# House Price Dynamics and Traffic Mode Choice: Three Essays in Real Estate and Urban Economics

Inaugural-Dissertation

zur Erlangung des Grades Doctor oeconomiae publicae (Dr. oec. publ.)

an der Ludwig-Maximilians-Universität München

2006

vorgelegt von

Thomas Maier

Referent: Prof. Sven Rady, Ph.D.

Korreferent: Prof. Ray Rees

Promotionsabschlussberatung: 07. Februar 2007

*To Tessa*

# Acknowledgements

First and foremost I would like to thank my supervisor Prof. Sven Rady. He steadily encouraged me during the development of this thesis, gave very valuable comments on all parts of this thesis and showed patience with me and my work in all stages of this dissertation. Thanks to his helpful guidance, the process of turning a mathematician to an economist was successfully completed. I am grateful to Prof. Ray Rees and Prof. Peter Egger who agreed to serve as members in my thesis committee.

I am also indebted to my colleagues from the Munich Graduate School of Economics, the Faculty of Economics and the Seminar of Dynamic Modelling. Especially, I thank Christoph Hofmann, Oliver Nikutowski, Tu-Lam Pham, Markus Reisinger, Ludwig Reßner and Ferdinand von Siemens for their helpful comments, suggestions and corrections on different parts of this work. Furthermore, I thank Gaurav Anand, Elmar Becker, Simon Gamperl, Florian Neubert, Thomas Nitschke and Georg Steinlein for proof-reading parts of this document.

Andreas Ertl and Konstantin Scholl shall be mentioned for improving some graphic elements of this thesis.

I would like to thank Ingeborg Buchmayr and Manuela Beckstein for providing an excellent working environment at the MGSE and at the Chair of Prof. Rady. The professional hard- and software service of our computer expert Dirk Rösing deserves to be positively alluded to here as well.

I would like to mention here as well the great hospitality of the University of Cambridge, Faculty of Economics, and Darwin College during my stay there in the year 2004.

Financial support from the DFG, the GRK 801 and SFB Transregio 15 is gratefully acknowledged. Finally, I thank my parents for their steady support in all matters of life. This doctoral dissertation is dedicated to Tessa, for keeping me smiling throughout.

# Contents

# List of Figures

# Chapter 1

# Preface

This thesis consists of three self-contained essays on urban and real estate economics. Chapter two studies the problem of traffic mode choice within an urban area in a spatial context. We examine and compare the effects of road capacity reduction and tolling schemes with respect to the urban traffic distribution and derive their welfare implications. Chapter three studies the consequences of the selection problem landlords face in order to attract tenants of appropriate quality. We show that rental prices exhibit an overshooting pattern after exogenous shocks to the cost structure of landlords. Finally, chapter four generalizes the model of Ortalo-Magné and Rady (2006) which considers the dynamics of house prices in the market for privately owned dwellings after exogenous income shocks. By using general utility functions, we show how the effect of consumption smoothing over the life cycle changes the relative prices and the transitional dynamics. Furthermore, the effects of transaction costs on relative prices are considered.

The general field of urban economics studies the interaction of households and firms within an urban area by the help of microeconomic theory. Location decisions, externalities and urban transport are some of the most important topics covered. While standard economic theory does not account for spatial relationships between organizations and individuals, urban economics incorporates these spatial relationships in order to understand the economic structure underlying the functioning, formation, and development of cities.

Considering the theory of urban transport, there are two main strands in the literature. The first strand focuses on the so-called two-mode problem. The seminal paper that studies this kind of traffic interaction is Wardrop (1952). In his model, there are two traffic links in the economy, both connecting point $A$ and point $B$. Agents have to decide which link to use. In the Wardropian equilibrium commuters continue to enter each line, until total costs including congestion costs of each line are equal. Depending on the characteristics of the two lines, traffic participants distribute themselves on the lines in such a way that in equilibrium agents are indifferent between both links. There is a large body of existing literature studying some modifications of this type of problem, see for example Arnott and Yan (1995) for a thorough overview. Strotz (1965) was the first to include user heterogeneity in the two-mode setting. In his model, agents from different groups have different benefit and average cost functions. Further, fruitful contributions on the optimal capacity in road networks are for example Wheaton (1978) and Wilson (1983). They study how optimal road capacity is changed when auto congestion is unpriced or underpriced. This whole body of literature assumes that marginal congestion costs are independent of the location of agents. However, it seems that road congestion decreases in the distance to the city centre - a fact which cannot be accounted for in models without spatial structure. Although it is possible to derive the optimal toll and capacity values of the two links in models of this kind[1], a problem arises when one wants to determine the optimal capacity dependent on location or when the optimal position of tollgates shall be determined.

In contrast, the second line of research considers the spatial structure of the urban area. Most of this literature makes use of the so-called monocentric approach. Alonso (1964) was the first to model a city as a circular disc with a central business district and a surrounding residential region. Many important contributions are based on this idea (e.g. Vickrey (1971), Solow (1972), Arnott (1979), Verhoef (2005)).[2] Recently,

[1] See for example Arnott and Yan (2000).

[2] It is generally recognized that the mono centric paradigm may no longer be universally valid. Several explanations for polycentric expansion have been proposed and summarized in models that

among others, Lucas and Rossi-Hansberg (2002) have generalized the monocentric city approach. They endogenize the emergence of business districts and housing areas by explicit consideration of agglomeration forces. In this equilibrium business districts and housing areas can both be located anywhere in the city. The literature of this kind either does not account for optimally designed traffic regulation schemes or does not consider the consequences of traffic mode choice within an urban area.

In chapter two, we construct a model which merges both strands of the literature described above into a spatial model of traffic mode choice and congestion regulation.

In our setting, we study a version of the two-mode problem in a spatial context. Agents who want to reach the centre are placed in the concentric area around the centre and differ with respect to their location - a feature that is in line with the monocentric city approach described above. The two transport lines can be considered as (congested) road and (uncongested) public transport. Agents' decisions between an uncongested link and a congested link depend on their location and the lines' cost characteristics.

In this setting, we show that the resulting urban traffic distribution has a unique equilibrium. Furthermore, in this equilibrium key features of actual urban traffic distributions are met.

With regard to traffic regulation, our main result states that it is optimal to point-wise[3] reduce the capacity of the congested line at certain positions when the possibility to toll the congested line is not given.[4] The intuition for this result is the following: By positioning the capacity reduction optimally, we distort only those traffic participants far away from the centre. They contribute the least to welfare on average but create the highest congestion externalities. After the capacity reduction, they either refrain from the trip to the centre at all or they use the uncongested line. In both cases they

account for factors such as utility gains from lower average land rents and increasing (or constant) returns due to economies of agglomeration.

[3]We model the capacity reduction by an artificial bottleneck which means that all agents who pass a certain point face additional costs, for example, due to the queueing time behind the bottleneck.

[4]Concerning the political feasibility of road pricing see for example Verhoef (1995).

do no longer contribute to congestion.

A further result is that the optimal distance to the centre of the capacity reduction is negatively correlated to its severeness. The intuition is that all road users who pass this position are distorted if the magnitude of the capacity reduction is chosen relatively low. However, only a few of them actually refrain from using the road. Therefore it would be better to position it farther away from the centre, such that the number of distorted road users who use the congested line anyway is minimized.

We also study the optimal number and position of tollgates as well as the optimal toll. It is found that the tollgate's optimal distance to the centre is decreasing in the toll. The reason for this effect is the maximization of toll income as we will explain in more detail in chapter two.

Interestingly, numerical results show that the optimal number of tollgates is two, whereas the optimal number of pointwise capacity reductions is one. With regard to tolls, the trade-off is to minimize congestion externalities and maximize user benefits but at the same time to maximize toll income. It turns out that this is achieved by a scheme of two tolls.

Chapters three and four of this thesis study the price dynamics on the real estate market.

Generally spoken, there exist at least two logic linkages between our model of urban transport in chapter two and real estate prices which we consider in chapters three and four. The first is that the savings in average commuting costs due to the improvement of the traffic system are capitalized into housing values, as indicated by empirical studies (Bajic (1983) and Henneberry (1998)). The second interrelation is based on the spatial structure within a city. Neglecting the possible negative externalities[5], areas close to public transport stations with connection to the city center exhibit higher property prices. In our model of urban traffic mode choice this fact is replicated by higher benefit values of households living close to public transport lines. Furthermore, we show that in equilibrium the commuter belt - which is the area around a city where

[5]Such negative externalities could be noise or pollution.

commuting takes place - is star shaped. This influences the spatial distribution of real estate prices as well.

Economic theory defines real estate markets by their salient characteristics. These characteristics include, for example, the property of indivisibility. This means that typically, it is not possible to rent, buy or sell less than the whole dwelling. A second important property is given by the fact that actors on the real estate market face high transaction costs including legal fees, search costs and moving costs. For example, the process of acquiring a house is much more expensive and time-consuming than most other types of transactions. Furthermore, real estate represents both a consumption and an investment good. In the market for privately owned real estate, properties are bought with the intention of using it (consumption good), or with the expectation of attaining a return (investment good), or both. This combination of consumption and investment properties, together with the fact that for most households real estate is the main investment during their life-cycles, leads to a situation where past prices in the market may have influence on current prices. The intuition is the following: Some buyers of real estate are already owners of a typically smaller entity of housing and try to climb up the property ladder. But due to borrowing constraints they have to sell the house they already own in the first period to get enough cash to finance the new dwelling in the second period. So the real estate prices before the actual time of buying the house determine how much they are able to invest in the new real estate and therefore determine their demand.

In contrast to standard microeconomic theory, housing markets show several odd results: Empirical studies prove that in many housing markets transaction volume moves with prices. This fact seems counterintuitive at first glance and can not be convincingly explained using standard theory. Although the depreciation rate of real estate of any kind is very low and the supply is almost fixed in short time, a high price volatility in real estate markets, especially in the market for privately owned dwellings, is observed.

In chapter three we study the price dynamics in the rental housing market after shocks to the cost structure of landlords. The rental housing market has a long tradition as a subject of interest to economists. Especially the imposition of rent control mechanisms, price ceilings and legal restrictions on evictions are salient features of this market which have found their place in the corresponding literature (for an overview see Bailey (2000) or Lind (2001)). While this literature concentrates mainly on the welfare effects on the economy, we study the pricing behaviour of landlords and the resulting transitional dynamics after a shock to their cost structure. In our model the assumption that rental prices cannot exceed a certain value is made for analytical convenience. However, we consider the costs of landlords associated with the process of attracting a new tenant, such as agency costs or legal fees in the case of eviction. These costs are strongly influenced by legal regulations.

A second branch of the literature on rental housing markets focuses on the contractual relationship between landlords and tenants. Among others, Read (1991, 1993), Hubert (1995) and Basu and Emerson (2003)[6] construct models of asymmetric information in order to explain the functioning of the rental housing market. These models construct situations where either moral hazard or adverse selection comes into play and influences the market outcome. Although we share some insights with these models, we do not need any information asymmetry in order to achieve our results. In our model, landlords are able to screen tenants perfectly and landlords are alike with respect to their quality.

Another stream in the literature highlights the searching aspect in rental housing markets. Salient representatives are Read (1988, 1997) and Wheaton (1990). In our model, we abstract from any problems arising in search models, such as the probability of not being matched. However, we model the demand side of the model by a function which is constructed as if it were the outcome of a search model.

Our stylized setting of the rental housing market comprises heterogeneity on both

[6]Interestingly, Basu and Emerson (2003) also formalize the idea of an efficiency rent, but in a different context. In their model, strict regulations imply that rental prices cannot be adjusted within the duration of a rental contract. If there is any inflation in the economy, the rental price is held below the market price in order to attract a tenant with *high* probability of moving out and therefore with low expected time in the dwelling. The reason is that by doing so, the landlord is able to adjust the rental price according to the inflation.

sides of the market. Landlords differ in their cost characteristics and tenants vary in quality which is represented by the probability of moving out. The rental price charged by a landlord determines the corresponding quality of the tenant he is able to attract. This structure enables us to formalize the intuition of efficiency rents[7]: Dependent on their costs, landlords set the rent that enables them to attract a tenant of appropriate quality.

In steady state, the vacancy rate is determined by the prevailing distribution of tenant qualities and it influences the relative market power of both sides on the market. We show that average rental prices tend to overreact after shocks to the cost structure of landlords. For concreteness consider the situation of increased costs. Then, we identify two effects.

The first effect influences the average price directly. After a shock to tenant switching costs, landlords increase their desired tenant quality. This corresponds to a decreased rental price quoted by each landlord.

The second effect concerns the vacancy rate which is determined by the actual distribution of tenant qualities. If more tenants of high quality are actual tenants, the vacancy rate is relatively low. This increases the relative market power of landlords to tenants and therefore tends to increase the rental price.

We show that under appropriate assumptions the first effect is stronger than the second one. However, the mitigating second effect only manifests itself in the long run, since the vacancy rate does not adapt immediately after a uniform shock to the cost structure of landlords. Directly after a shock to costs, the old vacancy rate prevails and the second effect does not come into play. Only in the long run the vacancy rate adopts a level which can mitigate the strong first effect. In summary, we observe an overshooting pattern of rental prices, caused by the fact that the equilibrium vacancy rate adjusts only slowly to its steady state level.

