



Railway Commuting Costs and Housing Prices in Norway: an Empirical Analysis

Haakon Gjersum and Per Stefan Præsterud

2017

MASTER THESIS

Department of Economics

Norwegian University of Science and Technology

Supervisor 1: Professor Kåre Johansen

Supervisor 2: Patrick Ranheim

Preface

This thesis is a part of our masters degree in economics at the Norwegian University of Science and Technology. The thesis is written in collaboration with The Norwegian Railway Directorate.

We would like to thank our supervisors, professor Kåre Johansen at the Department of Economics at NTNU and Patrick Ranheim, at the Norwegian Railway Directorate, for helpful inputs and guidance. We would also like to thank Wendy G. Bemrose for proofreading and Tora Uhrn for illustrations.

The master thesis is a joint work performed by Haakon Gjersum and Per Stefan Præsterud. Any mistake is our own and we take full responsibility for its content.

Trondheim, 2017-06-01

(Haakon Gjersum, Per Stefan Præsterud)

Abstract

In this thesis we present a new variable measuring the generalised time costs of commuting to and from central municipalities in Norway, from 1997 to 2016. We use this variable to estimate the effect of changes in railway transportation quality on housing prices in the central municipalities. The empirical analysis is performed by using a fixed effects regression model based on data for five of Norway's biggest municipalities; Oslo, Kristiansand, Stavanger, Bergen and Trondheim. We apply both static and dynamic models. We find that the generalised time costs of railway commuting has decreased in all municipalities between 1997 and 2016. Average commuting costs is highest in Bergen and lowest in Oslo. We find evidence of an inverse relationship between railway commuting quality to and from the central municipalities and housing prices within the municipalities. The short-run elasticity of housing prices with respect to changes in generalised time costs of commuting is about 0.2 per cent, while the long-run elasticity is about 0.4 per cent.

Keywords: housing prices; housing price dynamics; fixed effects; generalised time costs; transport economy; railway commuting.

Contents

Preface	ii
Abstract	iv
1 Introduction	1
2 Literature Review of Housing Prices	5
2.1 Modelling Housing Prices	5
2.2 Effects of Transport Quality on Housing Prices	6
3 Theory	9
3.1 Consumer Theory with Durable Goods	9
3.2 Aggregation and Total Demand	12
3.3 Housing Market Equilibrium	14
3.4 Generalised Costs	18
4 Data	19
4.1 Housing Prices	19
4.2 Gross Income	21
4.3 Unemployment Rate	23
4.4 Generalised Time Cost of Commuting	25
4.4.1 Converting NSB Time Tables	26
4.4.2 Trenklin III	27
4.4.3 Calculating The Average Generalised Time Cost for Each City	29
4.5 Descriptive Statistics	33

5 Empirical Specification	35
5.1 Baseline Model	35
5.2 Time Fixed Effects	37
5.3 Housing Market Dynamics	38
5.4 Serial Correlation and Driscoll and Kraay Standard Errors	39
5.5 Measurement Errors	40
6 Results	42
6.1 Static Housing Price Model	42
6.2 Housing Price Dynamics	44
6.3 Effects of Generalised Time Costs on 2016 Housing Prices	46
6.4 Alternative Functional Forms and Non-linearity	49
6.5 Summary of Results	54
7 Conclusion	56
Bibliography	60
A Additional Information	i
A.1 Estimation Gross Income 2016	i
A.2 The Process of Making the Generalised Time Costs Variable	i
A.2.1 Trenklin Transport Matrices	i
A.2.2 Intepretation of the Generalised Time Cost Matrix	v
A.3 Assumptions with Fixed Effects and Random Effects Models	vi
A.4 Unit Root Test	vii
A.4.1 Log Housing Prices	viii
A.4.2 Log Unemployment	ix
A.4.3 Log Gross Income	ix
A.4.4 Log PGC2	x
A.5 Durbin’s Test for Serial Correlation With Strictly Exogenous Regressors	x
A.6 Durbin’s Test for Serial Correlation Without Strictly Exogenous Regressors	xi
A.7 Testing Hypotheses About a Single Population Parameter: The t-test	xiii

List of Figures

- 3.1 Housing Market Equilibrium 15
- 3.2 Generalised Time Costs and Population 16
- 3.3 Housing Demand Shift 17

- 4.1 Development in Housing Prices 1992-2016 20
- 4.2 Gross Income 22
- 4.3 Development in Unemployment Rate 1999-2016 24
- 4.4 PGC Flow Chart 26
- 4.5 Day-distribution of Travels Used in Thesis 29
- 4.6 Central Municipalities and Commuting Radiuses 30
- 4.7 PGC 31

- 6.1 Observed Change in Housing Prices and Share Explained by PGC2 48
- 6.2 Estimated Relationship Between PGC2 and IHP in Model(3) 51
- 6.3 Estimated Relationship Between PGC2 and IHP in Model(4) 52

- A.1 Time Cost of Waiting iii

List of Tables

4.1	Housing Prices	21
4.2	Average Gross Income	22
4.3	Unemployment Rate	25
4.4	Timetable Entries	27
4.5	Descriptive Statistics PGC2(minutes)	32
4.6	Descriptive Statistics PGC1(minutes)	32
4.7	Descriptive Statistics	33
6.1	Fixed Effects Estimates With Driscoll and Kraay Robust Standard Errors	43
6.2	Change in 2016 Housing Prices With a 1 Percent Increase in PGC2, 95%-Confidence Interval	44
6.3	Fixed Effects Estimates with Driscoll and Kraay Robust Standard Errors Dynamic Model	45
6.4	Summary of the Effects on Housing Prices with Lower Transport Quality	46
6.5	Housing Price and PGC2 Variation	47
6.6	Effect of one Standard Deviation Railway Quality Improvement	47
6.7	Dynamic Model Estimation with Non-linearity	50
6.8	Comprehensive Model	53
6.9	Test of Nonnested Models	54
A.1	Average Waiting Time on Station Oslo S-Fredrikstad 2016	iii
A.2	Average On Board Time Oslo S-Fredrikstad 2016	iv
A.3	Average Transfers Oslo S-Fredrikstad 2016	iv

A.4 Average Waiting Time on Transfers Oslo S-Fredrikstad 2016 v

A.5 Generalized Costs Oslo S-Fredrikstad v

A.6 LLC Unit Root Test for IHp viii

A.7 LLC Unit Root Test for lunem ix

A.8 LLC Unit Root Test for IY ix

A.9 LLC Unit Root Test for IPGC2 x

A.10 AR(1) Serial Correlation Test xi

Chapter 1

Introduction

In this thesis we pose the question; Does improvements in railway quality for commuters diminish the housing price inflation in central municipalities in Norway? We aim to estimate a relationship between housing prices and variations in the time costs of travelling, also known as generalised time costs. The mechanism behind our hypothesis is the following; better railway quality for work commuters will increase the attractiveness of living in more peripheral areas relative to the central municipality, and thereby reducing its population. All else equal, a lower population in central city areas implies lower aggregate demand for housing and less pressure on housing prices.

The Norwegian housing market is an interesting subject with a broad specter of analyses to understand its developments. From [Anundsen and Jansen \(2011\)](#) estimating the interaction between housing prices and household borrowing, to [Fiva and Kirkebøen \(2011\)](#) which looks at housing market reactions to published school quality reports. [Eide \(2015\)](#) investigates the connection between taxation and housing prices. All in all, there is a lot of focus on this area.

As the data will show, all major cities in Norway have experienced substantial rising housing prices over the past two decades. The development of housing prices in the major cities is also evident in the rest of Scandinavia. According to [Norges Bank \(2016\)](#), all Scandinavian countries have experienced a larger increase in housing prices in their major cities compared to the respective country overall. After the global financial crisis in 2008 and the following drop in the oil price in 2014, Norway has also experienced a rise in unemployment and lower investments. These two effects, coupled with a rising housing market and credit growth, has presented policy makers with a dilemma between financial stability and counter-cyclical monetary policy. This makes the

developments and drivers in the housing market an important issue for analysis.

In the early 2000s, Norway lagged behind other major cities on the European continent in terms of centralisation and migration trends. [Østby \(2001\)](#) discusses how Norway was still characterised by a great deal of urbanisation, and points out that international studies found Norway to differ from other European cities, where centralisation declined in the 1990's. [Brunborg \(2009\)](#) states that such migration trends in Norway continued towards 2009 and that the population share living in the most central municipalities, has increased from 61 to 67 percent from 1980 to 2009. It is reasonable to believe that these migration trends have contributed to the rising housing prices seen in the major cities in Norway. Around central cities and municipalities, people balance the decision between migration and commuting. Improving the railway quality for commuters to and from the central municipalities can enhance commuting as an alternative to migration, and reduce demand and housing prices in the central municipalities. [Holvad and Preston \(2005\)](#) finds that reduced commuting costs, caused by improved transportation infrastructure, can influence commuting and migration decisions. Commuters can travel longer stretches for the same amount of generalised time costs, leading to higher migration to areas with lower housing prices. They also argue that inflating housing prices in a region, which is a net importer of labor, could give commuting rather than migration to the district.

Our contribution to this subject is the construction of a generalised time cost variable for railway commuting. The variable illustrates the development in railway quality in Norway for the past 20 years. We will use this variable to explain the link between the costs of commuting to and from central municipalities and housing prices in these municipalities. The generalised time cost variable is generated based on railway quality components between the central stations in central municipalities in Norway, and stations outside the cities' municipality border. To isolate the railway quality improvements relevant for commuting to and from the municipality, we exclude the stations within the municipality border.

In a world with a continually growing population and limited land resources, urban structure and land use will become increasingly important in the future. In the National Transport Plan (NTP) for 2018-2029¹, the Norwegian government assumes a population growth of about 1 percent

¹See [Norwegian Railway Directorate, Norwegian Coastal Administration and Norwegian Public Roads Administration \(2017\)](#)

per year from 2016 to 2022. This implies a population of over 5.5 million people in 2022. The transport plan outlines how the Government intends to prioritise the transport sector over a twelve year period. One of the objectives in the NTP is the "Zero growth objective", this implies that, excluding cycling and walking, the growth in passenger transport in the cities is to be absorbed by public transportation. The Norwegian Ministry of Transport and Communications states that²:

"Together with other public transportation options, the train will contribute to improved accessibility in the metropolitan areas. The development of modern rail infrastructure between cities and towns provides reduced travel times, frequent and regular departures and better punctuality. This means that people can choose to live and work where they want and larger areas are linked to a common housing and labor market".

The statement underlines the relevance of future railway commuting. This leads the way for an interesting analysis of the link between railway transportation quality for commuters and aggregate housing demand and housing prices in the central municipalities. In this thesis we investigate the link empirically, using five central municipalities in Norway; Oslo, Kristiansand, Stavanger, Bergen and Trondheim.

The structure of the thesis is: Chapter 2 will provide relevant earlier literature and chapter 3 presents economic theory. In chapter 4 we describe the data and the process of calculating the generalised time costs variable. Chapter 5 explains the empirical specification and chapter 6 presents the empirical results. Chapter 7 concludes.

²See [Transport Communications Ministry \(2015\)](#)

Chapter 2

Literature Review of Housing Prices

The development in housing prices has been a popular subject for theoretical and empirical analysis over the past few years. In our research for this thesis we have encountered numerous national and international papers trying to explain the movements in housing prices. In this section we will review some of the literature regarding housing prices with respect to variables and methods of interest.

2.1 Modelling Housing Prices

In this thesis we use unemployment and income as control variables when estimating the effect of commuting costs on housing prices. Income and unemployment affect the attractability of the municipalities and affect aggregate housing demand. [Arestis and Gonzales \(2013\)](#) from the Levy Economics Institute models the housing price developments in 18 OECD countries from 1970 to 2011 using a vector error correction model (VECM). A VECM is a time-series model used for data with a cointegrating relationship between the underlying variables. The authors point to disposable income as the cornerstone of their model. In addition to disposable income, they model housing prices as a function of residential investment, banking credit, mortgage rate, taxation, unemployment and population. Their analysis focuses on both a long-run relationship as well as short-run dynamics, keeping housing supply fixed in the short run. Their empirical findings for the short run is that previous levels of real housing prices plays a significant role in explaining the development of housing prices in Norway. This is, as underlined by Asteris and Gonzales,

an effect of the role of expectations and speculation in the housing market. Additional findings for Norway include that unemployment is not relevant in explaining the short-run dynamics, but population is. Our analysis will build upon some of the same variables as Asteris and Gonzales. We will use income and unemployment as control variables to explain migration decisions and changes in aggregate housing demand in each municipality. Macro economic variables that vary over time, such as interest rates, also affect housing demand. Because all our entities are cities within the same country, these variables do not vary between entities and can be controlled for with econometric techniques.

2.2 Effects of Transport Quality on Housing Prices

[Kulish et al. \(2012\)](#) investigates the development of urban structure and housing prices in Australia. They apply a basic model of Alonso(1964), Muth(1969) and Mills(1967), which assumes a city with fixed population and a given income level. The population lives around a central business district(CBD) and travel into the city center for work. Since commuting is costly and increases with the distance to the CBD, households would choose to live closer to the city center. The model assumes, as we do in this thesis, that households are identical in preferences. Their empirical analysis points to the potential significance of transport infrastructure. This includes public transport infrastructure as well as infrastructure for cars. The authors calibrate transport costs a monetary cost as well as a time cost. The time cost is valued to 60 percent of the wage rate. The remainder is assumed to be spent on housing and other goods. This means that higher income leads to higher time costs. They find that in cities with better transport infrastructure, it is more feasible to live further away from the city where housing prices are lower, because of lower commuting costs. This supports the theory that improvement in transportation infrastructure to central business areas can reduce aggregate housing demand in the city. In relevance to our thesis, this would imply that better railway infrastructure, to for example Trondheim from Stjørdal, could reduce housing price pressure in Trondheim city. Lower commuting costs on the railway line from Stjørdal to Trondheim could make it more feasible to live in Stjørdal, 33 kilometers north-east of Trondheim.

[Grue et al. \(1997\)](#) uses the log of housing prices in a hedonic pricing model for houses in three regions around Oslo city centre. These regions are Oslo West, Grorudalen-Østensjø and

North-South. A hedonic model treats housing as a heterogeneous good where its value depends on different attributes of the houses (i.e. floor-space, age, number of rooms etc.). In our analysis we treat housing as a homogenous good and use variables affecting aggregate demand to model changes in housing prices. In [Grue et al. \(1997\)](#), the main focus of the analysis is on the exposure to road traffic as measured by outdoor noise levels and location in terms of distance from city center. The contributions of a number of additional variables are also estimated. One of these is the measure of the public transportation quality in form of generalised costs per kilometer distance to the city center. Their empirical analysis is based on two separate datasets. One from 1995 with condominium apartments and a second from 1988-1995 with flats and houses. The authors find a negative effect of higher generalised cost of public transportation per kilometer distance to the Oslo city center. This effect was only significant in the first dataset for condominium apartments. This means that houses outside the city center are negatively affected by higher generalised costs of travelling in to the city center.

To analyse households residential and job location choices [So et al. \(2001\)](#) applies a logit model using commuting costs, wages and housing prices in Iowa with 6214 household entries. In their analysis of binary outcomes, a logit model serves as a convenient tool to predict the probability of commuting decisions. Households are either in what is defined as the metropolitan area or the nonmetropolitan area. Each with its own labor market. This leaves four possible residential and work location combinations:

- live and work in the metropolitan area
- live in metropolitan area and commute to nonmetropolitan area
- live and work in the nonmetropolitan area
- live in nonmetropolitan area and commute to metropolitan area

The authors find that the probability of commuting decreases with commuting costs. A 10 percent increase in commuting time between metropolitan and nonmetropolitan areas, reduces the proportion of commuters across markets by 17 to 19 percent evaluated at their sample means. This implies that commuting costs affects the households' migration decisions when balancing between commuting or migrating. If commuting costs decrease, households will find it more feasible to live

in a different area than where they work and rather commute to and from work. [So et al. \(2001\)](#) also finds that the effect of commuting costs on the proportion of commuters is bigger between the labor markets than within the markets.

A characteristic of the relevant literature is the focus on the effects in the peripheral areas. The effects on housing markets inside the central areas has received less attention. [Grue et al. \(1997\)](#) uses generalised costs per kilometer distance to the city center in Oslo as an explanatory variable for housing prices around the city center. There is also a lot of literature on the effects on housing prices in close proximity to new investments in physical railway infrastructure¹. Railway transportation quality in different areas depends not only on the physical infrastructure, but also on changes in how the available infrastructure is utilised.

In this thesis we attempt to explain housing demand within Norwegian central municipalities using a variable for the generalised time cost of commuting, varying from year to year and between entities. There are, as far as we know, no such previous studies.

In the next chapter we present the relevant theory to explain and justify our stated hypothesis. The theory is relevant to explain how commuting costs and migration decisions affects the housing market.

¹See for example [Pagliara and Papa \(2011\)](#) and [Geng et al. \(2015\)](#)

Chapter 3

Theory

In this section we will present relevant economic theory. To substantiate our hypothesis, we present both basic consumer theory of the demand for durable goods and urban economic theory. The consumer theory will establish a framework on demand for housing, and help us understand how people make their decisions when investing in property. To further investigate how generalised time costs of travelling affect migration decisions, we will clarify parts of urban economic theory as well as theory relevant to calculate costs of commuting.

3.1 Consumer Theory with Durable Goods

Durable goods, such as houses, is an important part of total economic production. The nature of durable goods is that they serve the consumer over a period of time. [Rødseth \(1985\)](#) argues that with respect to durable goods such as housing, cars etc. the argument for the utility function is in the inventory of the durable good, not the transaction. Rødseth uses the following two-period example to illustrate how the demand for durable goods depends on prices, income and the implicit user cost of the durable good. He starts out with a general discrete utility function with durable and non-durable goods.

$$U = u(c_1, c_2, k_1, k_2) \tag{3.1}$$

Where c_1 and c_2 is consumption of the non-durable in period one and two. k_1 and k_2 is the inventory of the durable good in period one and two. The consumers maximises utility over the two periods subject to a budget constraint. The inventory of the durable good increases with new investments and decreases with depreciation.

$$k_t = j_{t-1} + (1 - \delta)k_{t-1} \quad (3.2)$$

$t = 0, 1, 2, 3$

Where k_t is the inventory level at time t . j_t is new investments and δ is the depreciation rate. The consumer can only buy the durable goods at the end of each period and obtains utility over the following period. The first transaction of the durable good occurs at the end of period zero and all inventories are sold at the end of the last period so that $k_0 = 0$ and $k_3 = 0$.

