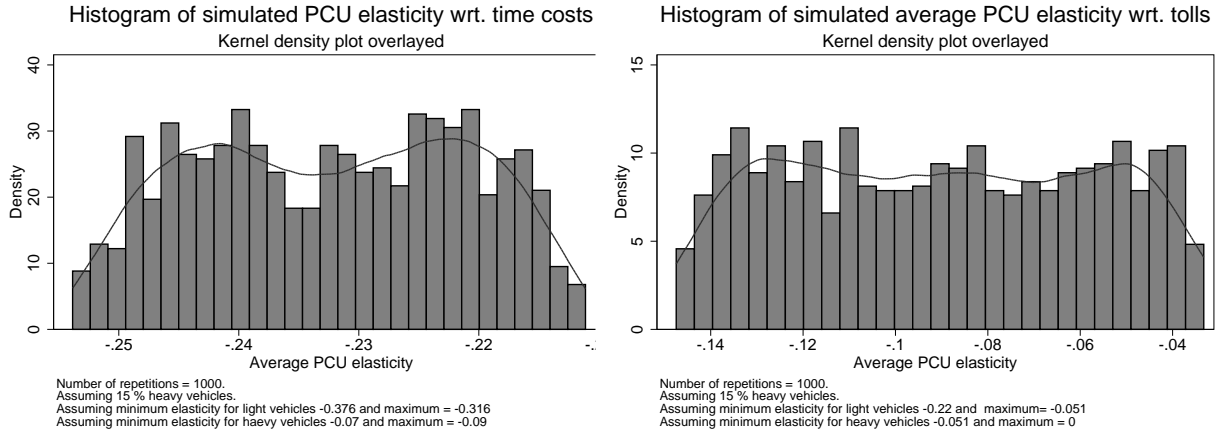
(a) Time cost elasticity.

(b) Toll elasticity.

Figure 12: Histogram of simulated elasticities. N = 1000 draws from a uniform distribution. Assuming $\lambda = .82, \alpha = 1.5, \pi = .15$. See section 6 "Data" for a description of these parameters.

Table 8:

Simulated elasticity of PCE flow with respect to time costs and congestion tolls. N = 1000 draws from a uniform distribution. Assuming $\lambda = .82, \alpha = 1.5, \pi = .15$. See section 6 "Data" for a description of these parameters.

	SIMULATION RESULTS				
	Mean	SD(ε)	Min	Max	count
$\varepsilon_{PCE-SIM}$	-.0912656	.0319861	-.1478978	-.0328515	1000
ε_{GC-SIM}	-.2324134	.0113784	-.2537215	-.2112068	1000
N	1000				