As noted above, an interesting property of housing markets is the high price volatility and the fact that transaction volume moves with prices. Concerning the correlation

[7]See Stiglitz (1974) for their counterpart in the theory of labour markets.

of transaction volume and prices several explanations can be found in the literature. Genesove and Mayer (2001) use prospect theory in order to explain this effect. Yet they rely on loss aversion and thus can only account for the fact that owners of real estate do not sell at market prices in times of recession. An approach which makes use of credit constraints is Stein (1993). In this model, multiple equilibria and multiplier effects due to down payment restrictions are generated in order to create the positive correlation of turnover and real estate prices. This model also serves as an explanation for the high volatility in the market. Ortalo-Magné and Rady (2006) build on this intuition and construct an overlapping generations model in order to study the interaction of different age cohorts in the market. They generalize the model of Stein in some aspects and show that the income of the young is the crucial driver of house prices in the presence of credit constraints. Further, an income shock to all age cohorts triggers an overshooting of housing prices.

Ortalo-Magné and Rady (2006) assume a linear utility function for numeraire consumption and zero transaction costs. These features keep their model analytically tractable but lead to unrealistic results of zero consumption of the numeraire good in all periods of life except in the last one.

In chapter four, we generalize the model of Ortalo-Magné and Rady (2006) by introducing transaction costs and weakening the restrictive assumption of linear consumption utility. By introducing the more general class of concave utility functions it is possible[8] to get positive consumption in any age period. Since the model is no longer analytically solvable under these more general assumptions, we make use of numerical simulations to derive our results.

We show analytically what the qualitative housing distribution looks like in the generalized case and provide upper and lower bounds for property prices. Furthermore, we derive sufficient conditions which determine the behaviour of marginal property buyers. In one case they strictly prefer the property to an increase in their consumption. In this case, their decision is driven entirely by credit constraints. In the other case,

[8]Not all concave consumption utility functions do allow for strictly positive consumption in equilibrium.

they are indifferent[9] between an increase in consumption and the property.

We also provide conditions which lead to the other extreme case of consumption behaviour. In this case, after potential housing transactions, all remaining wealth is consumed immediately. These conditions are met if the income rise is high enough in the first periods of life. Then property purchases increase future wealth because of the housing wealth transfer into the subsequent period. Hence, property buyers consume less numeraire good in the period of the purchase, but more in the subsequent period.

The numerical results show that concave utility functions exhibit the same overshooting pattern of prices as the linear specification does. However, the overshooting effect is reduced quantitatively. The reason is that the additional income by capital gains is not completely invested in real estate, but many households prefer to increase their consumption instead.

Furthermore, our numerical results indicate that the steady state effect to relative house prices is even higher in the case of a concave utility than in the linear case. The reason is that changes in the income structure implicitly change the willingness to pay for a property due to the changed marginal utility of consumption.

Concerning the effects of the introduction of transaction costs on steady state prices we find that flat prices are lowered in proportion to transaction costs, whereas house prices are changed overproportionally. Even in case of linear consumption utility it is now less attractive to climb up the property ladder. Households tend to stay in the flat they already own in order to avoid transaction payments.

In chapter five, we draw a conclusion by providing some additional comments and a discussion of the advantages, limitations and possible policy implications of the results derived in this thesis. Finally, we outline directions for future research.

[9] This indifference condition also has to account for changes in future consumption.

# Chapter 2

# Traffic Mode Choice and Optimal Congestion Regulation

## 2.1 Introduction

Participating in urban road traffic causes manifold positive and negative externalities. Examples of the latter are air pollution, risk of accident and the congestion externality.

Economic theory provides a simple answer to negative side-effects: Implementing a Pigouvian tax where the optimal tax rate is exactly the difference of the marginal social costs and the marginal private costs, evaluated at the social optimum. Nevertheless, the heterogeneity of traffic participants with different cost and benefit structures often makes it difficult to implement such a tax system in a tractable way. In this chapter we contribute to the theory of this problem.

Consider a situation typically occurring during the commuting peak hour. All traffic participants want to reach the central business district which is located in the centre of the city at the same time. Agents face a typical two-mode-problem.[1] They choose between driving on the road which is considered to be a congested link[2] in the model,

[1]There exist various interpretations of different traffic modes. When considering freight traffic, the congested line could be the road, and the uncongested line could be rail transport. When considering commuting in a car-free area the pair of modes could be cycling on the road and using the train.

[2]In the following sections we use the terms congested link/line and road/car interchangeably. The

or using public transport which is represented by an uncongested link. However, road users do not account for the negative externalities they exert on others by contributing to congestion. In this model we study the optimal spatial design of tolling schemes in urban areas. Furthermore, we consider the situation when tolls are not applicable and show that welfare gains can be achieved by an appropriate road capacity reduction.

In our model, agents are located in a monocentric city and differ only with respect to their location. All agents want to reach the centre and thereby face a typical network trade-off. By assuming that their benefit from the trip to the centre is independent of location, agents' decision problems only depend on the corresponding cost structure. We classify the different costs into four types: Free flow costs are associated with the uncongested use of a line. Congestion costs do not only depend on the number of new entries to the congested link at a certain distance, but also on the number of entries at any other position on the road. In order to capture the agents' spatial differences in location, we also consider reaching costs which are caused by the trip to the desired link. Finally, the participation in road traffic causes fixed costs. In our model, it is crucial that the marginal reaching costs for the road are lower than the marginal reaching costs of the public transport line.[3]

In equilibrium, agents will distribute themselves on the two links in such a way that for each agent the individual cost of traffic is minimized.

A salient feature of our model is the interaction of users of a congested line in case of an uncongested alternative: starting at the maximal distance and moving to the centre, an increasing number of agents will have chosen the congested line which leads to a (weakly) decreasing function of users of the congested line in distance to the centre. Therefore the marginal congestion costs are at a maximum at the centre, whereas at the border of the urban area they are minimal. Under appropriate conditions, there

same applies for uncongested link/line and public transport/train.

[3]This assumption is realistic. We do not consider switching traffic modes. Then almost by definition reaching the closest road to the centre by car is on average not as time consuming as travelling to the closest public transport station without a car. Furthermore, the road network is much denser in all spatial dimensions, such that reaching the optimal "road connection" should be faster than reaching the closest public transport station.

exists a certain distance from the centre that marks the inner region. In this inner region, it is always optimal to choose the uncongested line, due to high congestion costs and the fixed costs of the congested line. In this area, there are no "new entries" into the congested line and therefore the marginal congestion costs stay constant.

Our first contribution is to identify conditions under which we can match some key features of the observed distribution of road traffic within an urban area. It is an empirical fact that the farther away a commuter lives from the centre the higher is the probability that he owns a car and uses it for daily commuting trips.[4] Gleave (2002) studies the commuter traffic distribution within the city of London and finds that the ratio of the number of residents who use a car and the number of residents who use rail lines is nearly 1:1 in outer London while it becomes nearly zero (less than 1:20) in inner London. Ong (2004) analyses the traffic network density distribution and the traffic volume distribution within the city of Los Angeles. He finds that the downtown area has more than twice the number of vehicle miles per square mile than the city average. Furthermore, he shows that the average commuting distance of inner-city residents is much smaller than the commuting distance of residents on the outer ring. Consequently, large parts of car traffic in the city centre must stem from non-inner-city residents.

These findings can be matched in the present model: There emerges an inner city where residents use only public transport, and all traffic is generated by outer-city residents. Furthermore, traffic volume is maximal in the city centre and decreases continuously towards the centre.

The main intuition is the following: Consider the situation around the centre. The costs for reaching a public transport line are very small because of the spatial monocentric urban structure. An agent choosing the congested line faces high marginal congestion costs and has to bear the fixed costs as well. In contrast, an agent choosing the uncongested line faces no congestion costs and no fixed costs. Therefore, around the centre, an agent will opt for the uncongested line. Now consider the situation in the suburbs: The agent choosing the congested line will face low marginal congestion

[4] See e.g. Fujiwara et al. (2005)

costs at the beginning of his trip and high marginal congestion costs at the end, leading to medium congestion costs for the whole trip. Because of higher overall costs for the trip, the fixed costs the agent has to bear lose importance. But because of relatively high distances to the links reaching costs become more dominant. Since we assume that the uncongested link exhibits higher marginal reaching costs than the congested line, more and more agents will opt for the congested line with increasing distance to the centre.

Our second contribution is our main result. By abstracting from any other heterogeneity within the group of traffic participants, we are able to design simple and tractable location-specific road capacity schemes in order to improve total welfare of the network. First, we prove that the resulting equilibrium is inefficient. Then, we show how capacity reduction at certain locations leads to overall welfare improvements when the implementation of tolling mechanisms is not possible. We model capacity reduction as the installation of an artificial bottleneck which leads to further time costs for the potential road user. Examples[5] of these artificial bottlenecks could be traffic lights with extra long red phases or more generally a setup of corresponding traffic lights which creates situations where traffic flow is minimized because car drivers have to stop successively at all traffic lights.

Again, the intuition is easy to understand: By forcing potential long-distance commuters to pass the artificial bottleneck on the road at the beginning of their trip they will either refrain from the trip or will switch to the uncongested line. In both cases, they will no longer contribute to congestion. In summary, the effect of congestion reduction for all road traffic participants should outweigh the negative distortion for the agents at outer locations. Therefore, the trade-off for the optimal position of the artificial bottleneck is the following: By shifting its location towards the centre, the utility gain from the reduction in congestion costs increases. At the same time, though, there is an increasing utility loss by the fact that more and more (potential) road users have to pay the additional "bottleneck" costs.

The third contribution is to examine the allocative and distributive consequences

[5] Other examples are significant speed limits or road works with extensive barriers.

of location-dependent distortion schemes including tolling. In general, tolling schemes differ from distortion by capacity changes by the fact that road users have to pay a certain fixed toll when passing the tollgate. In contrast, the additional costs associated with a bottleneck depend on the number of road users at the position of the bottleneck.

We show that the optimal toll is a falling function in the tollgate's distance to the centre. In comparison with the situation without distortions, we find that at all distortion schemes, road traffic participants with a relatively high distance to the centre lose most, and road traffic participants with a location closer to the centre are the main winners. Users of the uncongested line gain only by the distributive effects of tolls. In the case where the number of toll points is not restricted to one, we show numerically that it is never optimal to have more than two tollgates. This result is due to the fact that (potential) road users have to pass all imposed toll gates successively.

We shall proceed as follows. An overview of the related literature is given in section 2. In section 3, we present the model framework. In section 4, we describe our equilibrium concept formally, and prove some results on the benchmark model. Our main results are presented in section 5. Finally, we conclude in section 6.

## 2.2 Related Literature

We are not the first to study second best efficiency of imperfect substitutes in urban traffic networks. In our work, we meld the theoretical foundations of different strands in the corresponding literature.

By using the monocentric city structure we rely on already well-developed literature initialized by Alonso (1964). Building on his model, the urban economic theory of transportation and land use (Vickrey (1971), Solow (1972), Kanemoto (1976) and Arnott (1979)) developed in the seventies generalized the theory of agricultural land rent and land use to the urban case. The main aim of this theory is to describe the different types of land use and their distribution in equilibrium.[6] Mirless (1972) showed

[6]Some important results from the analysis are that land rent, housing rent and housing density are monotonically decreasing with distance to the centre.

in the basic type of model that the resulting equilibrium density patterns are pareto-optimal. This is mainly due to the fact that there are no congestion externalities, in contrast to our work. With very few exceptions (as Verhoef (2005)), their main focus is on optimal land allocation, land rents, densities and city size. However, none of them considers the problem of traffic mode choice within an urban area as we do. This strand of the literature shares two important features with our model: First, the user cost of travelling depends on a location-dependent capacity term (which is defined in models of this kind as the amount of land allocated to road use). Secondly, this literature makes use of a circular monocentric urban area with a central business district located in the middle of the city.

By modelling the choice between different lines of transport, our work is related to the so-called two-mode problem. This literature considers two traffic modes which connect two different points, as first described by Wardrop (1952).

Many papers of this line of research give similar intuitions about the road pricing as we do. Rouwendal and Verhoef (2004) study second-best pricing mechanisms as they arise through the failure of tolling a link. In their model, they also allow the links to be imperfect substitutes. Arnott and Yan (2000) examine a classical two-mode problem where both lines are congestible but only one link can be tolled. They study whether the capacity of the links should be increased or decreased and identify the optimal toll. De Borger, Proost and Van Dender (2005) study tax competition in a congestible two-link network. In their parallel-road network model, different governments have tolling authority. Although their paper is mainly designed to study transit traffic on highways, it gives a good heuristic insight in possible effects on our model when relaxing the assumption that both possible lines are managed by the same operator. Kanemoto (1999) examines the second-best optimal capacity of the uncongested line. There cost-benefit analysis is the means to provide optimal investment conditions in the long-run.

Wheaton (1978) and Wilson (1983) study optimal road capacity with suboptimal congestion pricing. Although some of the intuitions of their results carry over to our model, in their work only the congested mode of transport is taken into account and they do not account for the spatial structure of the problem.

All these models have in common the simplification that marginal congestion costs are location-independent, whereas in our model the location-dependent congestion cost function is the main driver of our results. Moreover, this literature accounts for the problem of traffic mode choice, but does not integrate this problem into a spatial monocentric structure.

A third stream in the literature our work is related to is represented by Vickrey (1969). He established the so-called bottleneck model, where he could incorporate trip-timing decisions in the canonical model. This theory has been elaborated on and extended in several directions. One contribution which shows some similarities to our work is Arnott (1998). In his paper he incorporates a discrete spatial structure into the standard framework and examines the consequences of tolls in this setting. However, in contrast to our model, in his work no continuous heterogeneity of traffic participants is given. The bottleneck approach turned out to be consistent with the macroscopic approach concerning the marginal external congestion cost and with empirical results as well. Nevertheless, we abstract from any timing decisions in our model and hold our model static in this aspect. Vickrey also presented the standard type of congestion function which in the most simplified way can be defined as a function of the ratio of the number of traffic mode participants and the capacity of the traffic mode.[7] There, capacity is assumed to be a parameter influencing the congestion cost function. Referring to car traffic one would for example define capacity as road width. In public transport, one can interpret capacity as the number of maximally possible passengers per time. In contrast, we argue that we have to consider the spatial structure of the urban area as well. By introducing reaching costs, we replicate the intuition that the farther the commuter lives from the centre, the more time he spends on average to reach the next public transport station. Like most of the existing literature, we deal with a partial equilibrium model and do not consider any other distortions to the economy. This simplifies the structure considerably and helps us understand the intuitions.