The consumer finances the investment in the durable good by borrowing. This means that the value of the first purchase at the end of period zero plus interest equals the consumer's debt in period one.

$$\alpha_1 = q_0 j_0 (1 + i_1)$$

With q_t being the price of the durable good. After period 1 the consumer's debt inclusive interest is:

$$\alpha_2 = (p_1 c_1 - y_1 + \alpha_1 + q_1 j_1)(1 + i_2)$$

It decreases if the value of consumption in period 1 is less than income and increases with new investments in period 1.

At the end of period 2 the consumer must settle all his debt and, to maximise utility, spend all capital on consumption. This assumes no utility from bequeathing. This gives us the following budget constraint:

$$y_2 - (p_1 c_1 - y_1 + q_0 j_0 (1 + i_1) + q_1 j_1)(1 + i_2) - p_2 c_2 - q_2 j_2 = 0$$

Using (3.2) solved for the new investment j_t we have:

$$y_1 + \frac{y_2}{1 + i_2} = p_1 c_1 + [q_0(1 + i_1) - q_1(1 - \delta)]k_1 + (p_2 c_2 + [q_1(1 + i_2) - q_2(1 - \delta)]k_2) \frac{1}{(1 + i_2)} \quad (3.3)$$

From (3.3) it becomes evident that the consumer in effect is renting the durable good each period for a given price. The term $q_{t-1}(1+i) - q_t(1-\delta)$ is the implicit user cost of the durable good. Equation (3.3) tells us that the present value of consumption and renting the durable good for both periods must be equal to the present value of the sum of income. To simplify we can define the user cost for period t as s_t , so that:

$$p_1c_1 + s_1k_1 + (p_2c_2 + s_2k_2)\frac{1}{(1+i_2)} = y_1 + y_2\frac{1}{(1+i_2)} \quad (3.4)$$

The user cost of the durable good is in effect the real interest rate and depreciation rate multiplied by the price of purchasing the good. Reformulating the equation for the user cost, we have:

$$s_t = q_t\left[(1+i_t)\frac{q_{t-1}}{q_t} - (1-\delta)\right]$$

Defining the real interest rate expressed in the durable good, r_t , as $1+r_t = (1+i_t)\frac{q_{t-1}}{q_t}$ we have:

$$s_t = q_t(r_t + \delta)$$

In this two period example, the consumer chooses the levels of durable and nondurable goods, in order to to maximise utility (3.1) with respect to the intertemporal budget constraint (3.4). Maximising gives the following marginal conditions:

$$u'(k_1) - \lambda s_1 = 0 \quad (3.5)$$

$$u'(k_2) - \lambda \frac{s_2}{(1+i_2)} = 0 \quad (3.6)$$

$$u'(c_1) - \lambda p_1 = 0 \quad (3.7)$$

$$u'(c_2) - \lambda \frac{p_2}{(1+i_2)} = 0 \quad (3.8)$$

(3.5) - (3.8) together with the budget constraint gives 5 equations defining the five unknowns; c_1, c_2, k_1, k_2 and λ . Eliminating the Lagrange parameter we have 4 equations which defines optimal

consumption of each good in each period as a function of the exogenous variables:

$$c_1^* = f(p_1, p_2, i_2, s_1, s_2, y_1, y_2) \quad (3.9)$$

$$c_2^* = g(p_1, p_2, i_2, s_1, s_2, y_1, y_2) \quad (3.10)$$

$$k_1^* = h(p_1, p_2, i_2, s_1, s_2, y_1, y_2) \quad (3.11)$$

$$k_2^* = i(p_1, p_2, i_2, s_1, s_2, y_1, y_2) \quad (3.12)$$

Equation (3.11) and (3.12) defines demand for the durable good for the individual, or in relevance to our thesis, individual housing demand. Housing demand for the individual in each period depends on price level on non-durable consumption, interest rate, the user cost of housing and income. Without specifying the demand function, it is reasonable to expect that higher prices on non-durable goods reduces disposable income available for housing investments. Higher income levels increases disposable income for housing and increases demand, while higher user cost of housing decreases the demand for housing. For a given set of the exogenous variables and consumer preferences, total demand in a specific municipality will depend on the sum of the individuals' demand and hence variations in the population.

3.2 Aggregation and Total Demand

Rewriting the equations (3.11) and (3.12) on general form, we can write:

k_{it} = Demand for individual i in period t . The sum of all consumer's demand gives us total demand for housing in a specific municipality:

$$k_t^{total} = \sum_{i=1}^N k_{it}$$

Where N is the municipality's population. As explained in the last section, variations in N will lead to variations in the each municipality's total demand for housing. If we assume that consumers are

identical in their preferences and can be represented by a representative consumer we have:

$$k_{it} = k_t$$

and total demand becomes:

$$k_t^{total} = N * k_t$$

When people move to and from the municipality, the population changes and total demand for housing shifts. In this thesis we consider 5 cities: Oslo, Kristiansand, Stavanger, Bergen and Trondheim. These represent the central business municipalities relevant for railway commuting in each region; East, South, South-West, West and Middle Norway respectively¹. The size of the N in these cities, and implicitly total demand for housing, depends on the attractiveness of living in the city, which in turn depends on economic factors and city specific attributes. Such economic factors include income and unemployment rate and city specific attributes include factors such as location and climate. In addition, we conjecture that peoples migration decisions, and implicitly total demand for housing in the cities, depends on the generalised time cost of commuting to the cities from more rural areas.

Following the framework of Alonso(1964), Brueckner (1987) argues that housing prices near the city center will increase, if the commuting cost increases. Commuters will want to move toward the city center to reduce their commuting costs. This movement will cause an excess in demand for housing. This excess in demand bids up housing prices near the city center, and reduces them at suburban locations. Let \bar{x} denote the distance from the city center to the urban-rural boundary. At this urban boundry, urban land rent r , equals the agricultural rent r_A . Brueckner shows that $\frac{dr}{dx} < 0$, so that urban rent will exceed r_A inside \bar{x} , and fall short of r_A beyond \bar{x} . If r exceeds r_A , there will be housing production. Brueckner defines y as the income of a representative household, t as commuting costs and u as each households utility. At \bar{x} the land rent equals the agricultural rent, more formally

$$r(\bar{x}, y, t, u) = r_A \tag{3.13}$$

In equilibrium, the urban population exactly fit inside the city limit \bar{x} . To formalize this condition,

¹We consider the basis of railway commuting to be too small in regions North of Trondheim

let θ equal the number of radians of land available for housing at each distance x inside the city boundary, with $0 < \theta < 2\pi$. The population of a narrow ring with inner radius x and width dx , will approximately equal $\theta x D(x, y, t, u) dx$. The condition that the urban population N fit inside \bar{x} may then be written

$$\int_0^{\bar{x}} \theta x D(x, t, y, u) dx = N \quad (3.14)$$

The equation says that the total demand for housing within the city circle, equals the total population. The term $\theta x D(x, t, y, u)$ is the amount of available land within the circle, multiplied with the demand for housing. Brueckner defines p as the housing price and shows that $\frac{dp}{dN} = \frac{dp}{du} \times \frac{du}{dN} > 0$. If the population increases inside \bar{x} , the demand increases for a given amount of land and the prices will increase. That is, if the generalised time cost of commuting from rural areas to the city increases, people will consider housing outside the city as relatively less attractive and the size of N in the city increases and bids up the equilibrium housing price.

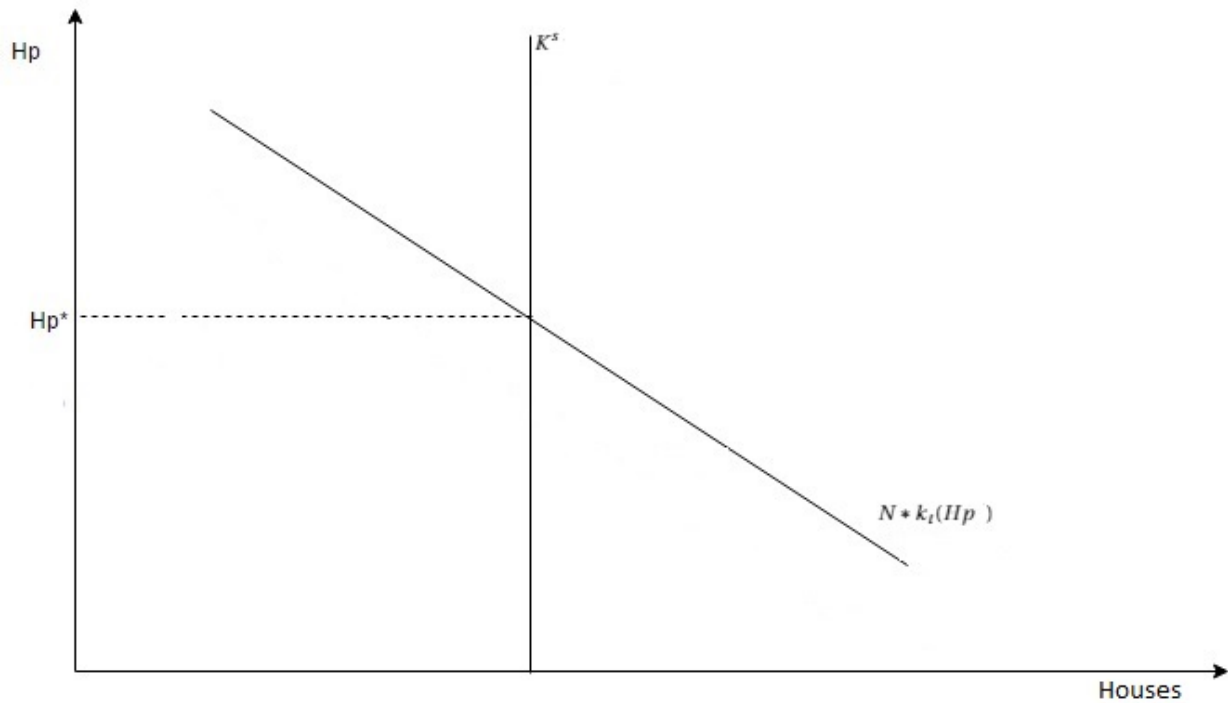
3.3 Housing Market Equilibrium

In equilibrium, total housing demand equals total housing supply and thereby defines the equilibrium housing price. The supply of housing will depend on factors such as price on input factors, wages and technology. Since building houses takes time, it can be assumed to be fixed in the short-run. Denoting the fixed housing supply as K^s we have equilibrium in the housing market when:

$$k^s = N * k_t(Hp) \quad (3.15)$$

Individual housing demand is as we know from (3.11) and (3.12) a function of the housing price denoted here as H_p . The equilibrium is illustrated in figure 3.1:

Figure 3.1: Housing Market Equilibrium

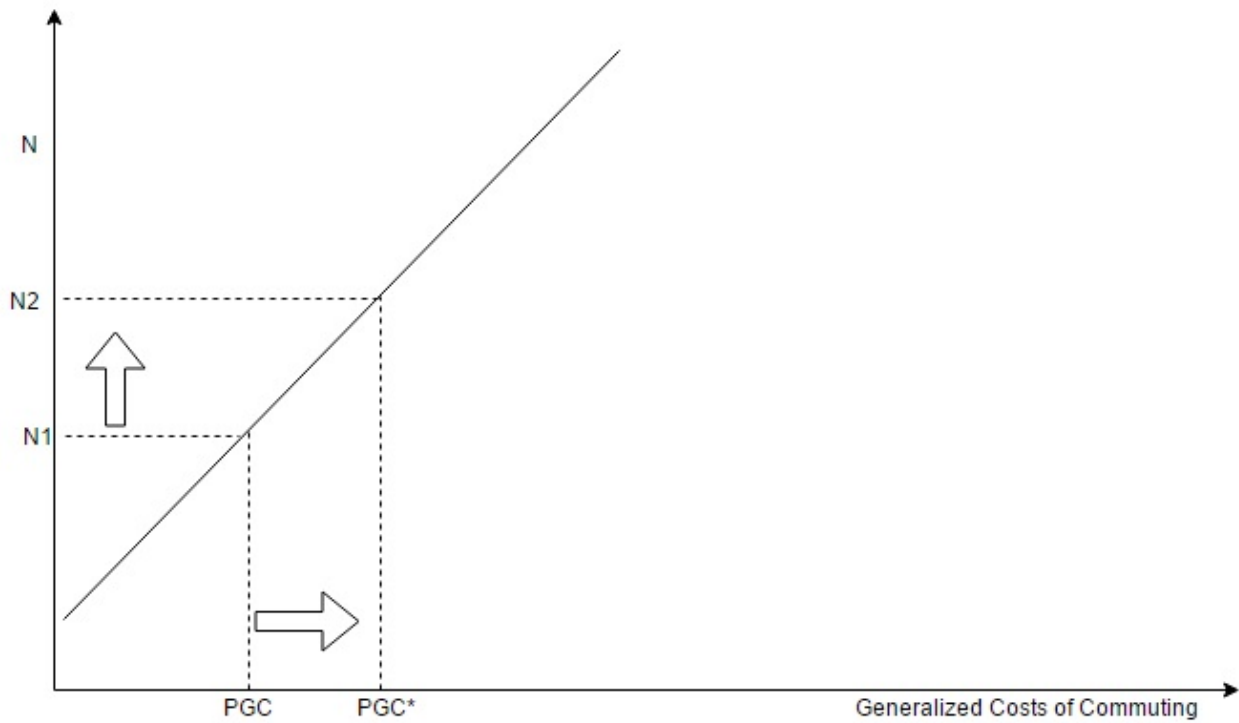


Using the market equilibrium in (3.15) we can solve for the housing price as a function of fixed total supply and total demand:

$$Hp^* = f(k^s, N * k_t) \quad (3.16)$$

Where higher population, an increase in N , shifts the demand curve up and results in a higher equilibrium housing price; $\frac{\delta Hp^*}{\delta N} > 0$. When we conjecture that the population in the city increases when commuting to and from the city is more costly, we have $\frac{\delta N}{\delta PGC} > 0$, where PGC is the private generalised time costs of commuting. The relationship between commuting cost and population in the city is illustrated in figure 3.2:

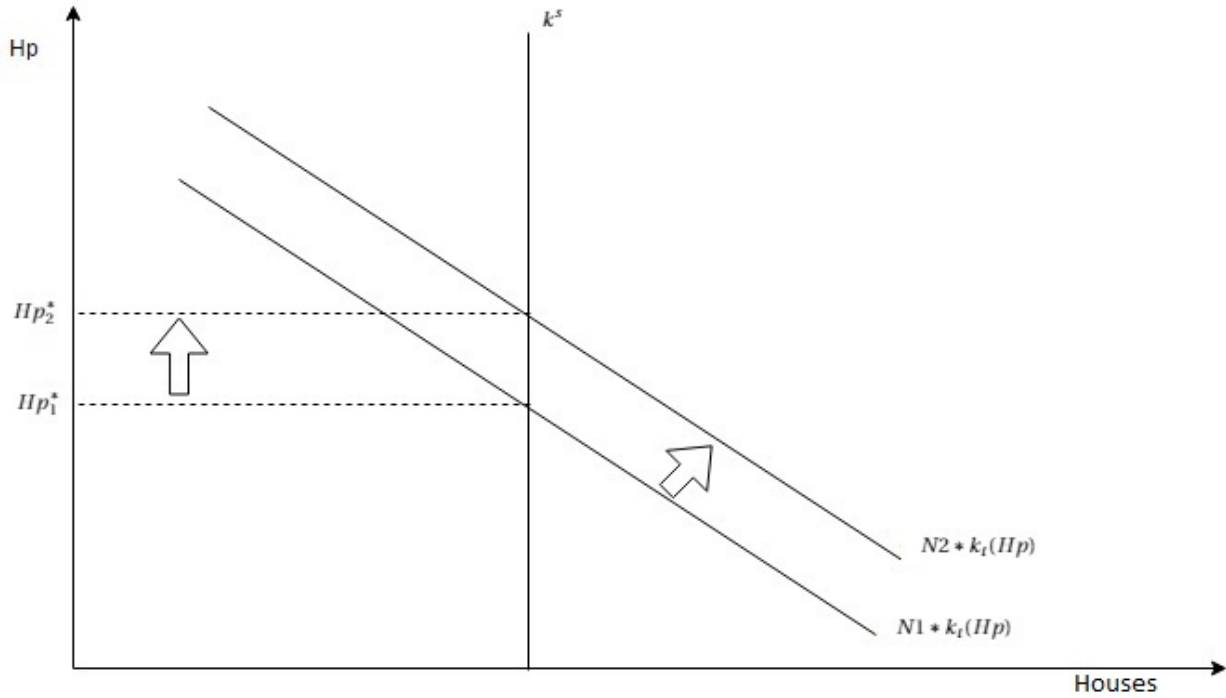
Figure 3.2: Generalised Time Costs and Population



The population in the city increases from $N1$ to $N2$ as less people want to live in the rural areas because of the higher generalised time cost of commuting to the city, PGC to PGC^* .

This again shifts the total demand curve and we get a higher equilibrium housing price:

Figure 3.3: Housing Demand Shift



The increase in the population from $N1$ to $N2$ shifts total demand for housing from $N1 * k_t(Hp)$ to $N2 * k_t(Hp)$ and increases the equilibrium price for housing from Hp_1^* to Hp_2^* . This means that the equilibrium housing price can be written as a function of variables that affect the population size. These variables include the cities' unemployment rate, income level and the commuting cost to and from the city. In addition, variables that are fixed over time but varies between the cities such as location, climate, governance etc. have effects on housing prices and should be included. The demand function on inverse form is written below:

$$Hp_{it} = f(\text{unemployment}_{it}^-, \text{income}_{it}^+, PGC_{it}^+, a_i, c_t) \quad (3.17)$$

Where a_i denotes the city specific fixed effects and c_t represents common time-varying macrovariables. As higher unemployment reduces the possibility of obtaining a job in the city, unemployment rate is expected to have a negative impact on housing prices. Higher income level increases the cities' attractiveness and is expected to have a positive relationship with housing prices. The implication of our hypothesis is that we expect commuting costs to and from the city to have a positive relationship with housing prices.