I   Sensitivity analysis of PCE flows and tolls with respect to elasticities

In this appendix, I provde different panels which shows the estimated optimal flow and congestion toll. I have recalculated the price and flow vector for different assumptions regarding the elasticity of PCE flow with respect to generalized costs and tolls. I have used the max, mean and min
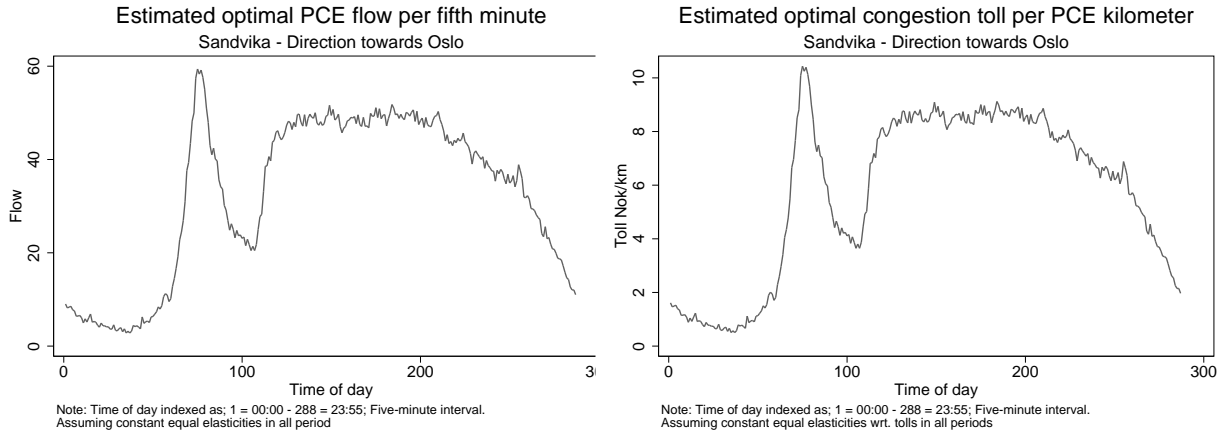
value from the simulation trial in appendix F. The panels indicate that results are fairly stable
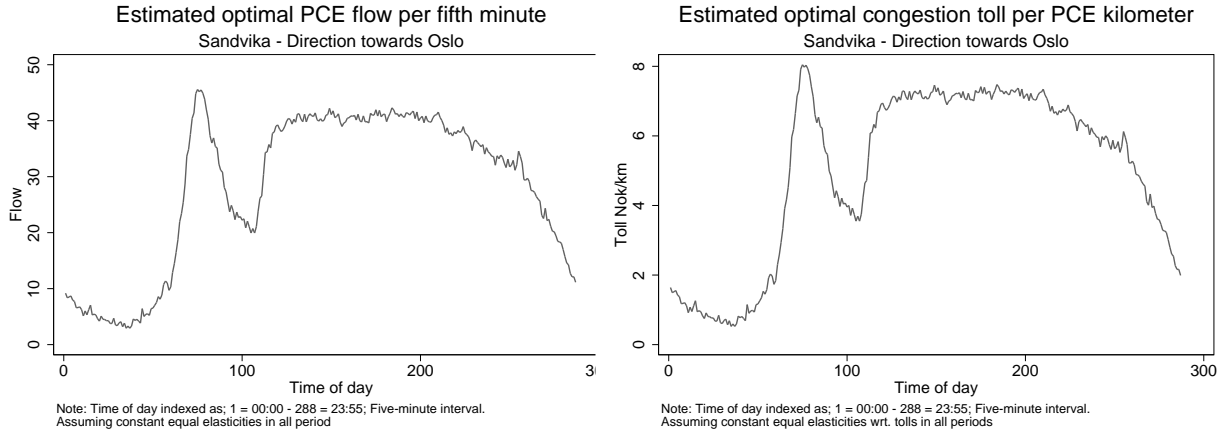
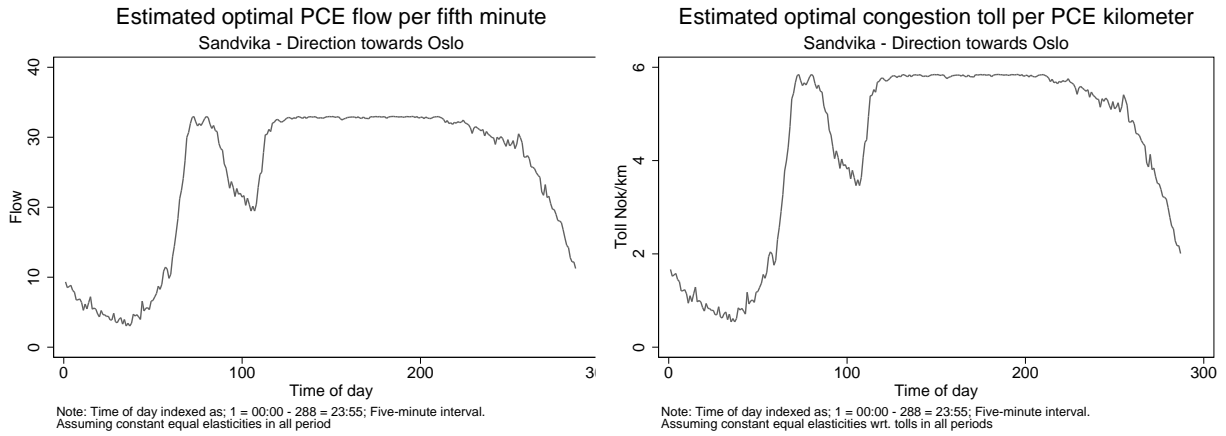

(a) Flow.

(b) Tolls per PCE kilometer (NOK).

Figure 13: Estimated optimal toll per PCE kilometer and flow. Assuming $\lambda = .82, \alpha = 1.5, \pi = .15$, and elasticities $\varepsilon_H = -.2324134$, $\varepsilon_L = -.0912656$. See section 6 "Data" for a description of these parameters.

(a) Flow.

(b) Tolls per PCE kilometer (NOK)

Figure 14: Estimated optimal toll per PCE kilometer and flow. Assuming $\lambda = .82, \alpha = 1.5, \pi = .15$, and elasticities $\varepsilon_H = -.2112068$, $\varepsilon_L = -.0328515$. See section 6 "Data" for a description of these parameters.

(a) Flow.

(b) Tolls per PCE kilometer (NOK)

Figure 15: Estimated optimal toll per PCE kilometer and flow. Assuming $\lambda = .82, \alpha = 1.5, \pi = .15$, and elasticities $\varepsilon_H = -.2537215, \varepsilon_L = -.1478978$. See section 6 "Data" for a description of these parameters.