[7] The general form of the Vickrey congestion cost function is $c(\frac{N}{r}) = a + b(\frac{N}{r})^d$, where $N$ is the number of road users, $r$ is referred to as road width or capacity and $a, b, d$ are non-negative constants.

From a mathematical point of view, Braess (1967) shows with the help of graph theory that under certain circumstances, an increase in capacity of a road network may lead to an increase of travel times for consumers. This paper is related to our work in the sense that a change in the capacity does not lead to the result which one would expect at a first glance, but exactly to the opposite. However, the setting of his traffic system is completely different to ours. In particular, he does neither consider a monocentric city structure, nor does he account for the possibility to choose between different traffic modes.

Some alternative yet practical strategies to improve welfare in urban transport systems are described in Schley (2001). In a more general context, Arnott, Rave and Schöb (2005) present some new strands in the theory of urban traffic congestion. Moreover, they provide a good overview of the existing literature on this topic.

## 2.3 The Model

Our aim is to construct a model which enables us to analyse the interaction of traffic participants entering into a congestible road at different starting positions. Furthermore, we want to study the allocative effects of tolls and artificial bottlenecks on the traffic system. By considering heterogeneity of the agents with respect to their location, we bring these interactions into a spatial context.

**Economic Environment.** In the following, we describe the formal model.

**The Transport System.** The urban area is represented by a circular disc of radius $\ell$. There are congested lines and uncongested lines, all leading to the central business district which is located in the middle of the disc.[8] The uncongested links are represented through $q$ lines which are placed symmetrically in the circular disc, crossing the outer circle line orthogonally such that every line leads directly to the

[8]See Anas, Arnott and Small (1997) for a critical discussion of the monocentric city approach.

centre.[9] The congested lines are given in a sufficiently high number, allocated in the circular disk. As we will see later, all our results are invariant to their exact location and number. Hence there is no need to specify them in detail.

**Population.** We assume a population density of 1 for all points in the city area. This leads to a total mass of population of $\pi \cdot \ell^2$. The urban area is assumed to be closed and static. It is closed in the sense that the population density is constant. Nobody enters the urban area from outside and nobody leaves the area. It is static in the sense that agents cannot choose their location.

**Utility.** At any point in the city area, agents can derive benefit $\bar{v}$ from a trip to the centre.[10] Agents choose between the congested line ($\mathcal{C}$), the nearest uncongested line ($\mathcal{U}$) and "no trip" ($\varnothing$). A decision for "no trip" leads to an overall utility of zero. Otherwise the agent's utility is calculated by $\bar{v}-$costs. The cost function depends on the travel mode the agent chooses and on his location.

**Location.** As described above, the uncongested lines divide the city in $q$ symmetric sectors. Each of these sectors can be partitioned into two symmetric half sectors. Due to their symmetry, we will examine only a representative half sector in the following. When considering the number of traffic mode users, we always refer to the corresponding number in this representative half sector. In order to get the corresponding result for the whole city, one needs to multiply the value by the factor $2q$.

[9] By assuming such a type of line distribution, we can match the fact that the further away the agent is from the centre, the higher is the average distance from a public transport station. It is not necessary to assume that there are no side-connections of the public transport lines. Indeed, if one did so, one would only complicate the calculations without changing the qualitative results.

[10] The constant benefit setting represents a relatively strong assumption. Nonetheless, all qualitative results of the chapter would stay the same if we introduced a distribution over the possible benefits. The main reason for this assumption is simplicity, and to abstract from any other heterogeneity than location.

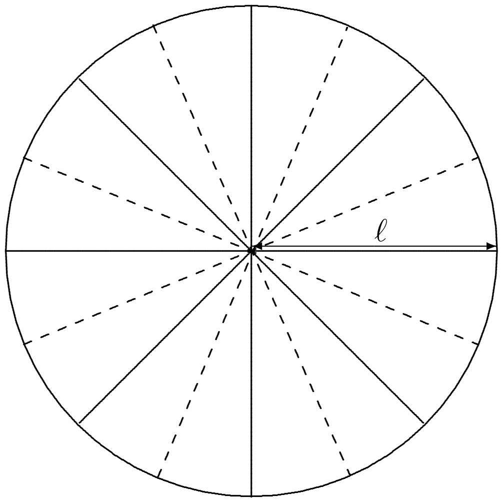


Figure 2.1: Depiction of the Urban Area with q=8 Uncongested Links

The dashed lines divide every sector into two halves. Each half represents a half sector.

The position of an agent is given by $(x_1, x_2)$, where $x_1$ is the distance to the centre point, hereinafter just called "distance"[11]. $x_2$ is the ratio of the length of the shortest way along the appendant circle line to the next uncongested line and $\frac{\pi \cdot x_1}{q}$, where $\frac{\pi \cdot x_1}{q}$ is the length of the circle line of the corresponding half-sector.[12] So we have $0 \leq x_1 \leq \ell$ and $0 \leq x_2 \leq 1$. Therefore, $x_2 = 0$ for an agent who is located exactly on an uncongested line. An agent at distance $x_1$ from the centre who is positioned exactly in the middle of two uncongested lines, is located at $(x_1, 1)$. We will call the agent's home location also his "starting distance". In order to keep the calculations simple, we assume that the agent can move only towards the centre or along a circle line.[13]

[11]When $x_2$ is not relevant, we use the terms "distance" and "position" to indicate the distance to the centre.

[12]The length of the appendant circle line of a full sector is $\frac{2\pi x_1}{q}$.

[13]Many roads in cities of European style are built in circles around the centre. Nevertheless this assumption is not crucial for our results, but it simplifies the calculations considerably.

**Cost Structure.** In the basic model, there are mainly four different types of costs for an agent.

**Fixed costs.** Fixed costs are independent of the agent's location.[14] Denote by $m_c$ the fixed costs for the congested mode, and by $m_{uc}$ the fixed costs for the uncongested mode. Let $m = m_c - m_{uc}$ be the difference in fixed costs. To simplify the calculation henceforth, we will only use $m$ on the side of the costs of the congested line.

**Reaching costs.** Reaching costs arise from the fact that agents have to spend time and money to get to the desired line. Reaching costs are location-dependent since a higher reaching distance to the preferred line leads to higher costs. In order to simplify the calculations, we assume marginal reaching costs of zero for the congested line and marginal reaching costs of one for the uncongested line.[15]

**A Half Sector**

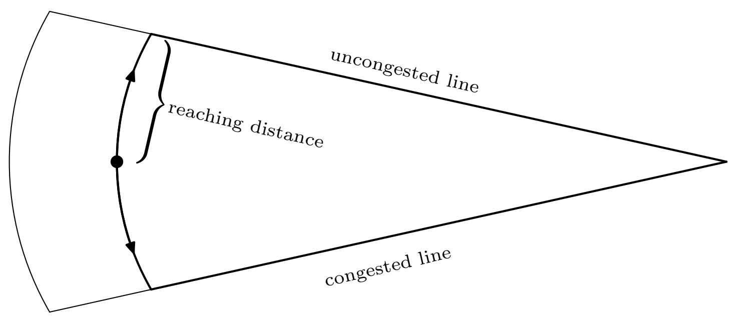


Figure 2.2: The Reaching Distance

[14]Fixed costs for the congested line are, for example, the costs of buying a car. For the uncongested line the fixed costs could be interpreted as the opportunity costs of not owning a car, or the costs for a bicycle to reach the next public transport station. The components included in fixed costs are arbitrarily chosen, and depend only on the point of view. But one should be very careful to add the corresponding costs for both modes of traffic in the right way.

[15]For our results this strong assumption of zero reaching costs for the congested line is not necessary, but it simplifies the calculation considerably. Indeed, in order to get the qualitative results we derive, the marginal reaching costs for the congested line only have to be strictly lower than the marginal reaching costs for the uncongested line.

**Free flow costs.** Free flow costs result from the uncongested and undistorted use of a traffic line. In case of road traffic, free flow costs include the petrol price and time costs. In case of public transport, they include the ticket price plus time costs. Again, for simplicity, we assume marginal free flow costs of one for both traffic modes.

**Congestion costs.** Congestion costs represent the externalities a road user is exerting on other traffic participants. Naturally, they arise only on the congested mode. Since the traffic volume is dependent on the location at the congested line, marginal congestion costs depend on the location as well. Marginal congestion costs are modelled by a function $c(N(r))$, with $N$ being the total number of users of the congested line who start their journey to the centre beyond distance $r$. We assume the marginal congestion cost function $c$ to be consistent with the Vickrey congestion cost structure. For each traffic mode, we normalize the appendant capacity term in the numerator to one.

When choosing the uncongested line, the agent faces reaching costs and free flow costs. An agent at location $(x_1, x_2)$ has free flow costs of $x_1$ and reaching costs of $x_2 \cdot \frac{x_1 \cdot \pi}{q}$. When choosing the congested line, the agent at location $(x_1, x_2)$ faces free flow costs and congestion costs. Furthermore the fixed cost difference $m$ will be added in this case. Let $N(x_1)$ be the total number of road users whose starting distance is $x_1$ or higher. So $N(x_1)$ is the mass of all people living beyond $x_1$ and using the congested line. By construction, $N$ is a (weakly) decreasing function in $x_1$. The marginal congestion costs at distance $x_1$ are given by the value $c(N(x_1))$. Denote by $x_1^0$ the distance from the centre where the road user enters the congested line, then the congestion costs are $\int_0^{x_1^0} c(N(x_1))dx_1$.

**Agent's choice.** Denote by $D(x_1, x_2)$ the choice function for an agent at location $(x_1, x_2)$. Then $D(x_1, x_2)$ is defined by the following optimization problem.

$$\begin{aligned} D(x_1, x_2) = \underset{x \in \{\mathcal{C}, \mathcal{U}, \varnothing\}}{\arg\max} \Big\{ & \Big(\bar{v} - m - \int_0^{x_1} \big(1 + c(N(r))\big)dr\Big) \cdot \delta_{\mathcal{C}}(x) \\ & + \Big(\bar{v} - x_1 - x_2 \cdot \frac{x_1 \cdot \pi}{q}\Big) \cdot \delta_{\mathcal{U}}(x)\Big\}, \end{aligned} \tag{2.1}$$

where $\delta$ is defined as follows:

$$\delta_s(t) = \begin{cases} 1 & \text{if } t = s \\ 0 & \text{if } t \neq s \end{cases}$$

The function $N(r)$ is defined as

$$N(r) = \int_r^l \int_0^{\frac{\pi z}{q}} \mathbb{1}_{\{D(z,\frac{kq}{\pi z})=\mathcal{C}\}} dkdz \tag{2.2}$$

The function $N(r)$ is defined by two nested integrals. The inner integral adds all road users along the appendant circle line within a half sector. The outer integral considers the distance to the centre. At distance $r$, all possible road users behind $r$ have to be considered.

Note that due to the assumption that everybody travels to the centre, wheresoever an agent enters into the congested line, he exerts a negative congestion externality on every other user of this line. In sum, every road user has a negative impact on any other road user, no matter where he is located.[16]

**Assumptions.** The parameter assumptions and the conditions we set for the congestion function are as follows.

1. $c(0) = 0$
2. $c(N)$ is continuous, strictly increasing and concave.
3. $l = \bar{v}$
4. $m > 0$

Assumption 1 is commonly used in the corresponding literature, and ensures that without any further traffic, a user of the potentially congested line has to pay exactly

[16] Since we assume a continuous density function for the population, to be formally correct we should say "a representative mass of agents" instead of "agent". In equilibrium, of course, a single agent does not have any effect on other road users.

the free-flow costs. In assumption 2, we match the fact that an increasing number of road users leads to higher congestion externalities and therefore to more costs. The more agents use the congested line, the higher is the external effect they are exerting on other road users. Note that $c$ is the *marginal* congestion cost function[17], not the global one. The whole congestion costs for a road user who starts at $x_1$ are calculated by $\int_0^{x_1} c(N(r))dr$. Assumption 3 ensures that there are indeed agents in the city who prefer "no trip" and that the confines of the urban area are set by the last position where commuting traffic to the centre occurs. Hereby we define the factual confines of the urban area's commuting zone. Assumption 4 says that the fixed costs associated with road usage are higher than the fixed costs of public transport. Although one has to consider opportunity costs when opting for no car, it is quite intuitive that car use for commuting causes higher fixed costs than using public transport.

## 2.4 Equilibrium

In this chapter we study the equilibrium properties of the model. First, we define our equilibrium concept. Then, the first theorem proves the existence of a unique equilibrium and provides conditions under which both traffic modes are in use. A description of the spatial distribution is presented in Proposition 2.4.2. Proposition 2.4.3 analyses the evolution of the number of congested-line users with respect to their distance. Proposition 2.4.4 shows how the equilibrium changes when the parameters are altered. All proofs are relegated to the appendix.

**Equilibrium Concept.** An equilibrium of the model consists of a pair of functions

$$\begin{aligned} D : [0,\ell] \times [0,1] &\rightarrow \{\mathcal{C}, \mathcal{U}, \varnothing\} \\ (x_1, x_2) &\mapsto D(x_1, x_2) \end{aligned}$$

[17] The function $c$ is the marginal congestion cost function with respect to distance.

and

$$\begin{aligned} N : [0, \ell] &\rightarrow \mathbb{R} \\ r &\mapsto N(r) \end{aligned}$$

such that equations (2.1) and (2.2) are simultaneously fulfilled for all $x_1, r \in [0, \ell]$ and $x_2 \in [0, 1]$.