3.4 Generalised Costs

The theory above shows how the commuting cost to and from the city is expected to affect the city's housing prices. In order to measure the cost of commuting it is important to remember that the total cost of commuting is not only the monetary cost of the travel, but also all other time consuming aspects of the travel. These time consuming aspects could have been used to something else if commuting was not needed. The sum of the monetary costs and these alternative costs are what is known as total generalised costs. [Grøvdal and Hjelle \(1995\)](#) gives the following formulation;

$$PGC = \sum_{i=1}^h q_i v_i + \sum_{j=1}^k w_j T_j^p \quad (3.18)$$

Where PGC is the private generalised costs, q_i is the price of a input factor (such as the ticketprice in the case of public transport or price of petrol in the case of private transport), v_i is the amount of input i , w_j is the time value of time component j and T_j^p is the amount of time component j . Both v and T will vary with aspects of the given travel, such as the total length of the travel, total volume of the travel and other factors related to the way the travel is performed. The equation takes into account that different time components (for example time waiting on transit stations, and time spent getting to the shuttle service) has different value. [Grøvdal and Hjelle \(1995\)](#) states that there is empirical evidence to use a higher time value on waiting time, than time spent on the actual travel.

In this thesis we will construct a generalised time costs variable in the same framework as [Grøvdal and Hjelle](#) based on time tables from Norwegian State Railways. This procedure is specified in chapter 4, which also includes detailed information of each variable used in the analysis.

Chapter 4

Data

In this section we will explain and illustrate each variable in our analysis. The analysis is based on a balanced panel data from 1997 to 2016 for Oslo, Kristiansand, Stavanger, Bergen and Trondheim. Unemployment data is obtained from The Norwegian Labour and Welfare Administration (NAV) with data from 1996 to 2016. Income data and housing data is collected from Statistics Norway (SSB) and ranges from 1993 to 2016 and 1992 to 2016 for income and housing respectively. The self-generated generalised time costs variable, PGC, is based on raw data collected from Norwegian State Railways(NSB) and The Norwegian Railway Directorate(previously Norwegian National Rail Administration). The data ranges from 1997 to 2016. The years 1997-2016 is also the time period where The Norwegian National Rail Administration were responsible for maintaining, owning, operating and developing the national railway network¹.

4.1 Housing Prices

Our housing price data is based on the sum of houses sold and the total value of houses sold. This gives us the average market housing price in each city. The data includes data for houses, town-houses, detached, semi-detached and apartments. The composition of the housing types vary over time so that the data does not illustrate the price development of specific types of housing. Properties without a stated purchasing price are not included. Analysing the data, we see that Norway

¹Norwegian National Rail Administration were responsible for operating and maintaining the Norwegian railway network from 1.12.1996 to 31.12.2016, before the agency was divided into The Norwegian Railway Directorate and Bane NOR

is no exception to the rising housing market seen elsewhere in the Nordic countries during the last couple of decades, (Norges Bank (2016)). Figure 4.1 shows the development in housing prices for our dataset. As expected, the global financial crisis which had most financial assets plummet, also had a significant impact on housing prices for all cities in our dataset. The years following the financial crisis had a continuous rise in housing prices with the exception of Trondheim in 2016 and Stavanger in 2015 and 2016. The most astounding development has been in the Capital Oslo, especially from 2011 to present.

In Stavanger, housing prices dropped by almost 6 percent from 2014-2016, after an average increase of over 8 percent per annum since 2010. This must be seen in light of the development in oil prices during this period. The price on one barrel of Brent Crude oil dropped by around 35 percent in 2014 from 109 USD to 70, before adjusting for currency depreciation(Norges Bank (2014)). According to the county-level national accounts for 2014, the oil and gas industry including services, contributed to one fifth of Rogaland's gross production at the time of the fall in the oil price.

Figure 4.1: Development in Housing Prices 1992-2016

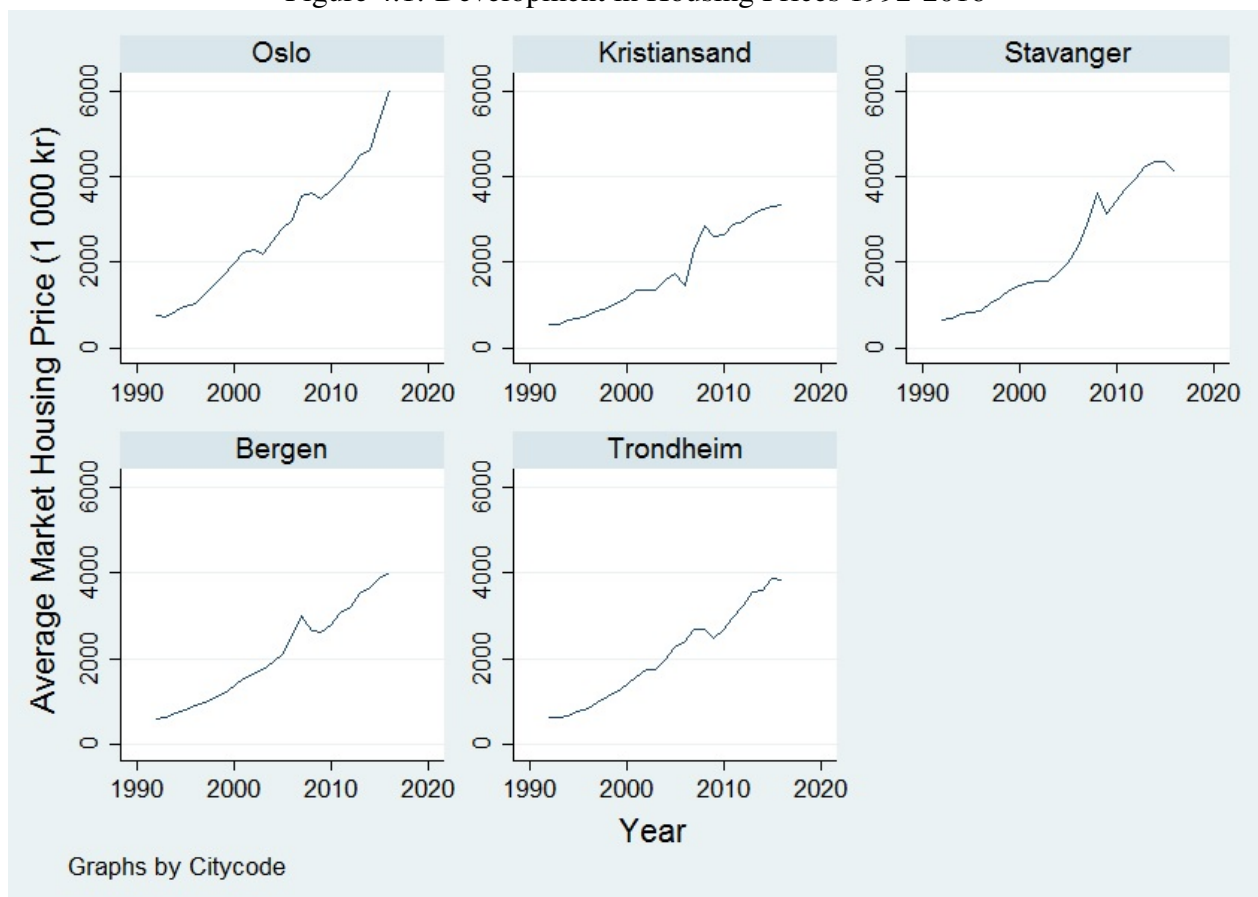


Table 4.1 shows the average housing price for three sub-periods of our dataset. Oslo has the highest price level in all three periods, followed by Stavanger. In the years 2009-2016 the average price for housing in Oslo was almost 4.6 million NOK, significantly higher than all the other cities.

Table 4.1: Housing Prices

	Average Housing Prices				
	Oslo	Kristiansand	Stavanger	Bergen	Trondheim
1992-1999	1.09 MNOK	0.75 MNOK	0.91 MNOK	0.85 MNOK	0.85 MNOK
2000-2008	2.58 MNOK	1.68 MNOK	2.08 MNOK	2.05 MNOK	2.05 MNOK
2009-2016	4.59 MNOK	3.01 MNOK	3.92 MNOK	3.34 MNOK	3.28 MNOK

4.2 Gross Income

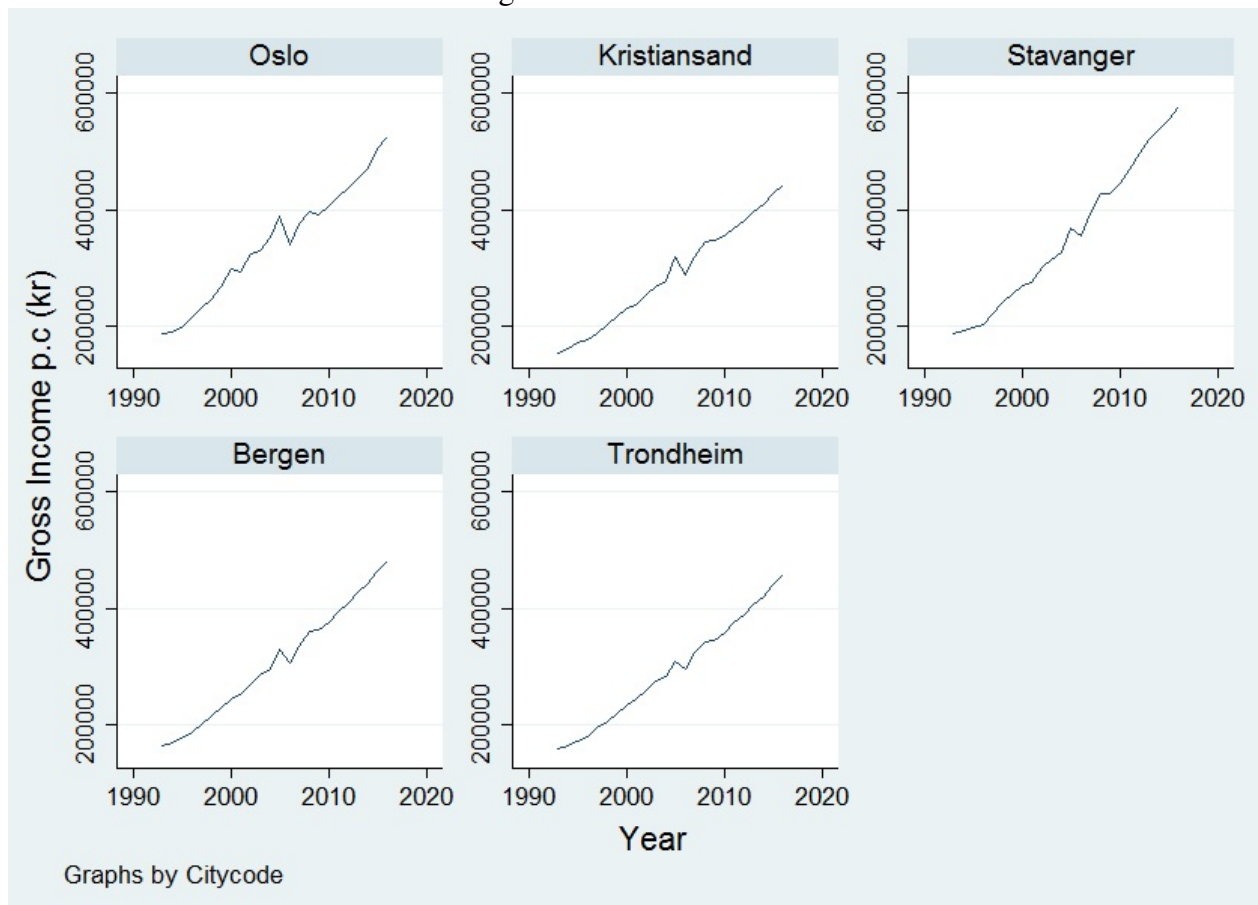
Figure 4.2 shows the development in average per capita gross income in Norwegian Kroner for the population over 17 years of age. The data is reported in the annual income tax returns. The data available for 2016 is incomplete and not on municipality levels. Gross Income for 2016 is therefore estimated based on the average annual growth rate the past five years for each municipality². Gross income includes labour income, trade income, pensions and capital income. Stavanger stands out as the highest earning municipality in the dataset followed by Oslo and Bergen. The data ranges from 1993 to 2016, and the five cities have had an average nominal gross income growth of 4.85 percent per annum in this period.

The large decline in gross income in 2006 is due to a change in the tax laws. From 2006, there were assessed taxes on dividends received for personal tax payers. Previous years this was tax-free. In 2005, dividends received amounted to NOK 99.3 billion, and in 2006 dividends received amounted to NOK 7.4 billion³.

²The estimation formula is showed in Appendix

³See Melby (2007)

Figure 4.2: Gross Income



All five cities lie above the national average for the entirety of our dataset. In 2015 the average per capita gross income was 25.8 percent higher in Oslo compared to the national average. Kristiansand, Stavanger, Bergen and Trondheim had an average gross income of 6.6, 37.9, 15.4 and 9.7 percent higher than the national average respectively.

Table 4.2 shows average gross income per capita for three sub-periods of our dataset. Oslo is the highest earning municipality for the two first sub-periods, while Stavanger is the highest earning municipality from 2009 to 2016. Kristiansand is the lowest earning municipality in all periods.

Table 4.2: Average Gross Income

	Average Gross Income p.c.				
	Oslo	Kristiansand	Stavanger	Bergen	Trondheim
1993-2000	229175 NOK	187750 NOK	220812 NOK	198562 NOK	191375 NOK
2001-2008	349900 NOK	287862 NOK	345150 NOK	304887 NOK	291900 NOK
2009-2016	451190 NOK	390651 NOK	503354 NOK	419693 NOK	399531 NOK

Higher gross income per capita attracts residents to the city and increases aggregate demand for housing. We expect income to have a positive effect on housing prices.

4.3 Unemployment Rate

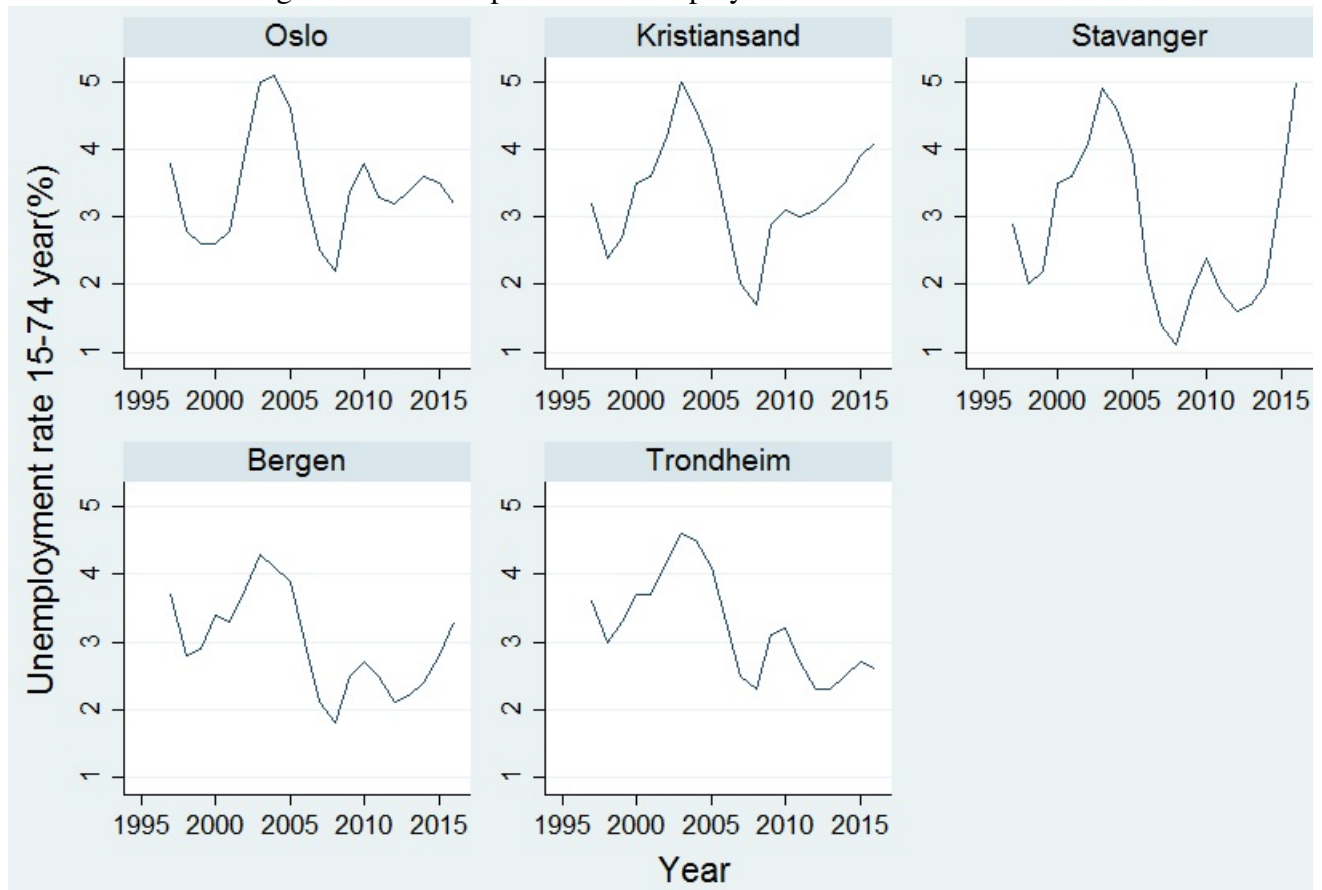
The unemployment rate is measured as the yearly average of total unemployed persons in percentage of the total labor force. The numbers are based on people that register themselves as unemployed through NAV. There are three criteria in international standards to be registered as unemployed⁴:

- People without income
- People that are searching for work
- People that are available for work

In addition to these criterias, NAV requires that the unemployed person also searches for a job through NAV. Our dataset consists of people in the ages of 18-74. Figure 4.3 illustrates the development in the unemployment rates in our five respective municipalities. All our selected municipalities has had the same trend, but with individual differences among percentage rates. The exception is the Stavanger region in year 2014 and 2015.

⁴See [Norwegian Welfare Administration \(2017\)](#)

Figure 4.3: Development in Unemployment Rate 1999-2016



Sørbo and Handal (2010) argues that the unemployment rate increased in affect of the declining business cycle in the beginning of the millennium. The unemployment continued to increase from 2001 to 2003/04, while declining towards the financial crisis in 2008. The financial crisis had an impact on the unemployment rate, in the form of higher unemployment. An economy stimulated by high oil prices and high investments gave a low unemployment rate from 2010 to 2012. Later, the oil price declined, which has resulted in strong unemployment growth in areas with close connections to the oil related industry, such as the Stavanger region. The average unemployment rate for Oslo is 3.36 percent for the whole period. Kristiansand, Stavanger, Bergen and Trondheim has an average of 3.38, 3.00, 2.95 and 3.10 respectively. Table 4.3 shows the average unemployment rate in percent for our five municipalities, in three sub-periods.

Table 4.3: Unemployment Rate
Average Unemployment Rate (%)

	Oslo	Kristiansand	Stavanger	Bergen	Trondheim
1996-2002	3.3	3.4	3.2	3.5	3.7
2003-2009	3.7	3.2	2.9	3.1	3.5
2010-2016	3.4	3.4	2.6	2.6	2.6

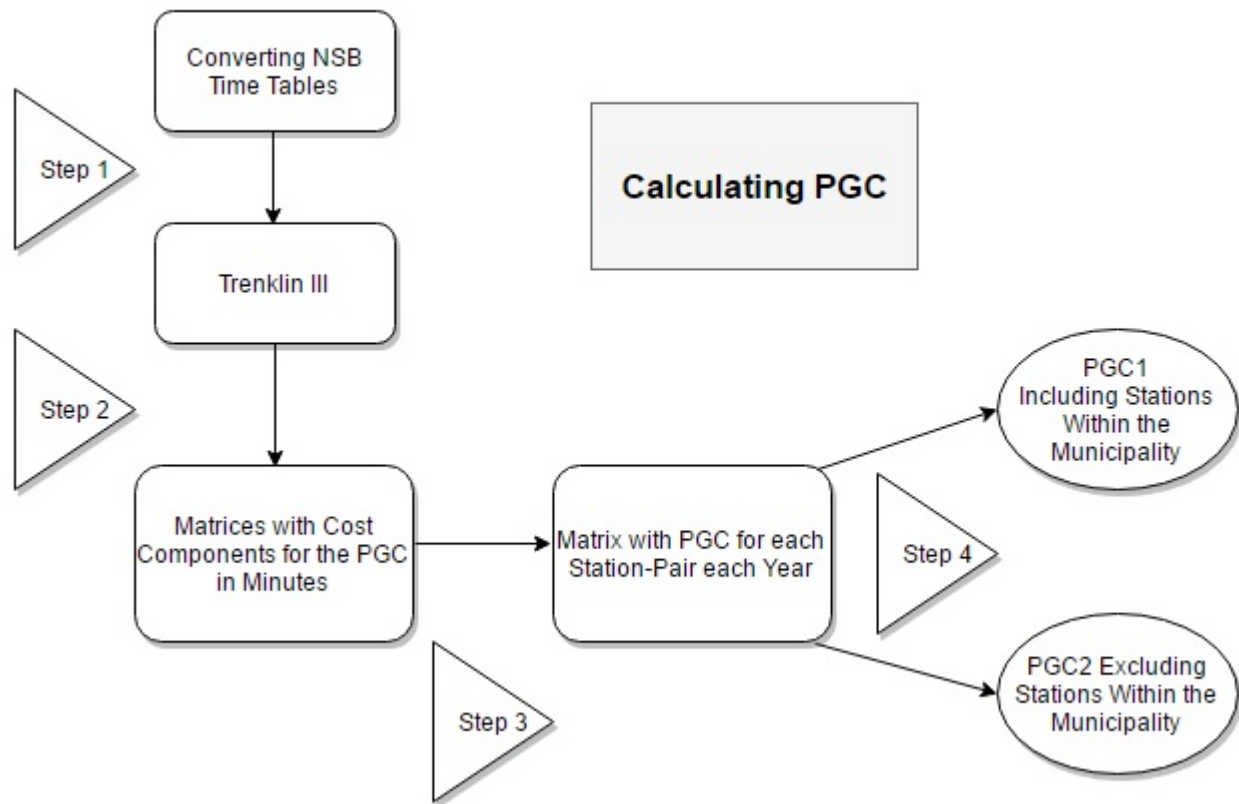
Higher unemployment is often correlated with lower demand in the housing market. In times of high unemployment, people would often be unsure about their future and be more careful with their investments. We expect the unemployment rate to have a negative effect on aggregate housing demand and housing prices.

4.4 Generalised Time Cost of Commuting

In this analysis we have calculated a variable for the generalised time cost of commuting by train in Norway. The variable is based on the time costs of different components of the travel. The variable does not include data for the ticket price as ticket prices varies with the same rate across entities and is accounted for in a fixed effects analysis by the same means as inflation. We therefore use the term generalised time cost and not generalised cost of commuting.

The dataset is based on timetables for every railway line in Norway from 1997 to 2016, provided by the Norwegian State Railways. Generating a variable from raw timetables is tedious and painstaking work, and because this variable will be unique for this thesis, we will dedicate a significant part of this section to explain the process. The appendix includes specific examples and more detailed explanation of the process of constructing the generalised time cost variable and its different components. The flow chart below describes the steps in the process of making the generalised time cost variable.

Figure 4.4: PGC Flow Chart



4.4.1 Converting NSB Time Tables

The first step in the process of making the generalised time cost variable was to convert time tables, obtained as PDF files from NSB, to a special Microsoft Excel format. The table below shows the format. The excerpt is from the railway line between Porsgrunn and Notodden called "Bratsbergbanen" in the Telemark county of Norway. The train stations along the railway line appears in the first column while the numbers represents the time point of departure and arrival at each station.

Table 4.4: Timetable Entries

Porsgrunn	x	x
Porsgrunn	x	0633
Skien	x	0645
Skien	0530	0646
Nisterud	0539	0655
Nisterud	0539	0655
Nordagutu	0559	0721
Nordagutu	0606	0722
Trykkerud	0616	0731
Trykkerud	0616	0731
Notodden	0625	0740
Notodden	x	x

The stations along a railway line is entered twice along the rows to capture arrival and departure time at each station. This is to measure eventual time barriers along the trip. The first station on a railway line would only have a departure time, while the last station on a line would only have an arrival time. The x symbolises no arrival or departure. As the table shows, the train leaves Skien at 0530 and arrives the last station Notodden at 0625. In the case of the station Nordagutu in the second column, the train arrives at 0559 and leaves at 0606. This means 7 minutes of waiting time on Nordagutu. These 7 minutes is a direct time cost for the traveller, and enters into the persons generalised time costs of travelling. The third coloumn shows the next departure on the line. The train now leaves the city of Porsgrunn at 0633, and arrives Notodden at 0740, with a 1 minute break at both Skien and Nordagutu station. Departures in the weekend and on holidays are not included in the dataset because of non relevans to commuters. The process of converting PDF time tables to excel spreadsheets was repeated for every line in Norway from the year 1997 to 2016.

4.4.2 Trenklin III

The Excel spreadsheets with every departure and arrival was ran trough a model made by the Norwegian Railway Directorate called Trenklin III. In Trenklin, generalised time costs are calculated between all station pairs in Norway based on the route plans that are added to the model and represents the train service. [Ranheim \(2017\)](#) describes in detail how Trenklin III is created and how it works. The structure of Trenklin, as applied in our thesis, is the following:

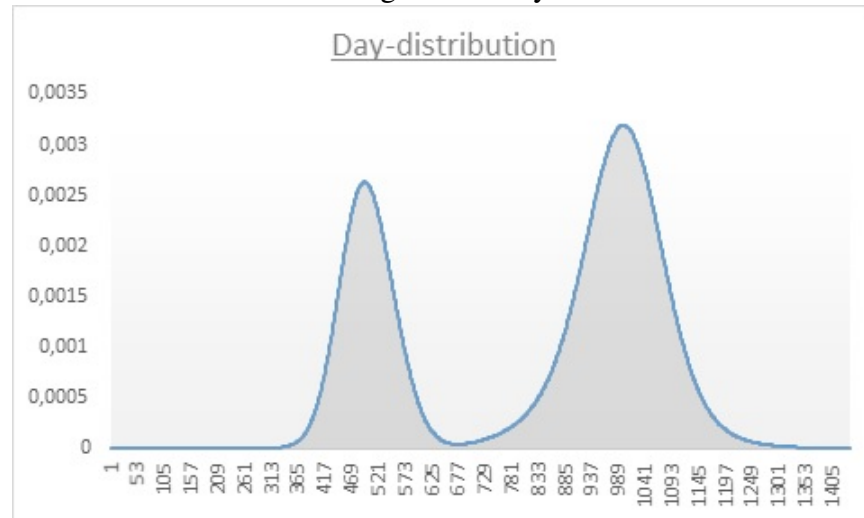
1. In-data and assumptions entered in to-and-from matrices

2. Train service in route-plan format added to the model
3. Trenklin III uses algorithms to calculate travel cost components

The cost components used to calculate the generalised time cost variable in our thesis are the cost of waiting time at the station, the cost of on board time, the time cost of train transfers and the cost of waiting time related to each transfer. Each of these components are described in detail in the appendix and are included in the total generalised time cost function. An important feature of Trenklin III is that the weights for waiting time is an increasing and diminishing function of the amount of minutes spent waiting. A one minute longer waiting time at the station has less impact on generalised time cost if the waiting time is already high than if the waiting time is short.

Trenklin operates with three different travelling purposes; work, leisure and business. These three have different weights for the different components of generalised time cost and different day-distributions of the relevant travelling times. For the purpose of our hypothesis, a special version of Trenklin III has been applied with the generous help of one of our supervisors, Patrick Ranheim, at The Norwegian Railway Directorate. In this version the generalised time cost is calculated using a special day-distribution in which only times relevant for work-related commuting is applied. The day-distribution describes travellers' preferred arrival time in discrete units, with a value for every minute within 24 hours (i.e a total of $60 \cdot 24 = 1440$ minutes). This means that the generalised time cost calculations in our thesis are foremost weighted in the periods people travel to and from work. The special day-distribution for our thesis is illustrated in figure 4.5:

Figure 4.5: Day-distribution of Travels Used in Thesis
Source: Norwegian Railway Directorate



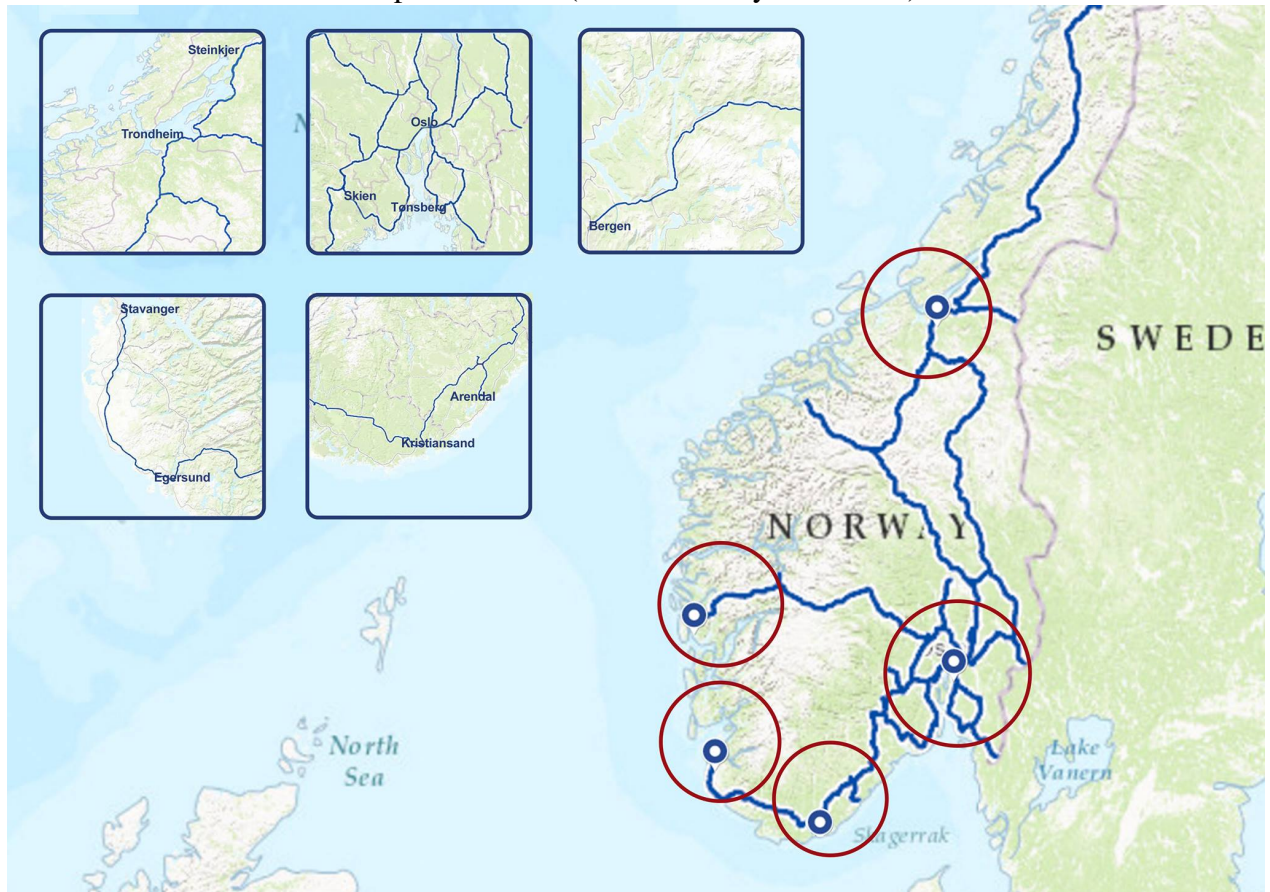
The minutes are along the x-axis and the weight for each minute on the y-axis. The weights sum to 1 over all the minutes within the 1440 minutes. This means that the first peak captures commuting to work and the second peak captures commuting from work. Commuting to work starts approximately at 0600 a.m (365/60) and lasts until approximately 1000 a.m(625/60). The period for commuting from work starts at approximately 1300 p.m(781/60) and lasts until approximately 1900 p.m(1145/60).

The generalised time costs estimates are presented in form of to-and-from matrices for each year. An example from one of the matrices is presented in the appendix. The matrices shows the generalised time cost of travelling between each individual station in Norway in minutes.

4.4.3 Calculating The Average Generalised Time Cost for Each City

After the generalised time costs matrices were obtained for each year(1997-2016), stations of relevance for our chosen municipalities were selected. In order to refine stations of relevance to commuting, a radius of 150 minutes of travelling from areas around the municipality in to the central station was implemented. We implemented the same radius as [Engebretsen et al. \(2012\)](#) uses in their report "Langpendling innenfor intercitytriangelet" on commuting flows within the intercity triangle in Oslo. Based on this commuting limit, each city's commuting area is illustrated in the picture below:

Figure 4.6: Central Municipalities and Commuting Radiuses
Map: Bane NOR (Customised by Tora Uhrn)

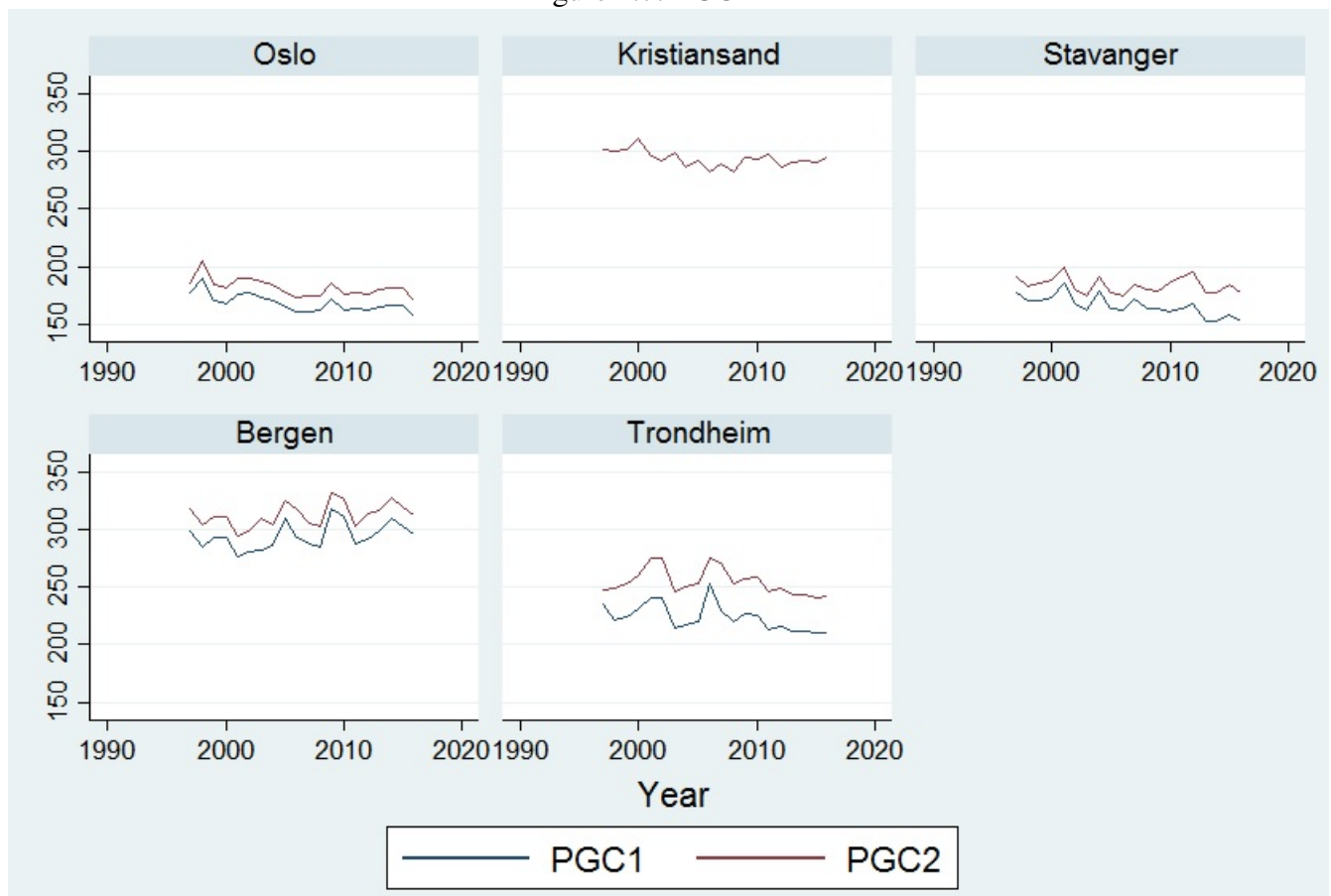


Since the characteristics of the journey is different depending on which way the journey is, the generalised time cost is also different. To take this into account, the generalised time costs of travelling to and from each station of interest was summed together. The average generalised time costs from the total number of stations within the 150 minutes radius was then calculated.