**Proposition 2.4.1** *There exists a unique equilibrium. In equilibrium both traffic modes are in use if and only if $mq < \pi l$. In this case, the marginal agent on the middle line of a sector who is indifferent between the uncongested and the congested line is positioned at location $(r^\star, 1)$ where $r^\star = \frac{qm}{\pi - qc(N(0))}$.*[18] *In the inner circle with radius $r^\star$, only the uncongested line is used.*

The main statements of Proposition 2.4.1 are to show that the equilibrium is unique and that for any parameter values there will be car commuting traffic if only the city size is large enough. Because of the uniqueness of the equilibrium, any policy implementation will lead to a well defined outcome. Furthermore, this proposition implies an inner-city region in which it is always preferable to use the uncongested line: it is simply a circular disc around the centre with radius $r^\star$. Since we will always consider the interesting case of having both commuting modes in use, we always assume the following condition:

$$mq < \pi l \tag{2.3}$$

Now we are able to describe the equilibrium distribution of traffic in case all three modes are in use:

**Proposition 2.4.2** *Under condition (2.3), in every half-sector three different connected areas emerge. In area I, agents only use the uncongested line. In area II, only the congested line is used. In area III, agents prefer "no trip". In the diagrammatic depiction of Figure 2.3, point A $(r_A, y_A)$ is characterized by*

$$\int_0^{r_A} c(N(r))\, dr = \frac{\pi r_A y_A}{q} - m = l - m - r_A.$$

[18] Note that $N(0)$ is an endogenous value. It is not possible to represent $r^\star$ by a closed explicit formula of exogenous parameters.

*Point B* $(r_B, y_B)$ *is characterized by*

$$r_B = r_A \text{ and } y_B = 1.$$

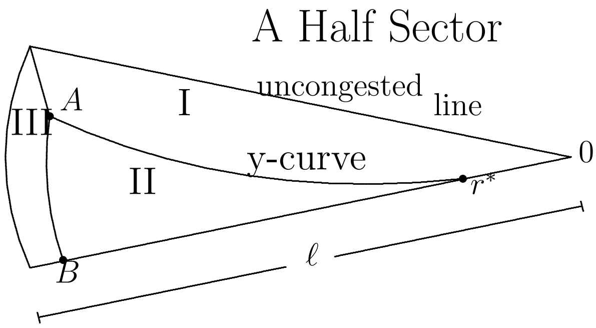


Figure 2.3: Depiction of the Equilibrium Distribution of Traffic

Note that this proposition also describes the city commuter belt as a star shaped area. The edges of the star are represented by the ends of the uncongested lines.

In the following two propositions we describe the functions $N(r)$ and $y(r)$ in equilibrium, and we state that they are solutions of integral equations. This will help us solve the model numerically.

**Proposition 2.4.3** *Under condition (2.3), the function* $N(r)$ *as defined in (2.2) is positive and constant in the interval* $[0, r^\star]$*, and equals zero in the interval* $[r_A, l]$*. On the interval* $[r^\star, r_A]$*,* $N(r)$ *is strictly decreasing.* $N(r)$ *is characterized as a solution of the following integral equation:*

$$N(r) = \int_r^l \int_0^{\frac{\pi z}{q}} \mathbb{1}_{\{\int_0^z c(N(x))\,dx \le \min(k-m, l-m)\}}\, dk\, dz, \quad r \in [0, \ell] \tag{2.4}$$

See Figure 2.4 for a depiction of the qualitative features of $N(r)$.

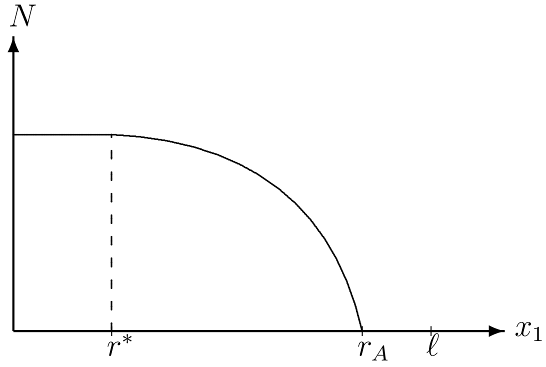


Figure 2.4: A Qualitative Depiction of the Function $N(r)$

**Proposition 2.4.4** *The curve $(r, y(r))$, separating area I from area II, is characterized as a solution of the following set of equations:*

$$\int_0^r c\Big(\int_z^{r_A} \frac{\pi k}{q}\big(1 - y(k)\big)\,dk\Big)\,dz = \frac{\pi r y(r)}{q} - m \quad \textit{for all } r \in ]r^\star, r_A[$$

*together with*

$$y(r) = 1 \quad \textit{for all } r \in ]0, r^\star[ \,\cup\, ]r_A, \ell[.$$

Note that in Proposition 2.4.4, $r^\star$ and $r_A$ are endogenous variables, determined by the function $N(r)$ which can be defined trivially through $y(r)$:

$$N(z) = \int_z^l \frac{\pi k}{q}(1 - y(k))\,dk.$$

Marginal road use is given by $1 - y$ and equals the ratio of the number of road users and public transport users at a certain distance to the centre. Let us now turn to the results of comparative statics concerning marginal road use.

**Proposition 2.4.5** *At any $r \in ]r^\star, r_A[$, marginal road use will decrease if the fixed cost difference $m$ or the city size*[19] *increases. If $c'$ is sufficiently small*[20]*, marginal road use will decrease at any $r \in ]r^\star, r_A[$ with an increasing number of uncongested lines.*

[19] Note that due to assumption (2), city size $\ell$ equals the benefit of a trip $\bar{v}$. Therefore an increase $\ell$ automatically causes a proportional increase in the benefit of a trip.

[20] This assumption is only of technical nature. Numerical simulations show that this assumption is not crucial for the result.

The results of Proposition 2.4.5 are quite intuitive, except for the last one. A higher fixed cost difference means higher costs of road use in comparison to public transport. Therefore, fewer people tend to use the car. If the number of uncongested lines increases, the average distance to the next public transport station will fall, leading to lower reaching costs. Therefore, the public transport line is used by more agents. The intuition for the last effect is the following: Take, for example, the starting distance $r_A$ as given. If the city size increases, more agents are already on the congested line (compared to none in the situation before). Therefore, the cohort at distance $r_A$ has less incentive to use the congested line. A similar reasoning applies for any point in $[0, r_A]$. Nonetheless, a second effect comes into play. Since from 0 to $r_A$ the rate of agents who use the road is lower than in the situation before, there also exists an opposite force on $y(r_A)$. This opposite effect also influences the agents beyond $r_A$, reinforcing the first effect and so on. The result shows us that in sum the first effect exceeds the second one.

### 2.4.1 Numerical Simulation of the Equilibrium

We simulated the equilibrium distribution numerically for the parameter values $\ell = 100$, $q = 3$ and $m = 5$, using a linear marginal congestion cost function $c(x) = 0.001 \cdot x$. If not indicated differently, these parameter values are used in all subsequent simulations.

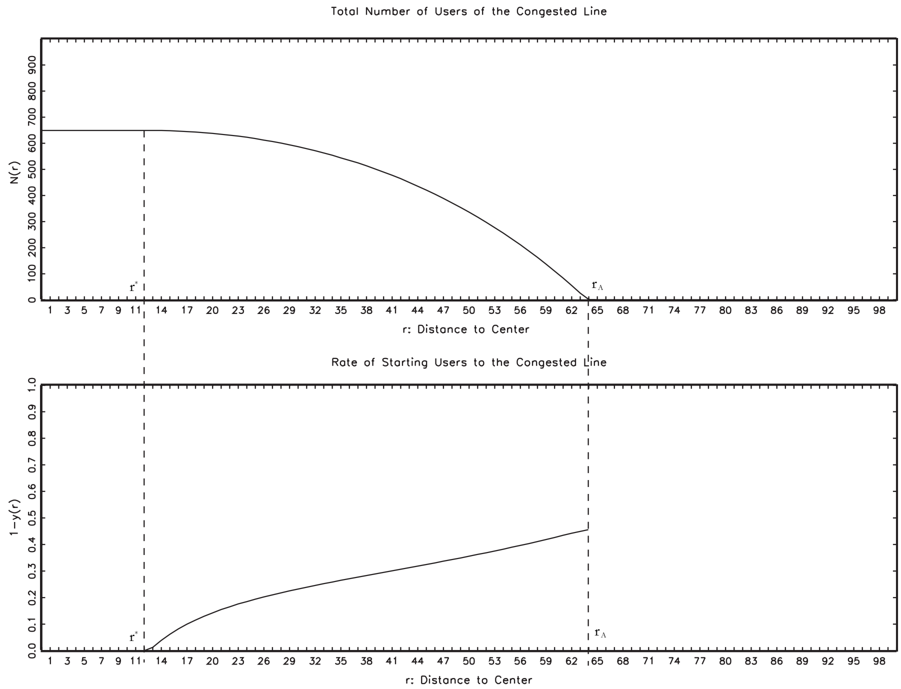


Figure 2.5: N(r) and y(r)

In this simulation, we get the following values: $r^\star \approx 12.5, r_A \approx 63.5$ and $N(0) \approx 645$. Note that beginning from $r^\star$, $1 - y(r)$ is concave and becomes close to a linear function when approaching $r_A$, indicating that the tendency to use the congested line increases faster for low-distance car users than for suburban car users.

## 2.5 Bottlenecks and Tolls

In this section, we consider the welfare of the whole traffic system. After defining a canonical welfare function, we show that the resulting equilibrium is inefficient. Then, we prove that it is indeed possible to increase total welfare by reducing road capacity. We model this capacity reduction by an increase of the costs of the congested link when passing a certain distance. In a subsequent numerical exercise, we show how and where these capacity reductions should be placed optimally. In the last part we do the same for potential tolls. Again, all proofs are relegated to the appendix.

### 2.5.1 Welfare of the System

We choose a very simple welfare function. It is just the integral over all agents' utilities plus potential toll income:

$$W = \int_0^{\ell} \int_0^{\frac{\pi x_1}{q}} \left[\bar{v} - \text{costs}(D(x_1, \frac{x_2 q}{\pi x_1}))\right] \cdot \Big(1 - \delta_{\varnothing}\big(D(x_1, \frac{x_2 q}{\pi x_1})\big)\Big) dx_2\, dx_1 + T(V, d)$$

with $D$ being the decision function, defined according to the appropriate situation. Potential costs of passing a bottleneck or a tollgate are included in the term for "costs". $T(V, d)$ represents the toll income when the toll $V$ is charged at position $d$. If we are not in the case of tolls, $T(V, d)$ equals zero.

In the following depiction, we see how welfare changes when the fixed costs difference $m$ is altered, holding $q$ constant ($q = 3$), and, when the number of uncongested lines $q$ is altered, holding $m$ constant ($m = 5$).

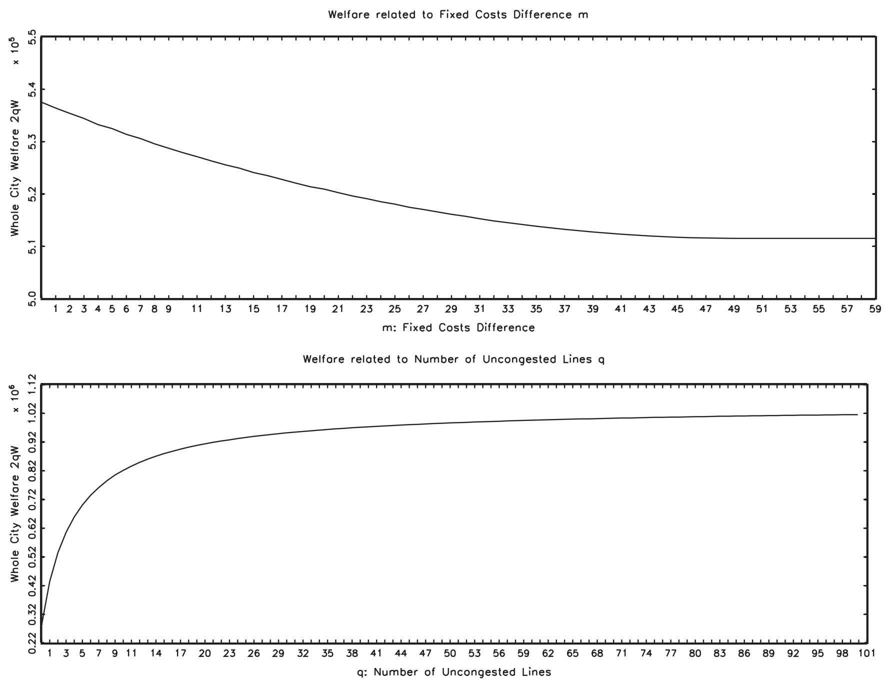


Figure 2.6: Welfare of Whole Urban Area Dependent on m and q

Increasing $q$ means reducing the average distance to the next uncongested link which leads to lower reaching costs, and therefore to higher total welfare. Increasing $m$ increases the costs of using roads, and therefore reduces welfare. Note that the welfare value in this figure is multiplied by $2q$ in order to capture the value of the whole city, and not just one half-sector. In all other figures, all values refer to a half-sector only.

Using the definition above we are able to state a crucial result:

**Proposition 2.5.1** *The resulting equilibrium is inefficient. Too many agents use the road.*

The intuition for this result is similar to the intuition of a standard congestion problem. In their personal decisions, road users do not account for the negative congestion externalities they cause on other traffic participants.