The next step was to exclude all stations within each municipality. Our analysis uses housing price data on municipality level. We analyse the effects of a change in the supply of train services and implicitly the population change. Failure to exclude stations within the municipality would lead to contradicting effects in housing prices related to improvements in generalised time costs. Improvements of the railway quality (i.e. a reduction in the generalised time cost) within the municipality would make living in the municipality more attractive and thereby increase aggregate demand. We therefore base the generalised time cost variable on travelling to the municipalities' central station from stations outside the municipalities' border, but within the commuting limit of

150 minutes. The separation gives two separate variables of the generalised time cost. PGC1 is the generalised time cost variable for all stations within 150 minutes to the central station and PGC2 is generalised time cost where stations inside the municipality border are excluded. Based on this rationale, PGC1 should have a smaller and less significant effect on the municipalities' housing prices. Figure 4.7 shows the development of the PGC variables for both within and outside the municipality:

Figure 4.7: PGC



Oslo and Stavanger has the lowest average PGC1 and PGC2 for the entire period. All cities have a lower PGC in 2016 than in 1997 implying an improvement of the railway quality during our time period in all municipalities. The reduction in generalised time cost is due to changes in the underlying parameters of the PGC variable. For example, the average on board time has been reduced from 54.7 minutes in 1998 and 57.7 minutes in 1999 to 52 minutes in 2016, when travelling by train to Oslo from close by areas. Also, the average amount of train transfers when

travelling into Stavanger has been reduced from 0.60 transfers in 1997, to 0.007 transfers in 2016. Both examples underline improvements in railway quality during our time period. The variation in PGC1 and PGC2 is lowest in Kristiansand, while highest in Trondheim. PGC2 lies above PGC1 in all cities implying that the average generalised time costs is higher when excluding the stations within the municipality border. The only exception is Kristiansand where no stations other than the central station lies within the municipality border so PGC1 equals PGC2. The descriptive statistics for PGC1 and PGC2 within each city is detailed in table 4.5 and 4.6:

Table 4.5: Descriptive Statistics PGC2(minutes)

	N	Std. deviation	Mean	Min	Max	2016-1997
Oslo	20	7.8	182	170.58	205.33	-14.71
Kristiansand	20	7.10	293.67	282.18	311.80	-6.5
Stavanger	20	7.17	184	174.30	199.62	-14
Bergen	20	10.22	312.63	294.21	331.99	-4.72
Trondheim	20	11.3	254.64	240.42	275.35	-5.24

Table 4.6: Descriptive Statistics PGC1(minutes)

	N	Std. deviation	Mean	Min	Max	2016-1997
Oslo	20	7.74	168.46	156.83	189.51	-21.17
Kristiansand	20	7.10	293.67	282.18	311.80	-6.5
Stavanger	20	8.85	166.15	152.23	185.56	-24.35
Bergen	20	11.10	294	276.57	317.25	-2.81
Trondheim	20	11.8	223.5	209.69	253.21	-24.44

Oslo has had the biggest improvement in the generalised time cost variable excluding stations within the municipalities since 1997. Including stations within the municipalities in PGC1, Trondheim has had the biggest reduction since 1997. The largest variation is also found in Trondheim.

Based on our hypothesis and the theory of Brueckner(1987), we expect both our generalised time cost variables to have a positive effect on housing prices, but that the effect of PGC1 to be smaller than PGC2. This implies an inverse relationship between railway transport quality to and from the municipality and housing prices within the municipality.

4.5 Descriptive Statistics

In this chapter we have presented and described our data with extra attention to the generalised time cost variable. Table 4.7 summarises the number of observations, global mean values, standard deviation, minimum value and maximum value for all our variables.

Table 4.7: Descriptive Statistics

Variable	Obs	Mean	Std. Dev.	Min	Max
Hp	125	2.198 MNOK	1.241 MNOK	0.535 MNOK	6.035 MNOK
unem	105	3.2 %	0.9 %	1.1 %	5.1 %
Y	120	318112 NOK	104207 NOK	154900 NOK	576534 NOK
PGC2	100	247.16 Min	56.17 Min	170.59 Min	332 Min
PGC1	100	229.17 Min	57.72 Min	152.23 Min	317.25 Min

Average housing price is almost 2.2 million NOK with the maximum of 6 million NOK found in Oslo in 2016. The average unemployment rate over the period 1996 to 2016 for all our municipalities is 3.2 percent. Average income from 1992 to 2016 is about 318000 NOK. PGC2 is higher than PGC1, meaning that generalised time costs of commuting is higher when excluding stations within the municipalities.

In the next chapter we will present and discuss our baseline model. We also discuss relevant issues regarding panel data analysis. We will present a static model as well as a dynamic specification which allows for sluggish adjustments in housing prices.

Chapter 5

Empirical Specification

In this chapter we will specify our empirical model for housing prices based on the panel data on our five municipalities. We model housing prices assuming a fixed supply as a function of demand side factors on total demand for housing.

5.1 Baseline Model

Our baseline model is a fixed effects model with log-specification, implying that all parameters are interpreted as elasticities:

$$lHp_{it} = \alpha_0 + \alpha_1 l unem_{it} + \alpha_2 l Y_{it} + \alpha_3 l PGC_{it} + a_i + u_{it} \quad (5.1)$$

Where:

- $lHp = \log(\text{Housing Prices})$
- $l unem = \log(\text{Unemployment Rate})$
- $l Y = \log(\text{Gross Income})$
- $l PGC = \log(\text{Private Generalised Time Cost of Commuting})$
- $a_i = \text{City specific fixed effects}$
- $u_{it} = \text{idiosyncratic error term}$

We estimate a fixed effects regression model to account for the regional specific attributes affecting housing prices and therefore exploit the within group variation in the data. The term a_i in equation (5.1) represents such fixed attributes. Failure to control for such regional fixed effects will cause biased estimators as these effects will be captured in the estimated coefficients for our explanatory variables. Dealing with panel data that includes unobserved regional specific effects, both fixed effects estimation and random effects estimation techniques can be applied. The difference between fixed effects and random effects models is the different assumptions relating to the unobserved effects. The random effects model assumes:

$$Cov(a_i, X_{itk}) = 0$$

Which says that the regional specific effects are uncorrelated with the other explanatory variables in the model. Fixed effects estimation on the other hand, assumes:

$$Cov(a_i, X_{itk}) \neq 0$$

We find it reasonable to assume that the regional specific effects are correlated to our regressors in equation (5.1), and that a fixed effects model is appropriate versus a random effects model. In our model, the regional specific effects would include factors such as the location of the municipalities, cultural aspects, political systems, climate etc. Economic intuition tells us that for example the location of Stavanger, on the west coast of Norway and close to the offshore oil industry, has had an effect on its unemployment level and income level. Another example could be that the topography around Oslo is more suitable for both housing and railway infrastructure, compared to the mountainous landscape on the west coast, and therefore have affected housing prices and railway quality.

Fixed effects estimation has certain assumptions¹:

- Something within the entities may impact or bias the estimates or outcome variables.
- The time-invariant characteristics are unique to the entity and are not correlated with the characteristics of the other entities

¹Full detail of Assumptions in Appendix.

Using a fixed effects model, we transform the data by deducting the time averages of each variable. Since a_i in equation (5.1) is constant it is removed from the model.

All variables enters in logarithmic form. This narrows the range of the variables' variation and gives intuitive interpretations of the coefficients. The limitations of using logs is that the variables cannot be zero or negative. For our dataset this limitation can be neglected. In addition, Wooldridge (2013) argues that taking the log of the variables also removes the problems of skewed distributions with heteroskedastic variance. In cases where both dependent and independent variables are in log form the, coefficients are interpreted as elasticities; the percentage change in the dependent variable with a one percent change in the independent variable. When the independent variable is on level form the coefficient is interpreted as a semielasticity. The percentage change in the dependent variable with a one unit change in the independent variable is calculated as $100 * (e^{\alpha_i} - 1)$.

5.2 Time Fixed Effects

There is reason to suspect that there are factors affecting housing prices that varies over time, but are common between the municipalities. These factors would include interest rate, price inflation and other macro economic variables. We can control for such factors by including dummy variables for each time period. The dummy variables takes the value 1 for its relevant period and zero otherwise. This allows for a separate intercept for each year. In equation (5.1) we have:

$$lHp_{it} = \alpha_0 + \alpha_1 lunem_{it} + \alpha_2 lY_{it} + \alpha_3 lPGC_{it} + \gamma_1 d_{1998} + \dots + \gamma_{20} d_{2016} + a_i + u_{it} \quad (5.2)$$

Where d_{1998} equals 1 in the year 1998 and 0 otherwise and similarly for the other years, making the year 1997 the year of reference. The dummy variables have no i subscript because they do not change across the cities. The intercept for year 1997 is α_0 and for 1998 the intercept is $\alpha_0 + \gamma_1$. Having the log of housing prices as the dependent variable would imply allowing for a different mean growth rate of housing prices in each year.

The downside of using yearly dummies is that it requires estimation of T-1 extra parameters. In our dataset T is large relatively to N and using yearly dummies might steal too many degrees of

freedom. To avoid these extra estimations, another way to account for the effect of time on housing prices, is to allow for a linear time trend using a variable that increases with one unit(year) for each year. This estimates an average effect of time on housing prices, rather than a single effect for each year, and saves degrees of freedom. Again returning to our baseline model (5.1), but including a time trend variable we have:

$$lHp_{it} = \alpha_0 + \alpha_1 lunem_{it} + \alpha_2 lY_{it} + \alpha_3 lPGC_{it} + \theta T + a_i + u_{it} \quad (5.3)$$

The effect on housing prices of adding one year is: $\frac{\delta lHp_{it}}{\delta T} = \theta$

5.3 Housing Market Dynamics

Equation (5.1) is a static model where the full effect of each explanatory variable is assumed to be realized in the current period. Including a lagged dependent variable in equation (5.1) allows for a time lag in the adjustment of the housing prices, and the explanatory variables to have a short-run effect as well as a long-run effect:

$$lHp_{it} = \alpha_0 + \beta_1 lHp_{it-1} + \alpha_1 lunem_{it} + \alpha_2 lY_{it} + \alpha_3 lPGC_{it} + a_i + u_{it} \quad (5.4)$$

In equation (5.4) the log of housing prices is assumed to depend on last periods value to account for delays in housing price adjustments. The long term effects of changes in unemployment, income and PGC on log housing prices is calculated by assuming that the housing market is in equilibrium with $lHp_{it} = lHp_{it-1} = lHp_i$ and solving for lHp_i :

$$lHp_i = \frac{\alpha_0}{1 - \beta_1} + \frac{\alpha_1}{1 - \beta_1} lunem_{it} + \frac{\alpha_2}{1 - \beta_1} lY_{it} + \frac{\alpha_3}{1 - \beta_1} lPGC_{it} + \frac{a_i + u_{it}}{1 - \beta_1} \quad (5.5)$$

The long-run effects of unemployment, income and PGC are:

$$\frac{\delta lHp_i}{\delta lunem_{it}} = \frac{\alpha_1}{1 - \beta_1}$$

$$\frac{\delta lHp_i}{\delta lY_{it}} = \frac{\alpha_2}{1 - \beta_1}$$

$$\frac{\delta lHp_i}{\delta lPGC_{it}} = \frac{\alpha_3}{1 - \beta_1}$$

The short-run effects are the coefficients α_1 , α_2 and α_3 . We assume that issues regarding biasedness of the within estimators, due to autocorrelation when including a lagged dependent variable in a fixed effects model, is negligible as our time dimension in our dataset is relatively large ($T=20$). Baltagi (2008) showed that as T gets large the fixed effects estimator becomes consistent.

5.4 Serial Correlation and Driscoll and Kraay Standard Errors

Working with data which is collected repeatedly over time and for several cross-sectional entities, rises the issue of correlation in the error terms or serial correlation. For equation (5.1) to give efficient estimators and valid t-statistics, the error term must be uncorrelated across time and cross-sections². From Wooldridge (2013) we know that if this assumption is violated, the estimated standard deviations of the coefficients will be smaller than the true standard errors. This could lead to the interpretation that the estimates are more precise than they actually are.

Wooldridge (2013) explains one way to solve this problem by estimating serial correlation-robust standard errors that account for the serial correlation in the errors, using the framework of Newey and West (1987). Starting with the standard multiple linear regression model:

$$y = \beta_0 + \beta_1 x_{t1} + \dots + \beta_k x_{tk} + u_t$$

It is shown that the the serial correlation-robust standard error of a coefficient in a multiple linear regression can be estimated as:

$$se(\hat{\beta}_1) = [se(\hat{\beta}_1) / \hat{\sigma}]^2 \sqrt{\hat{v}}$$

Where " $se(\hat{\beta}_1)$ " is the incorrect standard error from OLS. This standard error is then divided by the standard error of the regression and \hat{v} is defined as :

$$\hat{v} = \sum_{t=1}^n \hat{a}_t^2 + 2 \sum_{h=1}^g [1 - h/(g+1)] \left(\sum_{t=h+1}^n \hat{a}_t \hat{a}_{t-h} \right) \quad (5.6)$$

²See Appendix for assumptions for fixed effects and random effects estimation

Where $\widehat{a}_t = \widehat{r}_t \widehat{u}_t$ i.e. the product of the estimated residuals from the initial multiple linear regression (\widehat{u}_t) and the estimated residuals from an auxiliary regression of x_{t1} on $x_{t2}, x_{t3}, \dots, x_{tk}$ (\widehat{r}_t). The integer g defines the amount of terms included to correct for serial correlation. With AR(1) serial correlation the integer g equals 1 and equation (5.6) becomes simply $\widehat{v} = \sum_{t=1}^n \widehat{a}_t^2 + \sum_{t=2}^n \widehat{a}_t \widehat{a}_{t-1}$. When the order of serial correlation increases, the adjustment for the robust standard deviations increases.

The term $1 - \frac{h}{g+1} > 0$ in (5.6) ensures that higher order lags receives less weight.

[Hoechle \(2007\)](#) has developed a statistical package in Stata, which computes Driscoll and Kraay serial correlation robust standard errors³. Driscoll and Kraay robust standard errors equals the robust standard errors estimated in [Newey and West \(1987\)](#), but allows the sum $\sum_{t=h+1}^n \widehat{a}_t \widehat{a}_{t-h}$ in equation (5.6) to run to $n(t)$, depending on the amount of observations for each entity, allowing for unbalanced panels. Driscoll and Kraay robust standard errors account for both cross-sectional and time serial correlation as well as heteroskedasticity. Heteroskedasticity implies that the variance of the residuals is not constant, which violates the assumption of efficient estimators. The general idea is to estimate the regression models using OLS, but to adjust the standard errors to account for serial correlation in the error terms which then makes inference and interpretations valid. See [Hoechle \(2007\)](#) for a more complete understanding of the derivation of both Driscoll and Kraay standard errors and the application in STATA.

5.5 Measurement Errors

The self-generated PGC variables are a central part of our analysis. We therefore consider it important to address the issue of potential measurement errors and how these can affect the empirical analysis. The true generalised time costs of travelling is unobserved and the estimated PGC variables will always be exactly that; estimates. If the PGC is measured as the true unobserved PGC plus some measurement error we can write.

$$PGC_{it} = PGC_{it}^* + e_{1it}$$

³Command "xtscc" in STATA

Where PGC_{it}^* is the true measure of the generalised time costs and e_{1it} is the measurement error. Considering a simple model for housing prices as an illustration, we can write:

$$Hp_{it} = \alpha_0 + \alpha_1 PGC_{it}^* + \alpha_2 unem_{it} + \alpha_3 Y_{it} + a_i + u_{it} \quad (5.7)$$

unemployment(unem) and income(Y) are observable, but PGC^* is not. Using PGC to estimate this model could cause biased and inconsistent estimators because the error term would be correlated with the explanatory variable PGC:

$$Hp_{it} = \alpha_0 + \alpha_1 PGC_{it} + \alpha_2 unem_{it} + \alpha_3 Y_{it} + (a_i + u_{it} - \alpha_1 e_{1it})$$

If we assume that the measurement error is uncorrelated with the PGC, that the expected value of the measurement error is zero, and assume that the PGC variable is uncorrelated with the other regressors, the bias will only affect α_1 and be towards zero. This implies that the true effect of generalised time costs on housing prices is higher than in our estimates.

Any measurement error that is constant over time and between entities will be removed when estimating equation (5.7) using a fixed effects transformation. This means that if there are measurement errors in the parameters used to calculate the different weights as described in chapter 4, these errors would be removed with fixed effects transformation as these parameters are fixed. If the measurement error is constant but varies between entities, so that $e_{it} = e_i$, estimating equation (5.7) using fixed effect transformation would also eliminate the measurement error from the error term as with other fixed effects, a_i , such as city location etc. A more important issue arises if the measurement error varies from year to year, but as serially uncorrelated noise. [Black et al. \(2000\)](#) argues that in such cases, econometric methods like for example first differencing could actually increase the variation of the error, but also reduce the variation in the regressor if the regressor is serially correlated.