In the next part, we show the interesting result that it is possible to increase welfare by well-chosen capacity reduction.

### 2.5.2 Bottlenecks

Because of political, social or legal reasons it is often not possible for a local government to install toll systems. In the following, we show that there exists an incentive for governments to tackle the inefficient traffic allocation described above by adding some distortive elements to the traffic system, so called bottlenecks. By an (artificial) bottleneck, we mean a capacity reduction of the congested line at a certain distance $d$. Let us suppose that the installation of the bottleneck is possible at no costs. We model the bottleneck by adding a discrete cost component (which we will refer to as severeness) for all users who pass the bottleneck multiplied with the total number of users $N(d)$ at that distance.[21] Note that no matter how severe the bottleneck might be, there will always be a positive mass of agents beyond $d$ who use the road because the costs of passing the bottleneck are directly proportional to $N(d)$.

[21] This structure of additional bottleneck cost is chosen for simplicity. Note that the additional bottleneck cost is independent of the marginal congestion cost function $c$.

In the general case, we consider a number of bottlenecks. Let $V_i$ be the parameter indicating the severeness, and $d_i$ denote the distance to the centre of bottleneck $i$. Denote by $D(x_1, x_2)$ the choice function for an agent at location $(x_1, x_2)$. In the case with bottlenecks, $D(x_1, x_2)$ is defined by the following optimization problem:

$$
\begin{aligned}
&D(x_1, x_2) = \\
&\quad \arg\max_{x \in \{\mathcal{C}, \mathcal{U}, \varnothing\}} \Big\{ \Big( \bar{v} - m - \int_0^{x_1} \Big(1 + c(N(y))\Big) dy - \sum_i V_i \cdot N(d_i) \cdot \mathbb{1}_{\{x_1 > d_i\}} \Big) \cdot \delta_{\mathcal{C}}(x) \\
&\qquad\qquad\qquad\qquad\qquad\qquad + \Big( \bar{v} - x_1 - x_2 \cdot \frac{x_1 \cdot \pi}{q} \Big) \cdot \delta_{\mathcal{U}}(x) \Big\},
\end{aligned}
$$

This equation, together with (2.2), defines the solution of the equilibrium with bottlenecks. Note that the costs of undergoing bottleneck $i$ at distance $d_i$ are dependent on the endogenous equilibrium value of $N(d_i)$.

In the remaining part of this section, denote by $r_A$ and $r^\star$ the corresponding values of the undistorted equilibrium. Now we are able to prove one of the main results of the chapter:

**Proposition 2.5.2** *Let $V > 0$. Then a distance $d \in [r^\star, r_A]$ exists with the following property: If one installs an artificial bottleneck of severeness $V$ at distance $d$, overall welfare will be higher than in the situation without distortion.*

This proposition is a generalization of Proposition 2.5.1. It shows us that by a well-chosen reduction of road capacity it is possible to increase total welfare. Note that this result holds for any positive bottleneck severeness $V$.

By the installation of a bottleneck, we can discriminate users by their location. The capacity reduction at distance $d$ leads to increased costs for users of the congested line who start their journey beyond distance $d$. There are two reasons why we want to reduce the number of users with distance higher than $d$ from the congested line. First, they exert the highest externalities on the the traffic system, since they are travelling through most parts of the congested line. And secondly, all road users behind $d$ contribute only a small part to overall welfare, since they are relatively close to indifference to "no trip".[22] If one forces road users to undergo the artificial bottleneck, most of them either switch to the uncongested line or do not take the trip at all. In both cases, they do no longer contribute to congestion any more. The proposition shows us that if the bottleneck is installed at the right place, the total (positive) effect of reduced negative externalities on other road users will be stronger than the (negative) distortional effect on long-distance commuters.

We simulate the distorted equilibrium for $V = 0.06$, $d = 50$, all other numerical parameter values chosen as in Section 2.4.1.

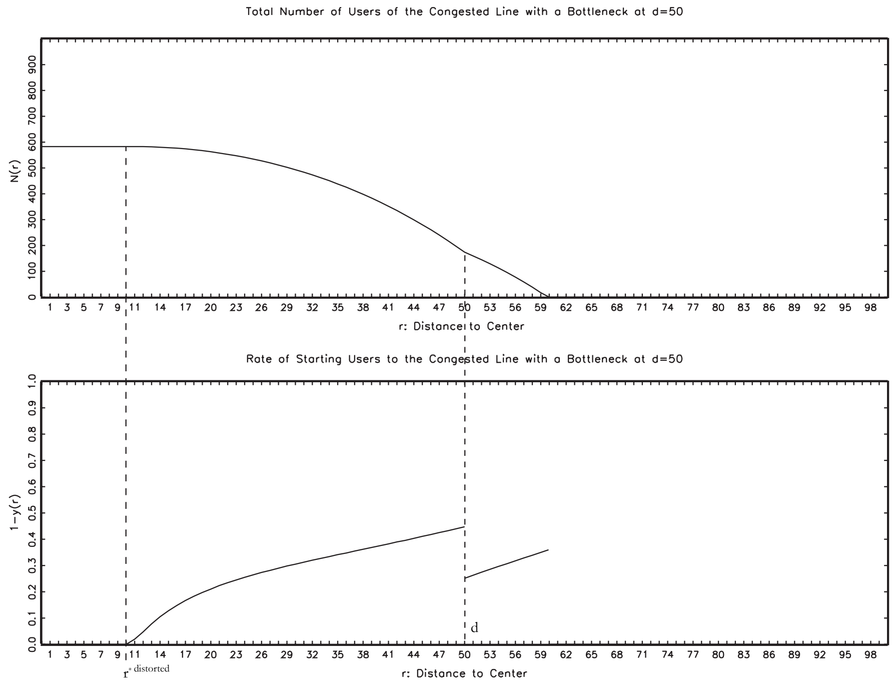


Figure 2.7: Equilibrium with a Bottleneck at d = 50

Of course, in the distorted equilibrium[23], we have a lower number of road users in total: $N^{\text{distorted}}(0) \approx 580 < N(0) \approx 645$, inducing a smaller inner-city region, estab-

[22] An effect in opposite direction is that a relatively high number of road users is distorted. The proposition shows us that if $d$ is optimally chosen, this opposite effect is lower than the described positive effects.

[23] The welfare in this equilibrium amounts to 88871 which is slightly lower than in the undistorted case (88903). However, by increasing the bottleneck severeness tremendously at this position, we can even create welfare gains.

lished by $r^{\star\text{distorted}} \approx 10.5 < r^{\star} \approx 12.5$.

In the following, we examine how the welfare of the system changes when the severeness and the position of the bottleneck are altered. Let us consider two benchmark cases: The case without any bottlenecks and the case where the severeness of the bottleneck tends to infinity. In the diagram below the welfare function is printed against the position of a bottleneck with prohibitively high[24] severeness.

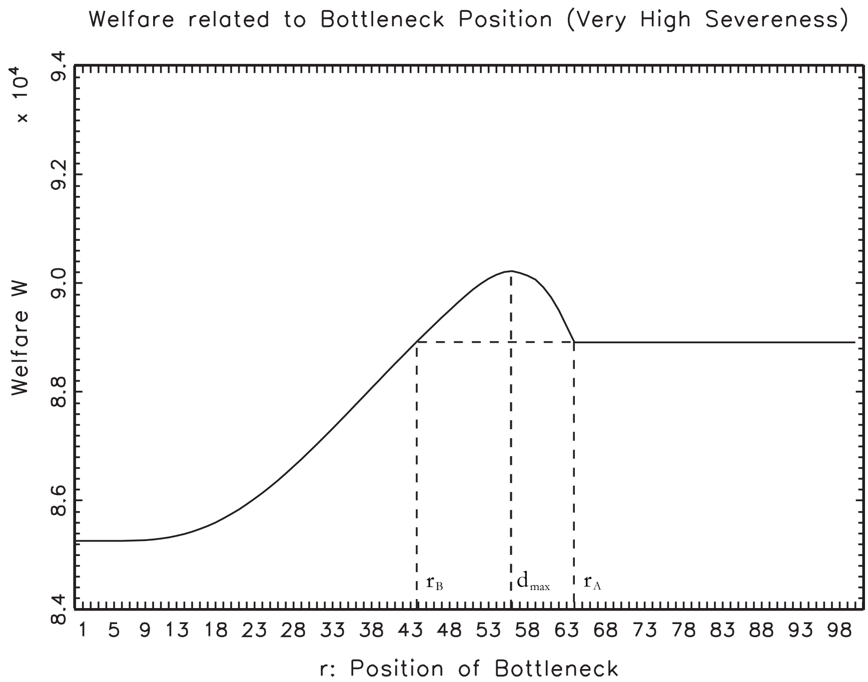


Figure 2.8: Welfare with a Very Severe Bottleneck

Let us define $d_{\max}$ as the position which maximizes the welfare function and $r_b$ as the lowest distance at which the welfare without distortion equals the welfare with a very severe bottleneck. In the simulation above we get $r_b \approx 43$ and $d_{max} \approx 56$.

The welfare function, evaluated at position of 0, tells us the welfare without any road use at all. This value stays constant on an interval of positive mass, since close to the centre, agents will never use the car because of the fixed cost difference $m > 0$.

[24]In the numerical simulation, we set the severeness to $10^7$ such that the number of road users passing the bottleneck is very close to zero.

Then welfare increases with $r$ due to the fact that more and more potential road users are no longer distorted in their decision.

The numerical simulation is also confirming the result of Proposition 2.5.2 by showing that welfare is higher than in the undistorted equilibrium value for bottleneck positions $r_A > d > r_b$. In this region, the reasoning of Proposition 2.5.2 comes into effect. Welfare is maximized at a bottleneck position of $d_{max} < r_A$. At $d_{\max}$ the marginal cost of distorting the group of agents living in $[d_{\max}, r_A]$ equals the marginal benefit of reduced congestion costs for all remaining road users. For bottleneck positions $d > d_{\max}$ too few road users are distorted compared to the optimum, welfare decreases in $r$. Since for bottleneck distances higher than $r_A \approx 63.5$, even in the undistorted equilibrium there is no road use, the welfare stays constant and equals the value of the undistorted equilibrium.

Now we turn to the relation between $V$ and the position of the bottleneck at the optimum. First, let us fix the position $r$ of the bottleneck, and determine the optimal value of severeness $V_{opt}(r)$. The following figure shows the result of a numerical exercise.

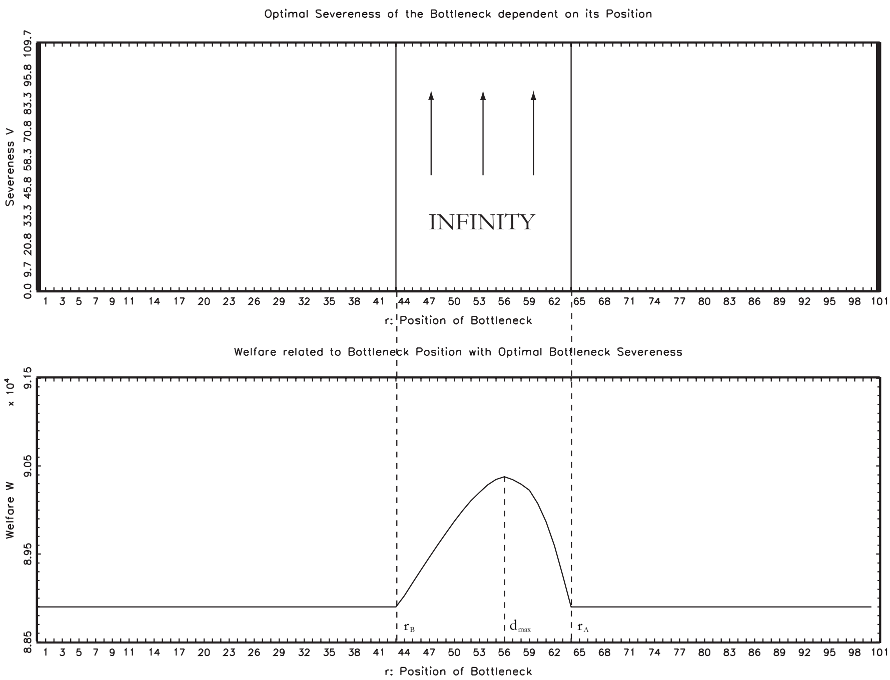


Figure 2.9: Optimal Bottleneck Severeness

In the upper part of figure 2.9, we see that at positions smaller than $r_b$, the optimal severeness is zero, showing that there it does not pay off to have any bottleneck at all because the distortional effects are too high. The two orthogonal lines at $r_b$ and $r_A$ indicate that in the interval $]r_b, r_A[$ the optimal severeness of the bottleneck tends to infinity.[25] Bottlenecks at positions higher than $r_A$ have no effect anyway.

The lower panel of the diagram shows us welfare as a function of the bottleneck position, given that the bottleneck severeness is chosen optimally. Since $V_{opt} = 0$ for all bottleneck positions smaller than $r_b$, total welfare equals the welfare of the undistorted equilibrium. For a bottleneck position in $[r_b, r_A]$, the optimal severeness tends to infinity. Therefore, in this region, welfare is equal to the benchmark case of prohibitively severe bottlenecks.