In this chapter we have specified our empirical fixed effects model, discussed issues relevant for panel data analysis and how measurement errors can affect the analysis. The next chapter presents our results as well as economic interpretations of the estimated relationships between the cost of commuting and housing prices.

Chapter 6

Results

In this chapter, we will present the results of our regression analysis. We first estimate the static model specification as in equation (5.1) before including a lag of the dependent variable to account for sluggish housing price adjustments as in equation (5.4). We apply a [Levin et al. \(2002\)](#) panel data unit root test to all our variables, and conclude that they do not contain unit roots and that they are stationary. We also test the models for serial correlation using the framework of [Durbin \(1970\)](#). All tests are included in the appendix.

6.1 Static Housing Price Model

Based on the Durbin's test we find evidence for AR(1) serial correlation in the error term. We therefore estimate the fixed effects model (5.1) using the Driscoll and Kraay robust standard errors. Table 6.1 shows the results using year dummies and time trend to account for the common time-varying factors. The estimation includes both PGC1 and PGC2:

Table 6.1: Fixed Effects Estimates With Driscoll and Kraay Robust Standard Errors

	(1)	(2)	(3)	(4)
	IHp	IHp	IHp	IHp
lunem	-0.113 (0.0472)	-0.161 (0.0456)	-0.107 (0.0470)	-0.162 (0.0426)
IY	0.876 (0.170)	1.241 (0.215)	0.976 (0.191)	1.281 (0.207)
IPGC2	0.346 (0.171)	0.357 (0.138)		
IPGC1			0.245 (0.131)	0.312 (0.0961)
time		0.0196 (0.00889)		0.0182 (0.00865)
cons	-5.578 (2.350)	-10.02 (2.062)	-6.233 (2.666)	-10.25 (2.156)
<i>N</i>	100	100	100	100
Year dummies	Yes	No	Yes	No
Within R^2	0.9857	0.9765	0.9853	0.9763

Standard errors in parentheses

Model (1) and (2) includes PGC2 as the generalised time cost variable, while in model (3) and (4) we use PGC1. The effect of unemployment on housing prices is as expected. In model(1) we find an elasticity of housing prices with respect to unemployment rate of -0.113 percent. In model (2) this elasticity is also negative with a value of -0.161. We find a positive income effect in both models. We estimate an elasticity of housing prices with respect to changes in income of 0.876 percent in model(1) and 1.241 percent in model(2), both significant at a 1 percent significance level. The elasticity of income is not significant different from 1 in either model (1) or (2)¹; a 1 percent increase in income, gives 1 percent increase in housing prices. This is in line with previous literature. [Larsen \(2002\)](#) finds that the income elasticity with respect to housing demand is close to unity in Norway over the time span 1986 to 1998.

The variable in focus for this thesis, PGC2, is estimated to have a positive effect on housing prices. This supports our hypothesis of an inverse relationship between changes in railway transportation quality and housing prices within the central municipalities. In model(1) the elasticity of housing prices with respect to generalised time costs is 0.346 percent, and in model(2) the elasticity

¹The t-test on general form is shown in the appendix

is estimated to be 0.357 percent. The effect of PGC2 is therefore practically identical when using time trend or time dummies. The effect is significant at a 5 percent significance level in both models. The table below illustrates the estimated effect of changes in generalised time costs on housing prices within a 95 percent confidence interval based on 2016 averages:

Table 6.2: Change in 2016 Housing Prices With a 1 Percent Increase in PGC2, 95%-Confidence Interval

	Lower bound	Upper bound
Model with trend	2943 NOK	27536 NOK
Model with dummy	-529 NOK	30070 NOK

The only case with a negative effect on housing prices is for the lowest bound in the confidence interval, using the model with yearly dummies. The upper bounds of the estimated effect is about 30 000 NOK. This illustrates the improbability that the effect of generalised time costs on housing prices is negative.

In chapter 4 we explained how we expect the PGC1 variable to have a smaller effect on housing prices than PGC2. When using PGC1, we do not exclude the stations within the municipalities where the housing prices are measured. As a consequence there are two contradicting effects of railway improvements. Railway transportation improvements within the municipality can increase the attractiveness of the municipality and lead to an increase in aggregate demand for housing. On the other hand, improvements outside the municipality can reduce aggregate demand in the municipality in favour of peripheral areas. Table 6.1 shows that the estimated effect of generalised time costs on housing prices is in fact lower using PGC1 rather than PGC2, all though more significant in the model with trend. We will continue our analysis focusing on PGC2.

6.2 Housing Price Dynamics

To account for lagged adjustments of housing prices, as explained in the empirical specification chapter, we estimate equation (5.4) which includes lagged value of log housing prices. We apply both time trend and yearly dummies. Durbin's test again find evidence of serial correlation. We therefore estimate the fixed effects model (5.4) with Driscoll and Kraay robust errors. The results

are shown in table 6.3:

Table 6.3: Fixed Effects Estimates with Driscoll and Kraay Robust Standard Errors Dynamic Model

	(1)	(2)
	IHp	IHp
lag_IHp	0.465 (0.136)	0.428 (0.131)
lnem	-0.132 (0.0386)	-0.128 (0.0421)
IY	0.317 (0.179)	0.720 (0.209)
IPGC2	0.223 (0.119)	0.210 (0.0955)
time		0.00840 (0.00367)
cons	-1.190 (1.690)	-5.792 (1.905)
<i>N</i>	100	100
Year Dummies	Yes	No
Within R^2	0.9891	0.9825

Standard errors in parentheses

The estimated coefficient for our lagged housing price variable shows that there is a significant lag in the housing price adjustment, in line with the findings of Arestis and Gonzales(2013). All else equal, log housing prices today will be about 43 percent of log housing prices last year using time trend and 46 percent using year dummies. As Arestis and Gonzales(2013) points out, this underlines the role of agents' expectations in the housing market. When expectations plays a significant role in determining market prices, housing prices will appear more sticky as expectations often depends on previous values.

Unemployment has a negative impact on housing prices with a short run elasticity of -0.132 and -0.128. The full effect of unemployment-changes on long-run equilibrium housing prices is estimated to be about -0.25 and -0.22 percent in model(1) and model(2) respectively. The long-run effects are calculated as explained in chapter 5.

The short run effect of gross income on housing prices is positive. Using yearly dummies in model(1) we estimate a short-run elasticity of 0.317 percent and 0.72 with time trend in model (2).

The long run effect of income is 0.59 and 1.26 percent respectively. This effect is fairly close to what we estimated in the static case in model (5.1).

The effect of railway transportation quality is still in line with our hypothesis. That is, we find an inverse relationship between railway commuting quality and housing prices. The short run elasticities with respect to generalised time costs are 0.22 and 0.21 in model (1) and (2) respectively. The full long-run effects are therefore estimated to be 0.42 and 0.37 percent. This implies an average long run effect of over 17000 NOK in model(1) and over 15000 NOK in model(2) on average housing prices for our five cities, with a one percent increase in generalised time costs in 2016. Using yearly dummies in model(1) the effect of log PGC2 is significant at a 10 percent significance level, while using trend in model (2) the effect is significant at a 5 percent significance level.

6.3 Effects of Generalised Time Costs on 2016 Housing Prices

Using the elasticities, estimated in both the dynamic and static models, we calculate monetary effects based on housing prices in 2016. Table 6.4 illustrates the estimated effects of a 1 percent and 1 minute increase in generalised time costs in the different models.

Table 6.4: Summary of the Effects on Housing Prices with Lower Transport Quality
Effect on housing prices from 2016 average

	Elasticity	1% increase in PGC	1 minute increase in PGC
Static model			
With trend	0,357	15229 NOK	6354 NOK
With dummy	0,346	14760 NOK	6158 NOK
Dynamic Model			
Short run			
With trend	0,210	8958 NOK	3738 NOK
With dummy	0,223	9513 NOK	3969 NOK
Long run			
With trend	0,367	15784 NOK	6586 NOK
With dummy	0,416	17916 NOK	7476 NOK

The lowest estimate of the full effect of a one minute increase in commuting time costs is 6158 NOK increase in housing price in the municipality. Accounting for lags in housing price adjustments, the immediate effect of a one minute increase in commuting costs is just short of 4000

NOK. The full long run effect of a 1 minute increase in the time cost of commuting is 6586 NOK using trend, and 7476 NOK using time dummies. Approximately 50 percent of the full long run effect of a one percent increase in generalised time costs will be realised within the first year.

Further, we wish to illustrate the monetary effects based on each municipality's historical variation in generalised time costs. Table 6.5 summarises the variation in both housing prices and generalised time costs for our five municipalities:

Table 6.5: Housing Price and PGC2 Variation

	Hp(NOK)			PGC2(Minutes)		
	Mean	St.dev.	Average annual change	Mean	St.dev.	Average annual change
Oslo	3211065	1302815	249785	182,0	7,8	-0,77
Kristiansand	2101177	897360	128525	293,7	7,1	-0,34
Stavanger	2679810	1222100	162811	184,1	7,2	-0,74
Bergen	2422009	970220	156245	312,6	10,2	-0,25
Trondheim	2403556	922285	150752	254,6	11,3	-0,28

The largest average annual change in both generalised time costs and housing prices is found in Oslo. Trondheim has the highest standard deviation in generalised time costs. To simulate the effect of a typical railway quality improvement, we estimate the long-run effect on housing prices if generalised time costs were reduced by one standard deviation in 2016. The calculations are based on the elasticity in the dynamic model(2) which is lower than in model(1), but more precise. The estimated effects for each city is shown in table 6.6:

Table 6.6: Effect of one Standard Deviation Railway Quality Improvement

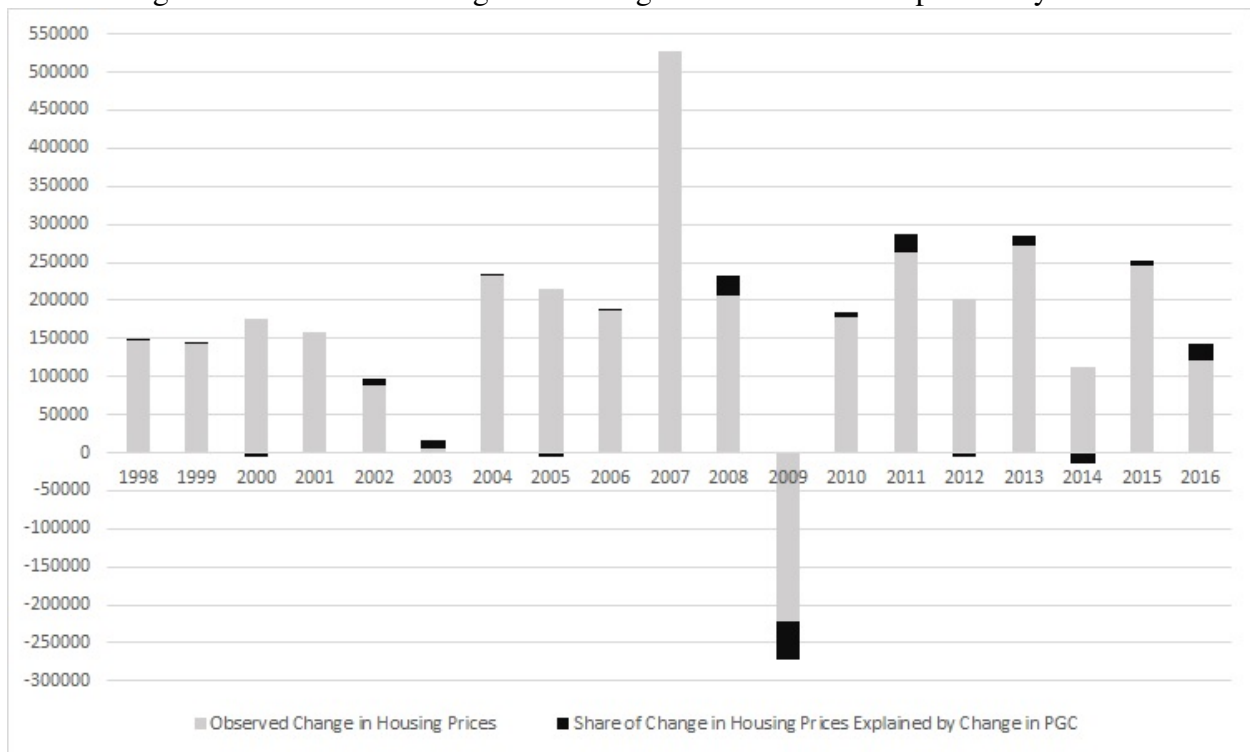
	Railway Quality Improvement	Long-run Effect on 2016 Housing Prices	Long-run Effect on 2016 Housing Prices (NOK)
Oslo	7,8 min	-1,67 %	-101060 NOK
Kristiansand	7,1 min	-0,88 %	-29503 NOK
Stavanger	7,17 min	-1,48 %	-61089 NOK
Bergen	10,22 min	-1,20 %	-48069 NOK
Trondheim	11,3 min	-1,71 %	-65485 NOK

The table shows the reduction in housing prices due to a typical reduction in commuting costs in 2016, based on each city's previous variation in generalised time costs. The large effect for the city of Trondheim is driven by its relatively large variation in generalised time costs. Because of

it's high housing prices in 2016, Oslo has the largest decline in housing prices with a one standard deviation reduction in PGC.

Based on our results we know that some of the variation in housing prices in the central municipalities in Norway is due to variation in the generalised time costs of commuting to and from these municipalities via its effects on population changes. Using our estimates for the effect of generalised time costs on housing prices we can calculate how the housing prices would have varied without the developments in commuting costs. As an illustration, figure 6.1 shows the average annual change in housing prices as observed in the data and the estimated change in housing prices accredited to changes in generalised time costs. We have applied the elasticity as estimated in the static model with time trend, which is the biggest short-run elasticity in all our models.

Figure 6.1: Observed Change in Housing Prices and Share Explained by PGC2



The grey columns are the observed annual change in housing prices. The black columns illustrates the change in housing prices due to changes in generalised time costs. This means that the sum of the grey and black columns shows how housing prices would have varied, without the effects of generalised time costs. In years with a decreased generalised time cost, the change in the housing price would have been bigger. In 2016 the grey column shows that the observed change

in housing prices was 121000 NOK. This year the generalised time costs decreased with around 3 minutes. Without this reduction in commuting costs, the change in housing prices would have been almost 150 000 NOK.

In 2009 we see that the decline in housing prices would have been bigger without the increased generalised time costs that year. Observed change in housing prices was about -222 000 NOK, while the average generalised time cost of commuting increased with 10 minutes. This means that the estimated decrease in housing prices in 2009, without the change in generalised time costs, is 272 000 NOK.

6.4 Alternative Functional Forms and Non-linearity

In the specifications above we have estimated constant elasticity models(CES), assuming a linear relationship between the log of PGC2 and log of housing prices. In this section we estimate an alternative model allowing us to identify potential non-linearity between the two variables. We want to analyse whether the relationship between log housing prices and generalised time costs is characterised by decreasing or increasing marginal effects. To do this we include PGC2 as well as PGC2 squared. This allows us to identify how the effects of generalised time costs on housing prices depends on the initial generalised time cost value. The dynamic model with non-linearity becomes:

$$lHp_{it} = \alpha_0 + \beta_1 lHp_{it-1} + \alpha_1 lunem_{it} + \alpha_2 lY_{it} + \alpha_3 PGC_{it} + \alpha_4 PGC_{it}^2 + a_i + u_{it} \quad (6.1)$$

The marginal effect of PGC_{it} on lHp_{it} is:

$$\frac{\delta lHp_{it}}{\delta PGC_{it}} = \alpha_3 + 2(\alpha_4 PGC_{it})$$

Table 6.7 shows the regression results:

Table 6.7: Dynamic Model Estimation with Non-linearity

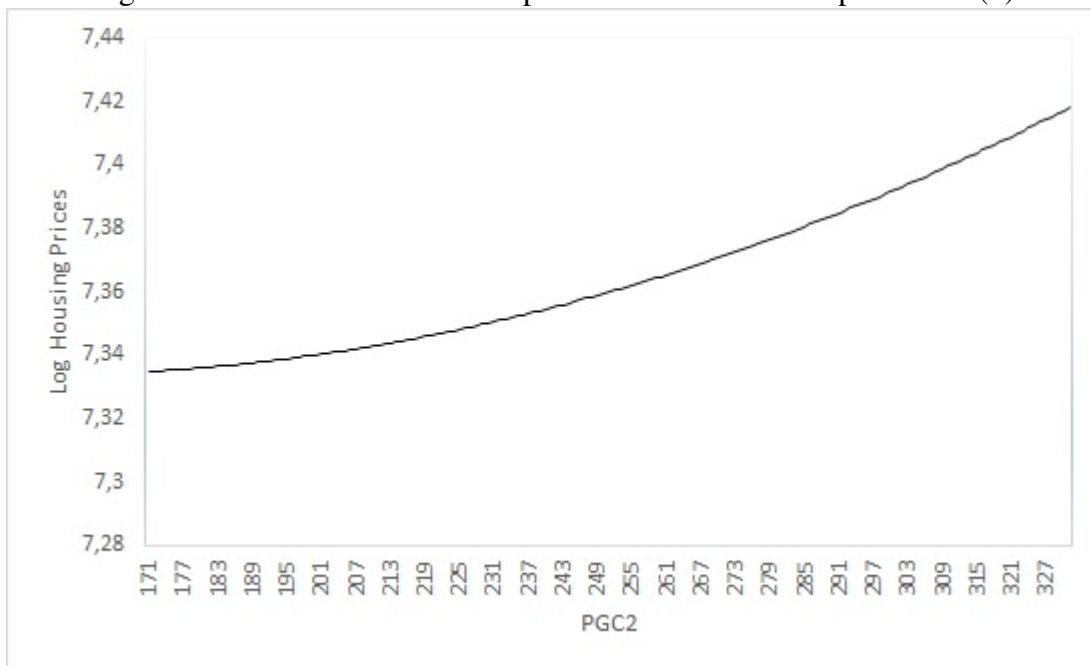
	(1)	(2)	(3)	(4)
	lHp	lHp	lHp	lHp
lag_lHp	0.465 (0.136)	0.428 (0.131)	0.497 (0.146)	0.443 (0.135)
lunem	-0.132 (0.0386)	-0.128 (0.0421)	-0.133 (0.0402)	-0.122 (0.0402)
lY	0.317 (0.179)	0.720 (0.209)	0.303 (0.189)	0.725 (0.216)
IPGC2	0.223 (0.119)	0.210 (0.0955)		
PGC2			-0.000775 (0.00356)	0.00158 (0.00330)
PGC2 ²			0.00000258 (0.00000766)	-0.00000248 (0.00000707)
time		0.00840 (0.00367)		0.00671 (0.00347)
cons	-1.190 (1.690)	-5.792 (1.905)	0.0174 (1.590)	-5.043 (1.872)
<i>N</i>	100	100	100	100
Year dummies	Yes	No	Yes	No
Within <i>R</i> ²	0.9891	0.9825	0.9888	0.9822

Standard errors in parentheses

Model (1) and (2) are earlier estimated models, but included for comparison. In model (3) and (4) the PGC level and squared values are included to allow for decreasing or increasing marginal effects.