The following result is robust in all parameter dimensions and summarizes the findings above:

**Numerical Result 2.5.3** *Suppose it is possible to place a bottleneck of arbitrarily high severeness at any distance $r \in [0, \ell]$. Then, there exists $r_b$ in $]r^\star, r_A[$ such that for all $r \in [0, r_b]$, the optimal severeness $V_{opt}(r)$ equals zero, and for all $r \in ]r_b, r_A[$, we have $V_{opt}(r) = +\infty$. There is a single maximum of the welfare function at position $d_{\max} \in ]r_b, r_A[$.*

Furthermore, we see that under any parameter combination, "high quality" traffic participants located next to $r^\star$ should never be distorted in their decision by road

[25] A surprising result is that even for positions in $]r_b, d_{\max}[$, the optimal severeness tends to infinity, although, at least for a part of the distorted road users, the welfare contribution is higher than the amount of negative externalities they exert on others. The reason is the following: In any case, the optimal severeness of a bottleneck should be high enough to prevent all agents behind $d_{\max}$ from using the road because their negative congestion externality is higher than their positive contribution to welfare. But then, the contribution to welfare of the last cohort of road users living behind the bottleneck (even at positions in $]r_b, d_{\max}[$) is very close to zero. However, their marginal congestion externality is strictly positive. At this point, the same reasoning like in Proposition 2.5.2 comes into play to show that they should be deterred from road use as well. Then, the reasoning is iterated up to bottleneck position $d$.

capacity reduction. Obviously, these agents contribute most to total welfare and cause the fewest externalities on others.

The following diagram shows a smoothed graph of the bottleneck's optimal distance from the centre dependent on its severeness $V$ for parameter values as in Section 2.4.1.

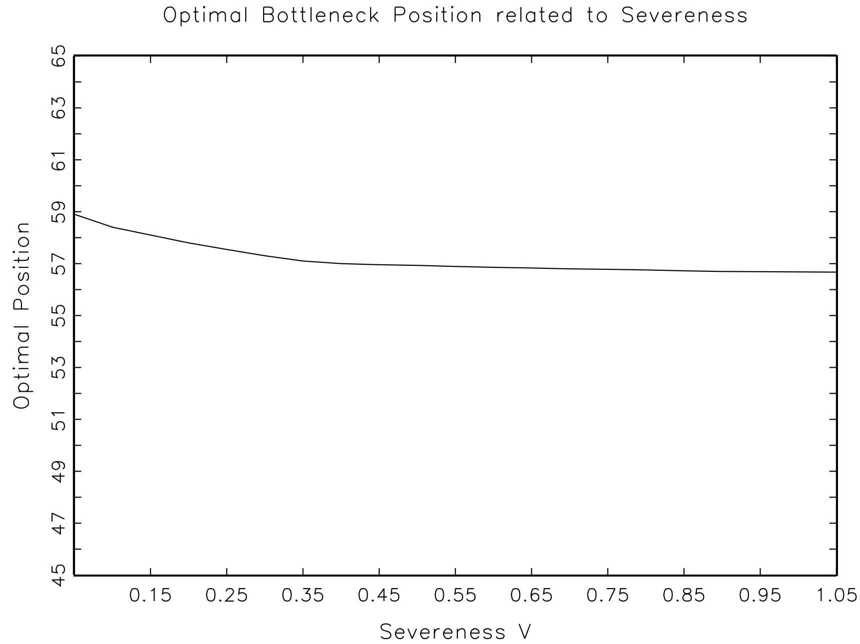


Figure 2.10: Optimal Distance of the Bottleneck Dependent on Bottleneck Severeness

If $V$ tends to infinity, the optimal bottleneck position obviously tends to $d_{\max}$. If the bottleneck severeness is too low, only few road users behind the bottleneck will be induced to leave the road, but all of them will be distorted. Hence, the trade-off is to induce many road users who exert high negative externalities to leave the congested line, without distorting too many road users who stay there anyway. In this situation the incentive to leave the road is too small for many road users who exert high negative externalities. If this is the case, it is optimal to deter only those road users who will actually refrain from using the road.

Now we turn to the situation with $n$ bottlenecks being possible and ask the question where they should be placed to maximize welfare. Obviously, maximized welfare (weakly) increases in $n$. The following numerical result again is robust under all tested parameter combinations.

**Numerical Result 2.5.4** *Using the parameter values of Section 2.4.1, in the situation with $n$ possible bottlenecks of severeness $V_i > 0, \quad i = 1 \ldots n$, all bottlenecks should be placed at the same position. The optimal position is in $[d_{\max}, r_A[$ and fully determined by $\sum_{i=1}^{n} V_i$.*

To explain this result intuitively, consider the situation when one bottleneck is already fixed. The question now is: Could it pay off to position a second bottleneck at a different place?

This can only lead to an optimal allocation if road users between these two positions should undergo only one bottleneck in optimum. Since all road users at a certain distance $r$ contribute the same value to welfare and cause the same negative externalities, we consider a representative agent at this distance. Consider a representative road user right behind the bottleneck closest to the centre. Suppose despite the bottleneck, his contribution to welfare is still higher than the externalities he causes. Then, in optimum, he should not face any bottleneck at all. If we are in the opposite situation, the second bottleneck should be placed in such a way that he is induced to leave the road, therefore right in front of him or even closer to the centre. Altogether, we see that different bottleneck locations never lead to optimal allocations.

### 2.5.3 Tolls

Let us now consider a situation where the government is able to charge all road users who are passing a tollgate. We assume that tolling is possible at no costs. There are two main differences between tolls and bottlenecks. First, toll income contributes to the welfare function, while there is no such extra income in the case of bottlenecks. Secondly, when passing a tollgate at distance $R$, the toll fee is independent of $N(R)$, while the additional costs caused by a bottleneck at distance $R$ are positively dependent on $N(R)$.

For toll fees $V_i$, charged at positions $R_i$, toll income $T$ equals

$$\sum_i V_i \cdot N(R_i),$$

and is added to the welfare function. If the toll income is redistributed equally to all residents of the urban area, the main winners of the installation of a toll will be road users living closer to the centre than the tollgate's position. They profit by the redistributed toll income and by lower congestion costs as well. Users of the uncongested line and residents who choose no trip profit only by their part of toll income, while road users behind the tollgate lose in total.

Denote by $D(x_1, x_2)$ the choice function for an agent at location $(x_1, x_2)$. In the case of tolls $D(x_1, x_2)$ is defined by the following optimization problem:

$$D(x_1, x_2) = \underset{x \in \{\mathcal{C}, \mathcal{U}, \varnothing\}}{\arg\max} \Big\{ \Big( \bar{v} - m - \int_0^{x_1} 1 + c(N(y))dy - \sum_i V_i \cdot \mathbb{1}_{\{x_1 > R_i\}} \Big) \cdot \delta_{\mathcal{C}}(x) \\ + \Big( \bar{v} - x_1 - x_2 \cdot \frac{x_1 \cdot \pi}{q} \Big) \cdot \delta_{\mathcal{U}}(x) \Big\},$$

where $V_i$ denotes the toll charged at tollgate $i$ at position $R_i$. This equation, together with (2.2), defines the solution of the equilibrium with tolls.

Some results of the case with bottlenecks carry over to the case when tolls are applicable. The analogon of the main Proposition 2.5.2 can be formulated in a slightly stronger sense:

**Numerical Result 2.5.5** *If $V > 0$ is sufficiently small and $d \in [0, r_A[$, overall welfare will be higher than in the undistorted equilibrium (if every road user who passes distance $d$ has to pay a toll of $V$).*

In contrast to the case of bottlenecks, the installation of appropriately chosen tolls always increases welfare. Note that we have to guarantee that a sufficiently high number of agents use the road and generate toll income. Hence, the toll $V$ cannot be arbitrarily high.

The reasoning exactly analogous to the proof of Proposition 2.5.2 works for the following assertion:

**Proposition 2.5.6** *Let $V > 0$. There exists a distance $d \in [r^\star, r_A[$ such that if every* *road user who passes distance d has to pay a toll of $V$, overall welfare is higher than in the undistorted equilibrium.*

The benchmark case of very high tolls (toll fee tends to infinity) leads to exactly the same result as in the case of very severe bottlenecks, as shown in Figure 2.8. If tolls are very high, the number of road users behind the tollgate will be equal to zero, leading to zero toll income.

As in the case with bottlenecks, we turn to the relation between $V$ and the position of the tollgate at the optimum. First, let us fix the position $r$ of the tollgate and determine the optimal toll fee $V_{opt}(r)$.

**Numerical Result 2.5.7** *In $[0, r_A]$, the optimal toll fee is a weakly decreasing function of the tollgate's position.*

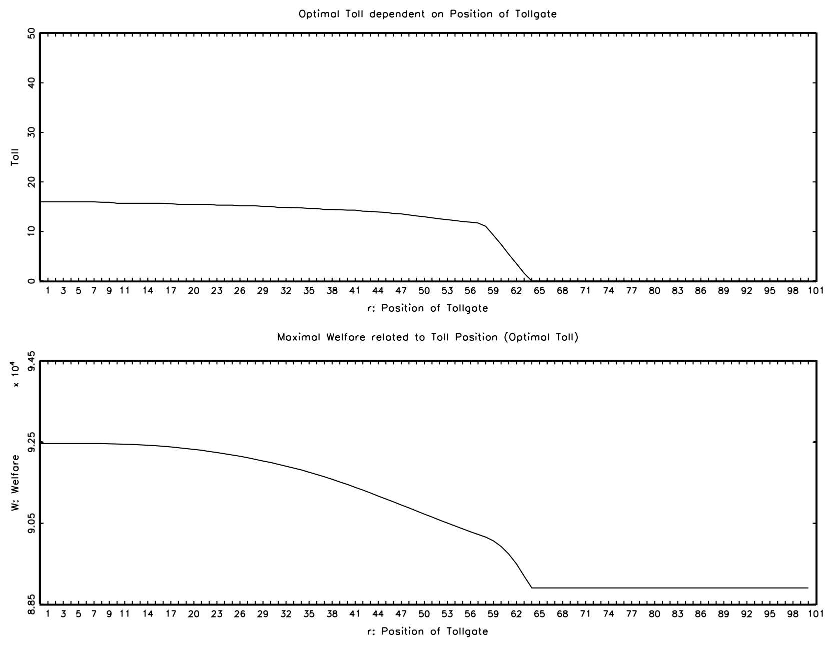


Figure 2.11: Optimal Toll Fee Dependent on the Position of the Tollgate

The upper panel of Figure 2.11 shows the optimal toll as a function of the tollgate's distance to the centre. The lower panel shows the welfare dependent on the tollgate position with optimally chosen toll level. It is intuitive that the equilibrium with optimal tolls leads to better welfare values than in case of optimally chosen bottlenecks because one does not only profit from the reduced congestion externality, but also from the (redistributed) toll income. In all parameter combinations we tested, the marginal value of toll income, $N(d)$, was always superior to the marginal congestion externality effect[26] on other road users.

Interestingly, the optimal toll is falling in the tollgate's distance from the centre. The driving force behind this result is the maximization of the toll income part of welfare. By having a relatively high toll right next to the centre, the number of road users is reduced severely. A tollgate close to the centre offers the advantage that everyone must pass the tollgate, and therefore toll income is maximal. However, the disadvantage is that it is no longer possible to discriminate potential road users by their position. With rising $r$, the influence on the road users decreases. Agents living closer to the centre than the tollgate do no longer contribute to toll income, but profit heavily from the reduced number of road users. An increased number of agents at these positions (compared to the undistorted equilibrium) will use the road which in turn weakens the positive effect of reduced congestion for road users behind the toll gate. In order not to lose too many road users behind the tollgate (who contribute to toll income), it pays off to slightly reduce the toll.

Welfare is maximal at toll positions close to the centre, where one can distort and charge all road users. It decreases with the number of road users who cannot be tolled, but always stays above the welfare function of the benchmark case of prohibitively high tolls.

In the situation when $n$ tollgates, each of toll fee $V > 0, \quad i = 1 \dots n$, are possible, and the result notably differs from the situation in the bottleneck case:

**Numerical Result 2.5.8** *Using the parameter values of Section 2.4.1, we see that in the situation with two possible tollgates, each of the same fee, the first toll should be placed in the centre, and the second at a position between $r_A$ and $r^\star$. For $n > 2$, the result does not change in the sense that further tollgates should be positioned at a place*

[26] In the proof of Proposition 2.5.2, $\Delta E$ is a lower bound and a good approximation for the *absolute* congestion externality effect.

*where they have no influence on the equilibrium (for example $d = \ell$).*

The equilibrium distributions when $n$ tollgates of the same toll fee (equal to 10) can be placed arbitrarily, are shown in Figure 2.12.

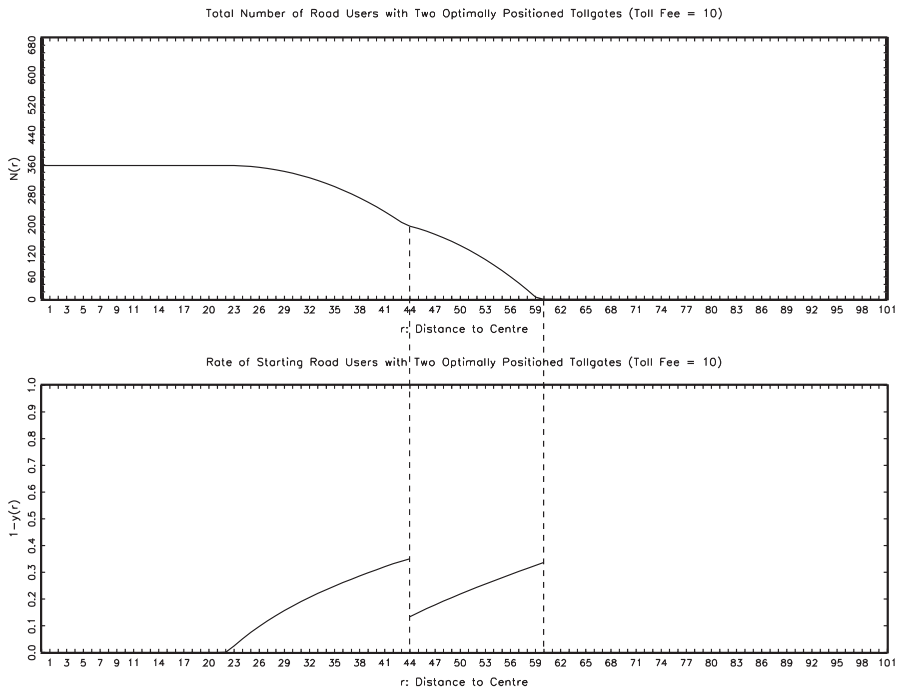


Figure 2.12: Equilibrium Distribution for Two Possible Tollgates

In this equilibrium, long-distance commuters (living behind $r = 43$) pay two times the toll fee, and therefore contribute the most to toll income.

Obviously, optimized welfare (weakly) increases in the number of possible tollgates. Note that due to the structure of the maximization problem, a road user's payments are (weakly) increasing in his distance to the centre because he has to pay successively at different tollgates. From an applied point of view, this shall not be a big constraint, since it is natural that the total toll positively depends on trip length. Nevertheless, if we relax this assumption, it is not clear whether a toll function of different type could be superior. In this situation, moreover, by continuous tolling[27], it should be possible

[27]In our setting, this would require to have different toll fees $V_i$ at different tollgates $(1, \ldots, n)$,

to further increase welfare, albeit difficult to implement in a real world traffic system.

## 2.6 Conclusion and Outlook

We presented a stylized model which enables us to study the two-mode problem in a spatial context. By introducing location-dependent marginal congestion costs, we analyse the interaction of traffic participants on uncongested lines (public transport) and congestible lines (roads) in a monocentric city. We proved the existence and uniqueness of the resulting equilibrium, and showed that the resulting traffic distribution can replicate several empirical facts. First, the use of the congested line pays off only beyond a certain distance to the centre. Secondly, agents who live next to the centre or next to the public transport line are not likely to use the congestible road. Thirdly, inner city traffic is caused by non-inner-city residents, traffic volume is maximal in the city centre, and decreases in the distance to the centre. Furthermore, we showed that an increase in the size of the urban area makes it more likely that an agent at a given distance to the centre uses public transport, although the same applies to any resident in the old confines. This means that at a given distance to the centre the fraction of households using the car is negatively dependent on city size.

We showed the inefficiency of the resulting equilibrium, and that by an appropriately chosen capacity reduction welfare improvements can be achieved. Two concepts of user distortion have been studied. First, road capacity is reduced at certain points through the installation of artificial bottlenecks, thereby causing additional travel costs for road users, and secondly, the distortion is managed by tolls, the income of which is redistributed.

It turns out that in the first case, especially long-distance road commuters should be distorted in order to maximize welfare. This is achieved by the installation of bottlenecks which are placed in such a way that exactly the agents with highest travel distance face them. Nevertheless, in order to maximize welfare, an appropriate design of spacial tolling is still superior in comparison to the use of bottlenecks. In the last

$n \to \infty$ and $V_i \to 0$.

part of the chapter, the optimal number and place of distortions are studied.

One of the most important findings of the model can be summarized as follows: By reducing road capacity at certain distances to the centre in order to protect the road users in the inner city from extensive congestion, total welfare increases. At least from a theoretical point of view, it is also possible to discriminate traffic participants not only by their distance to the centre but by their distance to the next public transport station as well. Very similar effects should be possible if the marginal road user who is indifferent between the uncongested line and the congested line is negatively distorted in such a way that he will choose the uncongested line. Another way to induce him to take the socially optimal decision may be achieved by the provision of positive incentives, for example by giving price reductions for public transport tickets. Nevertheless, because of political and administrative reasons, this seems to be less practicable than reducing road capacity to deter long-distance road users.

An interesting question for further research could be the introduction of heterogeneity in the agent's benefit structure. This would lead to equilibria where at any position within the urban area, both traffic modes are in use, also directly around the centre which seems to fit reality. Although almost all results presented here would only change in magnitude, the richer structure would make it possible to gain new insights in the distribution of road users within a city.

Further fruitful results could be gained if one relaxed the assumption that an agent can use only one line. For example, introducing park and ride facilities at certain locations could help us understand the optimal design of congestion charging zones, as successfully applied in London city centre recently. Another idea pointing in this direction is to relax the assumption that everybody wants to travel to the city centre. Introducing a probability distribution over the possible destinations, though, would complicate the model considerably.

## 2.7 Appendix

**Proof of 2.4.1.**

"$\Rightarrow$" Suppose both traffic modes are in use. Since in any case, the uncongested line is used around the centre, we must also have a mass of agents using the congested line. Note that the relative incentives to use the congested line increase with reaching costs of the uncongested line. So, if an agent at position $(x_1, x_2)$ uses the road, for all $1 - x_2 > \varepsilon > 0$, an agent at position $(x_1, x_2 + \varepsilon)$ will do the same. Then, with $t \in [0, \ell]$, there exists a point $(\ell - t, 1)$ on the middle line $\{(r, 1), r \in [0, \ell]\}$ of a sector and an $\varepsilon$-surrounding of $(\ell - t, 1)$ in which every agent uses the congested line. Every point in the intersection of the corresponding half-sector and this surrounding can be represented as $(\ell - t \pm \varepsilon_1, 1 - \varepsilon_2)$ ($\varepsilon_i \leq \varepsilon$ appropriately chosen, $i = 1, 2$). Then, we have for all $\varepsilon_i$:

$$\begin{aligned} m + (\ell - t \pm \varepsilon_1) + \text{congestion costs} &< \frac{\pi(\ell - t)(1 - \varepsilon_2)}{q} + (\ell - t \pm \varepsilon_1) \\ \Leftrightarrow qm + q \cdot \text{congestion costs} &< \pi(\ell - t)(1 - \varepsilon_2) \end{aligned}$$

If we let $\varepsilon_2 \to 0$, we get $qm + q \cdot \text{congestion costs} + \pi t \leq \pi l$. Now, because of the strictly positive mass of road traffic participants and assumptions (1) and (2), we have strictly positive congestion costs. Therefore, we get

$$qm < \pi\ell.$$

"$\Leftarrow$" This direction of the proof is shown by contradiction. Suppose in the whole sector, every agent opts for the uncongested line. Consider the mass of agents in an $\varepsilon$-surrounding of $(\ell, 1)$. Their travel costs for the uncongested line are $\frac{\pi(\ell-\varepsilon_1)(1-\varepsilon_2)}{q} + \ell - \varepsilon_1$, their travel costs for the congested line are $m + \ell - \varepsilon_1$ ($\varepsilon_i \leq \varepsilon$ appropriately chosen, $i = 1, 2$). Then, we have for all $\varepsilon_i$:

$$\begin{aligned} \frac{\pi(\ell - \varepsilon_1)(1 - \varepsilon_2)}{q} + \ell - \varepsilon_1 &< m + \ell - \varepsilon_1 \\ \Leftrightarrow \pi(\ell - \varepsilon_1)(1 - \varepsilon_2) &< mq. \end{aligned}$$

If we let $\varepsilon_i \to 0$, we get

$$\Rightarrow \pi l \leq mq,$$

which is a contradiction to our assumption.

Consider now the marginal agent on the middle line of a sector who is indifferent between the congested line and the uncongested line:

$$\begin{aligned} \text{costs}(\mathcal{C}) &= \text{costs}(\mathcal{U}) \\ \Leftrightarrow m + \int_0^{r^\star} \big(1 + c(N(r))\big)\, dr &= r^\star + r^\star \cdot \frac{\pi}{q} \end{aligned}$$

Since by definition nobody enters the congested line who is located in distance less to $r^\star$, $N(r)$ is constant on the interval $[0, r^\star]$. We use the notion $N(0)$ for this constant value. We get

$$m + r^\star \cdot c(N(0)) = r^\star \frac{\pi}{q}.$$

Rearranging terms leads to the desired result:

$$r^\star = \frac{qm}{\pi - qc(N(0))} \tag{2.5}$$

Note that because of our assumption $mq < \pi l$, we already know that $\ell > r^\star > 0$, therefore, the denominator of the term above never becomes zero.

Now we study existence and uniqueness of the equilibrium.

<u>1. Case: The congested line is not used at all.</u>

Then $N(r) = 0$ for all $r \in [0, \ell]$ and $\mathcal{D}(x_1, x_2) = \begin{cases} \mathcal{U} & \text{if } x_1 + \frac{\pi x_1 x_2}{q} < l \\ \varnothing & \text{otherwise} \end{cases}$,
and equations (2.1) and (2.2) are obviously fulfilled for all $x_1, r \in [0, \ell]$ and $x_2 \in [0, 1]$. The uniqueness of the equilibrium is obvious as well.

<u>2. Case: A positive mass of agents uses the congested line.</u>

We do not provide a formal proof of equilibrium existence here. Instead, we refer to numerical simulations which robustly confirm the existence of an equilibrium.

Uniqueness:

In order to show uniqueness, we cannot make use of standard theorems for nonlinear integral equations because the integrand function does not fulfill the required conditions.[28] Therefore, we have to exploit the structure of the model.

For any equilibrium define the function $y : [0, \ell] \to [0, 1]$ by

$$y(x_1) := \begin{cases} x_2 & \text{if there exists an agent at location } (x_1, x_2) \\ & \text{who is indifferent between } \mathcal{C} \text{ and } \mathcal{U} \\ 1 & \text{otherwise} \end{cases}$$

Note that $y$ is a well defined function because the concerning cost functions are dependent on the location values, and for all $x_1 \in [0, \ell]$, the cost function for the uncongested line is increasing in $x_2$, whereas the cost function for the congested line is invariant in $x_2$. It is almost everywhere continuous, since the cost functions are continuously dependent dependent on the position.

Now define the function $N$ through $y$:

$$N(z) \quad = \quad \int_z^l \frac{\pi k}{q}(1 - y(k))\, dk, \qquad z \in \ [0, \ell] \tag{2.6}$$

Because we are in the second case, we have a positive mass of agents using the congested line. Since $l = \bar{v}$ per assumption, we see that

$$\{(r, x_2) \in [0, \ell] \times [0, 1] \mid m + \int_0^r 1 + c(N(z))\, dz \leq \ell\}$$

is non-empty. Let us define $r_A$ as the maximal distance, where there exists at least one agent who uses the congested line:

$$r_A = \sup_{r \in [0,\ell]} \{(r, x_2) \in [0, \ell] \times [0, 1] \mid m + \int_0^r 1 + c(N(z))\, dz \leq \ell\},$$

then for $r \in [r^\star, r_A]$ the curve $(r, y(r))$ is characterized by

$$\begin{aligned} \text{costs(congested line)} \quad &= \quad \text{costs(uncongested line)} \\ \Leftrightarrow m + \int_0^r \big(1 + c(N(z))\big)\, dz \quad &= \quad r + \frac{\pi r}{q} y(r) \end{aligned}$$

[28]Most theorems of this kind require integrand function to be continuously differentiable on its whole definition interval which is not the case in our model.

Then $y$ is a solution of the following integral equation

$$\int_0^r c(\int_z^l \frac{\pi k}{q}(1-y(k))\,dk)\,dz = \frac{\pi r}{q}y(r) - m \text{ for all } r \in [r^\star, r_A], \tag{2.7}$$

together with

$$y(r) = 1 \text{ for all } r \in \,]0, r^\star[\, \cup \,]r_A, \ell[. \tag{2.8}$$

Now let $y_1$ and $y_2$ be two functions solving (2.7) and (2.8). Define $r_i^\star$ and $r_{Ai}, i = 1, 2$ accordingly.

Suppose that the corresponding $y$-lines are not identical. This means that there exists a subset $T \subset [0, \ell]$ such that $y_1(r) \neq y_2(r)$ for all $r \in T$.

First case: $r_1^\star = r_2^\star$.