In model (3), using yearly dummies, the first order effect of PGC2 on log housing prices is negative, but at a diminishing rate. This is not in line with our hypothesis, but evaluating the estimated effect against our dataset-range shows that the minimum point is lower than the minimum value of the PGC. The point of which the derivative of log housing prices with respect to PGC2 equals zero is approximately when PGC2 equals 150 minutes. The minimum value of the PGC2 variable in our dataset is 170.58 minutes. This implies that the marginal effect of PGC relevant for our dataset is in fact positive. Figure 6.2 shows the estimated relationship between PGC2 and log housing prices within our data-range for PGC2, assuming mean values for all other variables:

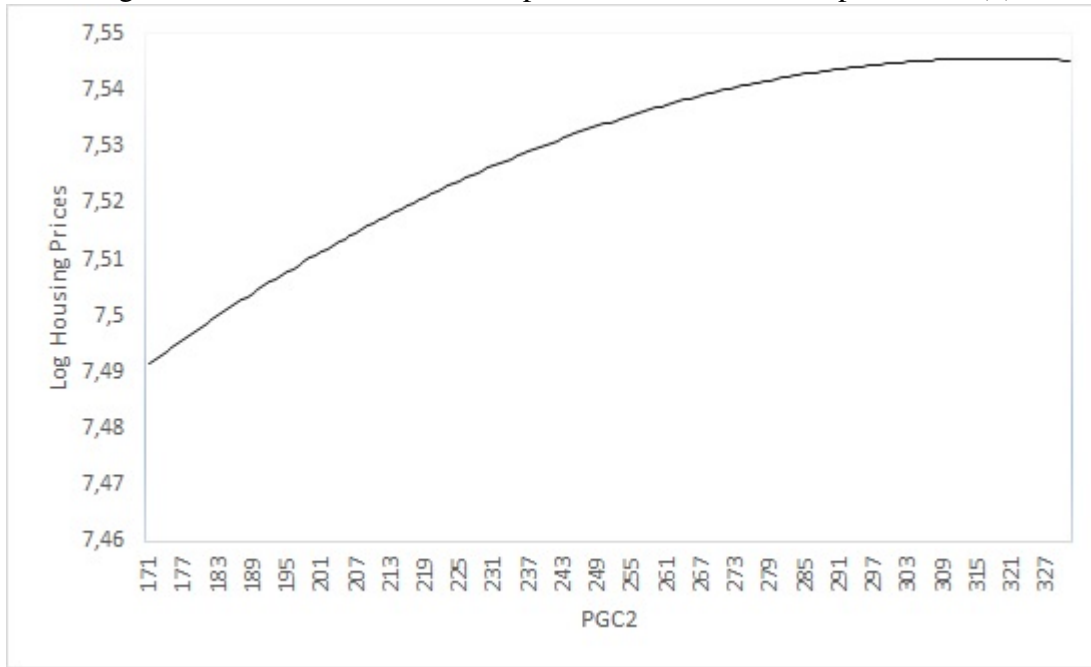
Figure 6.2: Estimated Relationship Between PGC2 and IHP in Model(3)



The figure illustrates the increasing marginal effects of PGC2 on log housing prices within our datarange. Increasing marginal effects means that when the initial levels of commuting costs are high, the effect of a marginal increase in commuting costs will be bigger than if the initial level is low. The reason for this relationship could be that when the cost of commuting between a certain station and the municipality's central station is high, the station usually is in a remote area with few other commuting options other than railway. That is, the households in more remote areas are more dependent on the railway as a means of commuting. This means that the marginal increase in PGC2 has a bigger impact on the households' migration decision, than if the station is located in a more well connected area in terms of other means of transport options, such as using the bike. As explained in chapter 3, the consequence of the bigger impact on migration is a bigger impact on housing prices.

Turning to model (4), with time trend, we see that the first order effect of PGC2 on log housing prices is now positive, but again at a diminishing rate. Evaluated within our data range we now find that the positive relationship holds until generalised time costs equals 318 minutes, with the maximum value of the PGC2 variable in our dataset being 331 minutes. This means that the positive relationship between PGC2 and log housing prices still holds true for (most) of our dataset. The diminishing marginal effect is illustrated in figure 6.3

Figure 6.3: Estimated Relationship Between PGC2 and lHp in Model(4)



In model(4) we know that the effect of a marginal change in the PGC2 on log housing prices is smaller, for higher initial levels of PGC2. This relationship between the initial level of PGC2 and the effect of a marginal change, is opposite of what we found in model(3). It could be argued that in areas where the generalised time costs of commuting to the central station is high, there are attributes in these areas other than the proximity to the city center, influencing the households migration decisions. Such attributes could be family relations and proximity to outdoors or recreational activities. This means that a marginal increase in the already high PGC2, will have a smaller effect on migration to the city compared to places with lower generalised time costs.

There are candidates to rationalise both increasing and decreasing marginal effects. Based on the results reported in table 6.7 we can not significantly identify neither increasing nor decreasing marginal effects, based on the variation available in the dataset. This leads us to favor the log specification. To test the log specification against the quadratic functional form we use the dynamic specification with dummy variables as in model(1) and (3), and perform a simple test for non-nested models as presented in Wooldridge (2013). This test is performed by estimating the comprehensive model:

$$lHp_{it} = \alpha_0 + \beta_1 lHp_{it-1} + \alpha_1 lunem_{it} + \alpha_2 lY_{it} + \alpha_3 lPGC_{it} + \alpha_4 PGC_{it} + \alpha_5 PGC_{it}^2 + a_i + u_{it} \quad (6.2)$$

Where both the log specification and the non-linear specification enters as special cases. The estimation results are shown in table 6.8:

Table 6.8: Comprehensive Model

	IHp
lag_IHp	0.448 (0.131)
lunem	-0.125 (0.0355)
IY	0.380 (0.193)
IPGC2	0.473 (0.186)
PGC2	-0.00435 (0.00397)
PGC2 ²	0.00000554 (0.00000805)
cons	-2.501 (2.166)
<i>N</i>	100
Year Dummies	Yes

Standard errors in parentheses

With a simple t-test, we can test if the quadratic specification as in equation (6.1) is the correct model by testing:

- $H_0^1: \alpha_3 = 0$
- $H_1^1: \alpha_3 \neq 0$

Similarly we can apply an F-test to test if the log specification as in equation (5.4) is the correct model by testing:

- $H_0^2: \alpha_4=0, \alpha_5=0$
- $H_1^2: \alpha_4 \neq 0, \alpha_5 \neq 0$

The results of the tests are shown in table 6.9:

Table 6.9: Test of Nonnested Models

Test IPGC2	
$H_0^1 :$	$\alpha_3 = 0$
t-obs=	2.54
$t_{crit}^{5\%} =$	1.99
Test PGC2, PGC2 ²	
$H_0^2 :$	$\alpha_4 = \alpha_5 = 0$
F-obs=	2.29
$F_{crit}^{5\%} =$	3,125

The test results show that we can reject hypothesis H_0^1 in favor of H_1^1 at a 5 percent significance level. We can not reject hypothesis H_0^2 at the same significance level. The tests therefore confirm that the log specification as in equation (5.4) is a better model than model (6.1).

6.5 Summary of Results

In this chapter we have estimated four static models using both yearly dummies and time trend as well as PGC1 and PGC2. Further, we have estimated four dynamic models using yearly dummies, time trend and log or quadratic specifications for PGC2. The analysis have shown that:

- In the static models we find a significant positive relationship between commuting costs and housing prices in the municipalities, using PGC2 as the generalised time costs variable.
- PGC1, which does not exclude stations within the municipality, has a smaller effect on housing prices than PGC2.
- In the dynamic models, we find a positive effect of commuting costs on housing prices at 10 percent significance level using year dummies and 5 percent using time trend. There is a significant lag in the adjustment of housing prices.
- We can not identify significant non-linearity between commuting costs and housing prices based on the variation in the data.

Chapter 7

Conclusion

In this thesis we have calculated a measurement of the year to year railway commuting time costs for Oslo, Kristiansand, Stavanger, Bergen and Trondheim in form of a generalised time cost variable. The generalised time costs variable's effect on housing prices has been modelled using a fixed effects model. To answer our hypothesis of an inverse relationship between railway quality for commuting and housing prices in the central municipalities, we estimate the effect of railway commuting time costs on housing prices during the years 1997 to 2016. We find that generalised time costs of railway commuting has declined in all cities in this period. The lowest average generalised time cost over our time span is found in Oslo, while Bergen has the highest. We also find that the average generalised time costs increases when the stations within our city of focus is excluded. Higher generalised time costs of commuting increases housing prices in the central municipalities, in line with our stated hypothesis and relevant economic theory. The full effect of one percent higher generalised time costs is estimated to be around 0.4 percent. This translates into an average monetary effect of around 7000 NOK per minute increase in generalised time costs. We find that the effect of generalised time costs on housing prices is smaller when including stations within the central municipalities. This substantiate the argument of contradicting effects by improving the commuting quality within the municipality border.

The analysis has focused on the inverse relationship between railway quality for commuters and housing prices. The mechanism is that improved railway quality leads people to consider living in peripheral areas more feasible. This reduces the population in the central municipalities and leads to a lower housing demand. With a fixed supply this leads to lower housing prices. A

continuation of this thesis could be to focus on migration developments, and use the generalised time cost variable to quantify these population changes.

Our generalised time costs variable is limited to the stations and time of day relevant for work commuters. It could also be argued that other purposes for a travel could pull in the same direction and give a similar effect. Better railway quality makes attributes such as cultural activities and big city services more available for people living in more peripheral areas. The generalised time costs variable is new for this thesis. Regardless of the time and work spent by the authors to make sure it is a precise measure, future work could be to make similar estimations and validate the computations.

In line with economic theory, unemployment in the municipalities is found to have a negative effect on housing prices, while gross income has a positive effect. This holds true in both the static and dynamic models. There is a significant lag in the adjustment of the Norwegian housing prices. About 40 percent of log housing prices today is explained by the log of housing prices last period. The results are not very sensitive between using yearly dummies or a time trend variable to control for common time factors. We can not identify potential increasing or decreasing marginal effects based on the variation in the data.

The findings of this thesis could be helpful in understanding current housing price developments as well as predicting future housing prices. It could also be useful with respect to urban city planning and developing public transportation strategies. In cases of housing price pressure due to migration in the central municipalities, improving the railway quality to and from the municipality could be seen as a demand reducing measure. This could be an alternative way to curtail the excess housing demand. The findings show that small changes to railway commuting quality can have an effect on central municipalities' housing prices. These changes need not be big capacity increasing infrastructure investments, but could simply be different allocation of resources within the current railway capacity.

Our analysis is constrained by limited data availability. When more data is available, further analysis could be able to not only yield more precise estimates, but also identify significant increasing or decreasing marginal effects. This could deepen the understanding of the effects of commuting costs depending on its initial levels. An extended analysis could include more time observations as well as more entities. The number of central municipalities in Norway is limited,

but an interesting analysis could be to collect data from several countries and perform a multilateral comparison.

Bibliography

- Anundsen, A. K. and Jansen, E. S. (2011). "Self-reinforcing effects between housing prices and credit, Evidence from Norway". *Journal of Housing Economics*, 13:3–20.
- Arestis, P. and Gonzales, A. R. (2013). "Modelling the Housing Price in OECD Countries". *International Review of Applied Economics*, pages 131–151.
- Baltagi, B. (2008). *Econometric Analysis of Panel Data*. John Wiley and Sons, New York.
- Black, D., Berger, M., and Scott, F. (2000). "Bounding Parameter Estimates with Nonclassical Measurement Error". *Journal of the American Statistical Association*, 95:739–748.
- Brueckner, J. K. (1987). "The Structure of Urban Equilibria: A Unified Treatment of the Muth-Mills Model". *A Handbook of Regional and Urban Economics*, 2:828–844.
- Brunborg, H. (2009). "Sentraliseringen fortsetter". *Valgaktuelt 2009*, 17/8:1–6.
- Durbin, J. (1970). "Testing for Serial Correlation in Least-Squares Regression When Some of the Regressors are Lagged Dependent Variables". *Econometrica*, 38:410–421.
- Eide, B. E. (2015). "Taxation of Housing. Killing several birds with one stone". *Statistics Norway; Discussion papers*, 829:4–22.
- Engebretsen, Ø., Vågane, L., Brechan, I., and Gjerdåker, A. (2012). "Langpendling innenfor intercitytriangelet" Institute of Transport Economics, report 1201/2012. pages 72–78.
- Fiva, J. H. and Kirkebøen, L. H. (2011). "Information Shocks and the Dynamics of the Housing Market". *The Scandinavian Journal of Economics*, 113:525–552.

- Geng, B., Bao, H., and Liang, Y. (2015). "A Study of the Effect of a High-Speed Rail Station on Spatial Variations in Housing Price Based on the Hedonic Model". *Elsevier, Habitat International*, 49:334–339.
- Grue, B., Langeland, J. L., and Larsen, O. I. (1997). "Boligpriser. Effekter av veitrafikkbelastning og lokalisering" Institute of Transport Economics, report 351/1997.
- Grøvdal, A. and Hjelle, H. M. (1995). *Innføring i transportøkonomi*. Fagbokforlaget Vigmestad Bjørke AS, Fagerdalen 6, 5035 Bergen-Sandviken.
- Hoechle, D. (2007). "Robust Standard Errors for Panel Regressions with Cross-Sectional Dependence". *The Stata Journal*, 7:281–312.
- Holvad, T. and Preston, J. (2005). "Road Transport Investment Projects and Additional Economic Benefits". *ERSA Conference Papers*, 5:4–19.
- Kulish, M., Richards, A., and Gillitzer, C. (2012). "Urban Structure and Housing Prices: Some Evidence From Australian Cities". *Economic Record*, 88:303–322.
- Larsen, E. R. (2002). "Searching for Basic Consumer Patterns; Is The Engel Elasticity of Housing Unity?". *Statistics Norway; Discussion Papers*, 323:3–11.
- Levin, A., Lin, C.-F., and Chu, J. C.-S. (2002). "Unit Root Tests in Panel Data: Asymptotic and Finite-Sample Properties". *Journal of Econometrics*, 108:4–14.
- Melby, I. (2007). Tax Statistics for Personal Taxpayers. URL: <https://www.ssb.no/en/inntekt-og-forbruk/statistikker/selvangivelse/aar/2007-11-29>. Visited 22/05-2017.
- Newey, W. and West, K. (1987). "A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix". *Econometrica*, 55:703–708.
- Norges Bank (2014). "Pengepolitisk Rapport 4/14" Norges Bank, report nr. 4/14.
- Norges Bank (2016). "Pengepolitisk Rapport 4/16" Norges Bank, report nr. 4/16.

- Norwegian Railway Directorate, Norwegian Coastal Administration and Norwegian Public Roads Administration (2017). "Nasjonal Transportplan for 2018-2029" The Norwegian Government, report nr. 5/4 2017. pages 1–36.
- Norwegian Welfare Administration (2017). Om statistikken: Arbeidssøkere. URL:<https://www.nav.no/1073745818/om-statistikken-arbeidssøkere?kap=6>. Visited 20/04-2017.
- Pagliara, P. and Papa, E. (2011). "Urban Rail Systems Investments: An Analysis of the Impact on Property Values and Residents Location". *Journal of Transport Geography*, 19:200–211.
- Ranheim, P. (2017). "Trenklin Versjon 3 Dokumentasjon og Brukerveiledning" Jernbanedirektoratet, Forthcoming 2017.
- Rødseth, A. (1985). *Konsumentteori*. Universitetsforlaget, Oslo, Norge.
- So, K. S., Orazem, P. F., and Otto, D. M. (2001). "Effects of Housing Prices, Wages, and Commuting on Joint Residential and Job Location Choices". *American Journal of Agricultural Economics*, 83:1036–1048.
- Sørbø, J. and Handal, J. (2010). "Nedgangskonjunktoren Til Nå-Hva skiller denne fra den forrige?". *Arbeid og Velferd*, 1:15–21.
- Transport Communications Ministry (2015). "Nye dobbeltspor gir bedre transport mellom byene" The Norwegian Ministry of Transport and Communications, report nr. 13/10 2015.
- Wooldridge, J. M. (2013). *Introductory Econometrics A Modern Approach*. SOUTH-WESTERN Cengage Learning, University of California, San Diego, Norge.
- Østby, L. (2001). "Norge i utakt med Europa". *Artikler og Publikasjoner SSB*, pages 1–6.

Appendix A

Additional Information

A.1 Estimation Gross Income 2016

$$GrossIncome_{i2016} = Grossincome_{i2015} * (1 + \beta)$$

$$\beta = \frac{1}{5} \sum_{t=2011}^{2015} \ln \left(\frac{Grossincome_t}{Grossincome_{t-1}} \right)$$

A.2 The Process of Making the Generalised Time Costs Variable

In this section we will explain in more detail how we computed the generalised time costs variable. Parts of the Trenklin III model relevant for the thesis are also summarised here based on the documentation report written by Patrick Ranheim.

A.2.1 Trenklin Transport Matrices

The algorithms in Trenklin calculate travel costs for all to-and-from combinations in the model. The travel cost concept in Trenklin is generalised time costs, ie all potential disadvantages with the travel weighted with the relevant time value. The relevant disadvantages used in the thesis are:

- Average cost of waiting time at the station
- Average cost of on board time

- Average cost of amount of train transfers
- Average cost of waiting time related to transfers

Average Cost of Waiting Time

Trenklin assumes that the travellers are indifferent between arriving before or after the most preferred point of time. Waiting time is the difference between actual and preferred arrival time, in other words hidden waiting time. As a random example; there will be people that wants to arrive the station at 09:40, but the best alternative is 09:45. This is a difference of 5 minutes that is a disadvantage for the traveller. The day-distribution gives the foundation for the distribution of the travellers and includes the travellers preferences for preferred arrival time.

In Trenklin III the time cost of waiting enters as an increasing, but diminishing function of the minutes spent on waiting. This is to capture the fact that the change in the disadvantage of waiting with a marginal increase in the waiting time should be smaller if the waiting time already is high. A traveller can easier find other things to do if waiting time is 40 minutes compared to if the waiting time is 10 minutes (for example read a book, go to a restaurant etc.) In Trenklin III, the average waiting time weight for the first 15 minutes is 1.8, from 15-30 minutes is 1.02 while 0.57 for 30-60 minutes.

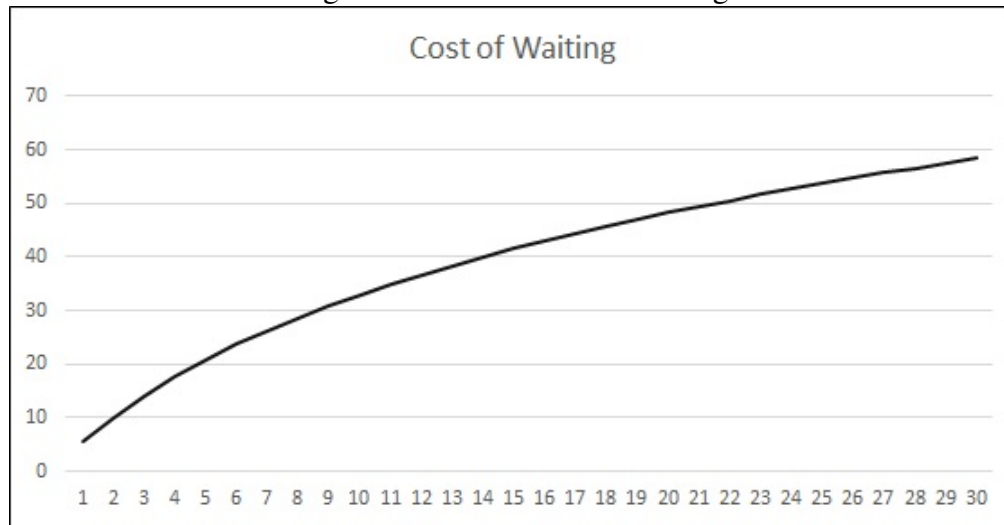
This is based on the following function:

$$U = v * (\ln(t + k) - \ln(k)) * \frac{timevalue}{60}$$

Where U is the waiting cost, t is the difference between preferred and actual arrival time, k is a constant(set to 10). Time value is set to 60 and v equals 30.

Figure [A.1](#) illustrates the increasing function of the cost of waiting.

Figure A.1: Time Cost of Waiting



An example from the computation of average waiting time cost on the stations between Oslo S and Fredrikstad is shown below. The highlighted number is the average of all the arrivals in Oslo S from Ski station in 2016.

Table A.1: Average Waiting Time on Station Oslo S-Fredrikstad 2016

	Oslo S	Ski	Moss	Rygge	Råde	Fredrikstad
Oslo S	0	2,736151938	3,716695458	6,306641399	6,523435145	6,669168068
Ski	2,555992761	0	6,416933774	7,584814004	7,444922911	7,46779108
Moss	3,948183509	3,790499193	0	6,287657181	6,500911896	6,650557805
Rygge	6,788507494	6,952218643	7,103965836	0	6,500911896	6,650557805
Råde	6,789735541	6,953472026	7,10660424	7,206988064	0	6,641779869
Fredrikstad	6,792558064	6,956023568	7,110444981	7,209806701	7,286623232	0

As the table shows, the average waiting time cost on the station when travelling from Oslo central station to Ski station is 2.5 minutes and the waiting time on Ski station when travelling to Oslo central station is 2.7 minutes.

Cost of On Board Time

On board time is the time spent on board the train. As underlined in the theory chapter a one minute longer on board time has a lower weight than waiting time on the station (up to 30 minutes) as it is more convenient to spend the minute on board the train than waiting on the station. The time

on board the train can easier be spent on other things. Commuters can for example use the extra minutes on work related pass-time.

An example of the computed on board time cost is illustrated below:

Table A.2: Average On Board Time Oslo S-Fredrikstad 2016

	Oslo S	Ski	Moss	Rygge	Råde	Fredrikstad
Oslo S	0	25,46975965	46,38062107	50,80441299	58,43430716	71,49736407
Ski	22,36886889	0	26,31498571	34,471179	42,67967666	56,52277049
Moss	44,34651581	23,54566266	0	7	14,65458875	27,70082514
Rygge	52,98082873	32,18193607	11,35009456	0	7,654588747	20,70082514
Råde	59,05708356	38,28333268	17,45837524	6,108653636	0	13,06391132
Fredrikstad	71,12960719	50,38001876	29,56136423	18,21211755	12,10507258	0

The matrix above shows the average on board time cost when travelling between stations in the south eastern part of Norway. From Oslo S to Ski the cost is 22.37 minutes, and from Ski to Oslo S the average is 25.47 minutes.

Average Cost Of Train Transfers

The time cost of transfers is the disadvantage of the transfer itself, in other words a fixed costs of transfers independent of the time spent on the transfer. The cost of a transfer is set to 10 minutes per transfer. An example of a transfer cost matrix from the year 2016 is shown below and the numbers are the time cost of the transfer.

Table A.3: Average Transfers Oslo S-Fredrikstad 2016

	Oslo S	Ski	Moss	Rygge	Råde	Fredrikstad
Oslo S	0	0	0	0	0	0
Ski	0	0	0	0,989783476	1	1
Moss	0	0	0	0	0	0
Rygge	0,003135435	0,001176662	0	0	0	0
Råde	0,003135435	0,001176662	0	0	0	0
Fredrikstad	0,003135435	0,001176662	0	0	0	0

Now, we can take look at the amount of transfers when travelling from Ski station to Fredrikstad station in the Østfold county. When travelling from Ski to Fredrikstad, there is an average of 0.00112 transfers. In the opposite direction, the average amount of transfers is 1 which then increases the generalised time cost with 0.01 and 10 minutes respectively.

Average Cost of Transfer Waiting Time

Transfer waiting time is the time spent related to a train transfer. A table of the average transfer waiting time cost is presented below.

Table A.4: Average Waiting Time on Transfers Oslo S-Fredrikstad 2016

	Oslo S	Ski	Moss	Rygge	Råde	Fredrikstad
Oslo S	0	0	0	0	0	0
Ski	0	0	0	32,80512199	33,36992021	33,70012724
Moss	0	0	0	0	0	0
Rygge	0,116011092	0,043536499	0	0	0	0
Råde	0,116011092	0,043536499	0	0	0	0
Fredrikstad	0,116011092	0,043536499	0	0	0	0

As the table presents, the average minutes of waiting when travelling from Ski to Rygge is 0.04 the average waiting time on transfers when travelling in the opposite direction is 33 minutes.

A.2.2 Interpretation of the Generalised Time Cost Matrix

Together, the different time components and their weights form a generalised time cost matrix. The matrix will have the same stations as the time component matrices from all the time tables within a certain year along the top rows and first columns. Rest of the cells are filled with the computed value of the generalised time cost of travelling between the stations in amount of minutes. The table (4.3) is a part of the spreadsheet from the year 2016 and shows the generalised time cost of travelling between different stations in the Østlandet region:

Table A.5: Generalized Costs Oslo S-Fredrikstad

	Oslo S	Ski	Moss	Rygge	Råde	Fredrikstad
Oslo S	0	37,32057666	61,64930863	73,0557633	81,04278359	94,40481529
Ski	33,64664207	0	48,83068504	91,33029464	100,1498868	114,4845768
Moss	60,43131017	39,10644495	0	29,22891532	37,2369679	50,58208563
Rygge	76,36317286	55,93563834	35,60707093	0	30,2369679	43,58208563
Råde	82,44087167	62,03891542	41,72015603	30,61911198	0	35,92979254
Fredrikstad	94,51798875	74,14002081	53,83007739	42,72703786	36,80766296	0

The upper row indicates the departure stations, and the left coloumn row is the destination. As an example, the generalized cost of travelling (monetary cost excluded) from Oslo Central Station

to Ski station in the Akershus county is 33.65 minutes. On the other hand, travelling from Ski to Oslo has a cost of 37.32 minutes. The generalized cost of travelling is different between two stations, dependent on which way the journey is.

A.3 Assumptions with Fixed Effects and Random Effects Models

Our analysis is based on the method of Ordinary Least Squares (OLS). This method estimates the coefficients by minimizing the sum of the squared residuals between predicted values and observed values. OLS will give unbiased and consistent estimates of the coefficients under the following assumptions¹;

- Linearity in the parameters

The true model in the population that we want to estimate can be written ;

$$y_{it} = \alpha_0 + \alpha_1 X_{it1} + \dots + \alpha_k X_{itk} + a_i + u_{it}$$

Where $\alpha_0, \dots, \alpha_k$ are the unknown parameters we wish to estimate and u_{it} is the unobserved random error.

- Random Sampling between the N entity observations (not necessarily over time)
- No Perfect Collinearity

There is no perfect linear relationship between the independent variables and no variable is constant.

- Zero Conditional Mean

The idiosyncratic error terms are on average unrelated to the explanatory variables for any time period.

$$E(u_{it} | x_{i1}, \dots, x_{ik}) = 0$$

¹Wooldridge 2013

- Constant variance (Homoskedasticity) and no serial correlation

$$Var(u_{it}) = \sigma_u^2, \sigma_u^2 > 0, \text{ finite}$$

$$Cov(u_{it}, u_{is}) = 0, s \neq t$$

The term a_i is an entity specific unobserved effect that does not vary over time. In our analysis this variable includes city specific effects like culture, location and political practice. These are effects that vary between entities, but are practically constant on short term. When these are uncorrelated with the explanatory variables we have what is called a random effects model(RE). In the case of correlation between the entity specific effects and the regressors we have a fixed effects model(FE). We assume further:

- Unrelated effects(RE not FE)

The entity specific variable a_i is a random variable which is uncorrelated with the explanatory variables at any given time.

- Constant variance of the entity specific effect(RE not FE)

$$Var(a_i|X) = \sigma_a^2$$

- Regressors and constant are not perfectly collinear, regressors have non-zero variance and not too many extreme values.

- Identifiability (FE)

No perfect linear combination between time-varying regressors, variance over time for a given entity is not zero and not too many extreme values

A.4 Unit Root Test

In testing for unit roots using Levin-Lin-Chu test for panel data we want to determine if any variable is integrated within each city. An integrated variable is non-stationary and must be differenced by the order of its integration. A variable that is integrated of order 1, I(1), must be differenced once to become stationary.

In the Levin-Lin-Chu framework we can consider the model:

$$y_{it} = \alpha_0 + \alpha_1 y_{it-1} + a_i + \epsilon_{it}$$

If α_1 equals 1 the variable is non-stationary and follows a unit root process. A simple t-test on α_1 is invalid because if $\alpha_1 = 1$ the test statistics does not follow a standard t-distribution. Subtracting y_{it-1} on both sides yields:

$$\delta y_{it} = \alpha_0 + \rho y_{it-1} + \alpha_i + \epsilon_{it} \quad (\text{A.1})$$

where $\rho = (\alpha_1 - 1)$. We can now test $H_0 : \rho = 0$ against $H_1 \rho \neq 0$. Failure to reject H_0 implies that we can not reject that $\alpha_1 = 1$ and that the variable is not stationary. Levin-Lin-Chu assumes that ρ is equal between entitites; $\rho_1 = \rho_2 = \dots = \rho_N = 0$. The results of the tests are shown below:

A.4.1 Log Housing Prices

Table A.6: LLC Unit Root Test for lHp
Levin-Lin-Chu unit-root test for lHp

Ho: Panels contain unit roots	Number of panels =5	
Ha: Panels are stationary	Number of periods =25	
AR parameter: Common	Asymptotics: N/T $\rightarrow 0$	
Panel means: Included		
Time trend: Not included		
ADF regressions: 1 lag		
LR variance: Bartlett kernel	9.00 lags average (chosen by LLC)	
	Statistic	p-value
Unadjusted t	-4.6400	
Adjusted t*	-4.2749	0.0000

Reject H_0 and conclude that log Housing Prices is stationary.

A.4.2 Log Unemployment

Table A.7: LLC Unit Root Test for lunem
Levin-Lin-Chu unit-root test for lunem

Ho: Panels contain unit roots	Number of panels =5	
Ha: Panels are stationary	Number of periods =21	
AR parameter: Common	Asymptotics: N/T \rightarrow 0	
Panel means: Included		
Time trend: Not included		
ADF regressions: 1 lag		
LR variance: Bartlett kernel	9.00 lags average (chosen by LLC)	
	Statistic	p-value
Unadjusted t	-7.0112	
Adjusted t*	-4.5522	0.0000

Reject H_0 and conclude that log unemployment is stationary.

A.4.3 Log Gross Income

Table A.8: LLC Unit Root Test for IY
Levin-Lin-Chu unit-root test for IY

Ho: Panels contain unit roots	Number of panels =5	
Ha: Panels are stationary	Number of periods =24	
AR parameter: Common	Asymptotics: N/T \rightarrow 0	
Panel means: Included		
Time trend: Not included		
ADF regressions: 1 lag		
LR variance: Bartlett kernel	9.00 lags average (chosen by LLC)	
	Statistic	p-value
Unadjusted t	-3.8009	
Adjusted t*	-3.6625	0.0001

Reject H_0 and conclude that log gross income is stationary.

A.4.4 Log PGC2

Table A.9: LLC Unit Root Test for IPGC2
Levin-Lin-Chu unit-root test for IPGC2

Ho: Panels contain unit roots	Number of panels =5	
Ha: Panels are stationary	Number of periods =20	
AR parameter: Common	Asymptotics: N/T \rightarrow 0	
Panel means: Included		
Time trend: Not included		
ADF regressions: 1 lag		
LR variance: Bartlett kernel	9.00 lags average (chosen by LLC)	
	Statistic	p-value
Unadjusted t	-5.1343	
Adjusted t*	-2.5870	0.0048

Reject H_0 and conclude that log PGC2 is stationary.

A.5 Durbin's Test for Serial Correlation With Strictly Exogenous Regressors

Estimating 5.1 allows us to obtain estimated residuals. In order to test the presence of serial correlation in the error terms we can run the regression of $u_i^{hat}{}_t$ on $u_i^{hat}{}_{t-1}$:

$$u_i^{hat}{}_t = \rho u_i^{hat}{}_{t-1} + e_{it}$$

To test for AR(1) serial correlation we can test the $H_0 : \rho = 0$ against $H_1 : \rho \neq 0$. Rejecting H_0 implies presence of serial correlation.

Table A.10: AR(1) Serial Correlation Test

	(1) u_hat
u_hat_1	0.918*** (18.94)
_cons	0.000239 (0.03)
<i>N</i>	94

t statistics in parentheses

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

. test ρ

$H_0: \rho = 0$

$H_1: \rho \neq 0$

F(1, 92)= 358.83

Prob >F = 0.0000

We reject H_0 and therefore find evidence of AR(1) serial correlation.

A.6 Durbin's Test for Serial Correlation Without Strictly Exogenous Regressors

Estimating 5.4 we obtain estimated residuals. In order to test for serial correlation we can run the regression of $u_i^{hat}_t$ on all dependent variables as well as $u_i^{hat}_{t-1}$:

$$u_i^{hat}_t = \alpha_1 lHp_{it-1} + \alpha_2 lY_{it} + \alpha_3 lunem_{it} + lPGC2_{it} + \rho u_{it-1}$$

To test for AR(1) serial correlation we test $H_0 : \rho = 0$ against $H_1 : \rho \neq 0$:

(1)	
u_hat	
u_hat_1	0.514** (3.41)
lag_IHp	0.0474 (0.44)
IY	-0.188 (-1.13)
lunem	0.0230 (0.83)
IPGC2	-0.000787** (-3.15)
time	0.00326 (0.48)
.cons	2.166 (1.23)
<i>N</i>	89

t statistics in parentheses

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

. test ρ

$H_0 : \rho = 0$

$H_1 : \rho \neq 0$

F(1, 82)= 11.60

Prob >F = 0.0010

We reject H_0 and therefore find evidence of AR(1) serial correlation.

A.7 Testing Hypotheses About a Single Population Parameter: The t-test

The statistics we use to test an hypothesis about the effect of our observed coefficients is the t-test. The most used hypothesis is: $H_0 : \beta_j = 0$. We want to check if our independent variable has any effect on our dependent variable. The t-value is then measured up against our critical values of t-distribution.

$$t_{\hat{\beta}_j} \equiv \frac{\hat{\beta}_j}{se(\hat{\beta}_j)} \quad (\text{A.2})$$

Although $H_0 = 0$ is the most common hypothesis, we sometimes want to test whether β_j is equal to some other given constant. Two common examples are $\beta_j = 1$ and $\beta_j = -1$. Generally the formula becomes:

$$t_{\hat{\beta}_j} \equiv \frac{\hat{\beta}_j - a_j}{se(\hat{\beta}_j)} \quad (\text{A.3})$$

Where a_j is our hypothesized value of β_j .