Since $c(N_i)$ is constant in $[0, r_i^\star]$, an agent located at $(r^\star, 1)$ is indifferent between the congested and the uncongested line in both equilibria:

$$r_1^\star + r_1^\star \cdot c(N_1(r_1^\star)) = r_1^\star + \frac{\pi r_1^\star}{q} = r_1^\star + r_1^\star \cdot c(N_2(r_1^\star)),$$

which leads to

$$c(N_1(r_1^\star)) = c(N_2(r_1^\star)).$$

Since $c$ is invertible, we see that the mass of agents using the congested line beyond $r_1^\star$ must be the same in both equilibria such that

$$N(0) := N_1(r_1^\star) = N_2(r_2^\star).$$

Note that as a solution of an integral equation, $N_i(r)$ is a continuously differentiable function on $]r_i^\star, r_{Ai}[$ and, because of its definition as integral function, everywhere continuous. Since $y(r)$ is continuous on $]r_i^\star, r_{Ai}[$, there exists a certain distance beyond $r_1^\star$ at which the $y$-curves of both equilibria intersect (semi-singularily) in one point. (Otherwise we could not have the same mass of agents using the congested line in both equilibria, or both $y$-curves are identical.) Now define

$$\begin{aligned} r^h \;:=\; & \min\{x > r_1^\star |\; y_1(x) = y_2(x) \text{ and} \\ & y_1(x-\varepsilon) \neq y_2(x-\varepsilon) \text{ for sufficiently small } \varepsilon > 0\} \end{aligned}$$

as the (semi-)singular intersection point closest to $r^\star$. Without loss of generality, we assume $N_1(r^h) < N_2(r^h)$. Therefore, we must have

$$y_1(r) \leq y_2(r) \text{ for all } r \in [r_1^\star, r^h[. \tag{2.9}$$

Note that the set $\{x \,|\, r^h \leq x \leq \ell \text{ and } N_1(x) = N_2(x)\}$ is bounded by 0 and $l$, closed and nonempty:

$$\max\{r_{A1}, r_{A2}\} \in \{x \,|\, r^h \leq x \leq \ell \text{ and } N_1(x) = N_2(x)\}$$

Therefore, it must have a minimal element. Define

$$r^k := \min\{x > r^h |\; N_1(x) = N_2(x)\}.$$

Then, we have

$$r^k \leq \max\{r_{A1}, r_{A2}\}. \tag{2.10}$$

Since $N_1(r^h) < N_2(r^h)$, and $r^k$ is the closest intersection point of $N_1$ and $N_2$ to $r^h$ in positive direction, we can easily see that the derivative of $N$ with respect to $r$, evaluated at $r^k$, in equilibrium 1 must be higher or equal than the corresponding derivative in equilibrium 2. By equation (2.6), we get

$$N_i'(r^k) = (y_i(r^k) - 1) \cdot \frac{\pi r^k}{q} < 0.$$

Altogether, we see that we must have

$$y_1(r^k) > y_2(r^k), \tag{2.11}$$

where the strict inequality follows from the fact that

$$\begin{aligned} &N_1(r) \leq N_2(r) \text{ for all } r \in [0, r^k] \text{ and} \\ &N_1(r) < N_2(r) \text{ for a positive mass of positions around } r^h. \end{aligned} \tag{2.12}$$

Hence, all agents located around distance $r_k$ face strictly lower congestion costs on their way to the centre in equilibrium 1 than in equilibrium 2.

Now consider the agent at position $(r^k, y_2(r^k))$. By definition and because of (2.10), in equilibrium 2, he is indifferent between the congested line and the uncongested line:

$$\begin{aligned} U(\mathcal{U}) &= U(\mathcal{C}) \\ \Leftrightarrow l - r^k - \frac{\pi r^k y_2(r^k)}{q} &= l - m - r^k - \int_0^{r^k} c(N_2(r))\,dr \\ \Leftrightarrow \int_0^{r^k} c(N_2(r))\,dr &= \frac{\pi r^k y_2(r^k)}{q} - m \end{aligned}$$

Because of $y_1(r^k) > y_2(r^k)$, in equilibrium 1, he opts for the uncongested line:

$$\begin{aligned} U(\mathcal{U}) &> U(\mathcal{C}) \\ \Leftrightarrow l - r^k - \frac{\pi r^k y_2(r^k)}{q} &> l - m - r^k - \int_0^{r^k} c(N_1(r))\,dr \\ \Leftrightarrow \int_0^{r^k} c(N_1(r))\,dr &> \frac{\pi r^k y_2(r^k)}{q} - m \end{aligned}$$

Combining the equations above, we get

$$\int_0^{r^k} c(N_1(r))\,dr > \int_0^{r^k} c(N_2(r))\,dr$$

But this is a contradiction to (2.12).

<u>Second case: $r_1^\star \neq r_2^\star$.</u>

Without loss of generality suppose $r_1^\star < r_2^\star$. From equation (2.5) we get:

$$\frac{\partial r^\star}{\partial N(0)} = \frac{mc'(N(0))}{(\frac{\pi}{q} - c(N(0)))^2} > 0$$

Therefore, we have $N_1(0) < N_2(0)$.

Let us now prove a helpful lemma.

**Lemma 2.7.1** *If $r_1^\star < r_2^\star$, there must exist a distance $r^j \in ]r_2^\star, l]$, where the corresponding graphs of $N_1(r)$ and $N_2(r)$ intersect: $N_1(r^j) = N_2(r^j) > 0$.*

**Proof.** Consider a marginal road user who is indifferent between no trip and the congested line. His distance to the centre in equilibrium 1 is $r_{A1}$. His overall utility is zero.

$$\begin{aligned} U(\varnothing) &= U(\mathcal{C}) \\ \Leftrightarrow \qquad 0 &= l - m - r_{A1} - \int_0^{r_{A1}} c(N_1(r))\, dr \end{aligned}$$

Suppose the two concerning curves for $N_i(r)$ do not intersect. Then, we have

$$c(N_2(r)) > c(N_1(r)) \quad \text{ for all } r \in [0, r_{A2}[$$

and

$$r_{A2} > r_{A1}.$$

In equilibrium 2, the agent at location $(r_{A2}, 1)$ gets an overall utility of

$$U(\mathcal{C}) = l - m - r_{A2} - \int_0^{r_{A2}} c(N_2(r))\, dr < l - m - r_{A1} - \int_0^{r_{A1}} c(N_1(r))\, dr = 0$$

but by choosing "no trip", he could guarantee an overall utility of 0. Contradiction! q.e.d.

By the lemma above, we see that both curves intersect in at least one point. Define

$$r^z := \min\{r \in [0, \ell] |\; N_1(r) = N_2(r)\}$$

Obviously, $N_1(r) < N_2(r)$ for all $r \in [0, r_z[$. It is easy to see that at the intersection point, we must have $N_2'(r^z) \leq N_1'(r^z) < 0$. By equation (2.6), we get

$$N_i'(r) = (y_i(r) - 1) \cdot \frac{\pi r}{q}.$$

This equation shows us that $y_2(r^z) < y_1(r^z)$, where the strictness of the inequality follows as in case 1. Now consider the agent at location $(r^z, y_2(r^z))$. In equilibrium 2, he is indifferent between the congested line and the uncongested line:

$$\begin{aligned} U(\mathcal{U}) &= U(\mathcal{C}) \\ \Leftrightarrow l - r^z - \frac{\pi r^z y_2(r^z)}{q} &= l - m - r^z - \int_0^{r^z} c(N_2(r))\,dr \\ \Leftrightarrow \int_0^{r^z} c(N_2(r))\,dr &= \frac{\pi r^z y_2(r^z)}{q} - m \end{aligned}$$

We immediately get the following inequality:

$$\int_0^{r^z} c(N_1(r))\,dr < \int_0^{r^z} c(N_2(r))\,dr = \frac{\pi r^z y_2(r^z)}{q} - m < \frac{\pi r^z y_1(r^z)}{q} - m \tag{2.13}$$

In equilibrium 1, the agent at location $(r^z, y_1(r^z))$ is indifferent between the congested line and the uncongested line:

$$\begin{aligned} U(\mathcal{U}) &= U(\mathcal{C}) \\ \Leftrightarrow l - r^j - \frac{\pi r^j y_1(r^j)}{q} &= l - m - r^j - \int_0^{r^j} c(N_1(r))\,dr \\ \Leftrightarrow \int_0^{r^j} c(N_1(r))\,dr &= \frac{\pi r^j y_1(r^j)}{q} - m \end{aligned}$$

But this is a contradiction to (2.13).

q.e.d.

**Proof of 2.4.2.** Point $A$ $(r_A, 1)$ is defined by $U(\mathcal{U}) = U(\mathcal{C}) = U(\varnothing)$. From this equality we get

$$\begin{aligned} \Leftrightarrow \ell - r_A - \frac{\pi r_A}{q} &= \ell - m - r_A - \textstyle\int_0^{r_A} c(N(r))\,dr &&= 0 \\ \Leftrightarrow -\frac{\pi r_A}{q} &= -m - \textstyle\int_0^{r_A} c(N(r))\,dr &&= -\ell + r_A \\ \Leftrightarrow \int_0^{r_A} c(N(r))\,dr &= \tfrac{\pi r_A}{q} - m &&= \ell - r_A - m \end{aligned}$$

Point $B(r_B, 0)$ is defined by $U(\mathcal{C}) = U(\varnothing)$. Since congestion costs are invariant in the second argument of the location, we get $r_B = r_A$.

Agents around position $(\ell, 1)$ always opt for $\varnothing$. Agents around the centre always choose the uncongested link. Since we impose the condition $mq < \pi\ell$, we have a positive mass of road users in equilibrium. Hence, since all cost functions are continuously dependent on location, three connected areas with different modes of use will occur.

q.e.d.

**Proof of 2.4.3.** It is clear that $N(r)$ is monotonically decreasing in $[r^\star, r_A]$. The fact that $N(r)$ is strictly decreasing follows from condition $mq < \pi l$, since the term $\int_0^{\frac{\pi r}{q}} \mathbb{1}_{\{D(r, \frac{kq}{\pi r}) = \mathcal{C}\}} \, dk$ is strictly positive for all $r \in [r^\star, r_A]$.

It remains to be shown that for all $z \in [0, \ell], k \in [0, \frac{\pi z}{q}]$ we have:

$$D(z, \frac{kq}{\pi z}) = \mathcal{C} \Leftrightarrow \int_0^z c(N(x))\, dx \leq \min(k - m, l - m).$$

The agent opts for the congested line if and only if

$$\text{costs(congested line)} \leq \text{costs(uncongested line)} \quad \text{and} \quad \text{costs(congested line)} \leq \bar{v}.$$

Because of $\bar{v} = l$ we get

$$\Leftrightarrow m + \int_0^z \big(1 + c(N(x))\big)\, dx \leq z + k \quad \text{and} \quad m + \int_0^z \big(1 + c(N(x))\big)\, dx \leq l$$

$$\Leftrightarrow \int_0^z c(N(x))\, dx \leq k - m \quad \text{and} \quad \int_0^z c(N(x))\, dx \leq l - m$$

$$\Leftrightarrow \int_0^z c(N(x))\, dx \leq \min(k - m, l - m)$$

q.e.d.

**Proof of 2.4.4.** The proof was already given in the existence part of Proposition 2.4.1.

q.e.d.

**Proof of 2.4.5.** We have to show that an increase in the parameters $m$, $\ell$ or $q$ leads to a (weak) increase in $y(r)$ for all $r \in [r^\star, r_A]$. Concerning the city size $\ell$, we do not prove this result formally, but refer to numerical simulations confirming it.

Consider the representation of $y(r)$ given in Proposition 2.4.4. Note that this equation must hold for all $r \in ]r^\star, r_A[$.

$$\int_0^r c\Big(\int_z^{r_A} \frac{\pi k}{q}(1 - y(k))\, dk\Big)\, dz - \frac{\pi r y(r)}{q} + m = 0. \tag{2.14}$$

The function $y(r)$ is characterized by this equation for all $r \in ]r^\star, r_A[$ together with $y(r) = 1$ for all $r \in ]0, r^\star[ \cup ]r_A, \ell[$, where $r_A$ and $r^\star$ are defined as in the proof of Proposition 2.4.1. Now fix an $r^0 \in ]r^\star, r_A[$ and apply the Implicit Function Theorem on this equation.

$$\frac{\partial y(r^0)}{\partial m} = -\frac{1}{\underbrace{\int_0^{r^0} c'\big(\int_z^{r_A} \frac{\pi k}{q}(1-y(k))\,dk\big)\,dz \cdot \int_z^{r_A}(-1)\frac{\partial y(k)}{\partial y(r^0)} \cdot \frac{\pi k}{q}\,dk - \frac{\pi r^0}{q}}_{<0}} > 0$$

By multiplying equation (2.14) with $q$, we get

$$\int_0^r q \cdot c\Big(\int_z^{r_A} \frac{\pi k}{q}(1 - y(k))\, dk\Big)\, dz - \pi r y(r) + mq = 0 \tag{2.15}$$

In order to prove the last part of the proposition, we make use of the following well known fact: $\frac{\partial a f(\frac{x}{a})}{\partial a}(a) > 0$ for all $a, x \in \mathbb{R}^+$ if $f : \mathbb{R}^+ \to \mathbb{R}^+$ is concave. By the concavity assumption (2) we get

$$\int_0^r \frac{\partial(qc(\frac{1}{q} \cdot \ldots))}{\partial q}\, dz + m > 0$$

and therefore

$$\frac{\partial y(r^0)}{\partial q} = -\frac{\int_0^r \frac{\partial(qc(\frac{1}{q}\cdot\ldots))}{\partial q}\,dz + m}{\underbrace{\int_0^{r^0} c'\big(\int_z^{r_A} \frac{\pi k}{q}(1-y(k))\,dk\big)\,dz \cdot \int_z^{r_A}(-1)\frac{\partial y(k)}{\partial y(r^0)} \cdot \frac{\pi k}{q}\,dk - \frac{\pi r^0}{q}}_{<0}} > 0.$$

q.e.d.

**Proof of 2.5.1.** In Proposition 2.5.2 an even stronger result will be proved. q.e.d.

**Proof of 2.5.2.** Let $V > 0$ and $\varepsilon > 0$. Let the index $b$ refer to the "new" equilibrium with a bottleneck of severeness $V$ at distance $r_A - \varepsilon$. Then $N^b(r, \varepsilon))$ denotes the total number of agents at distance $r$ using the road in the distorted equilibrium with a bottleneck at distance $r_A - \varepsilon$.

Let us first prove a helpful lemma.

**Lemma 2.7.2** *Under the conditions above, we have*

$$(i) \quad \lim_{r \to r_A -} \quad \lim_{\varepsilon \to 0} \frac{\partial N^b(r, \varepsilon)}{\partial \varepsilon} = \frac{\partial N^b(r, e)}{\partial e}(r_A, 0) < 0$$

*and*

$$(ii) \quad \lim_{\varepsilon \to 0} \frac{\partial N^b(r, \varepsilon)}{\partial \varepsilon} = \frac{\partial N^b(r, e)}{\partial e}(r, 0) \leq 0 \text{ for all } r \in [0, r_A - \varepsilon].$$

**Proof.** The first equalities in (i) and (ii) (the continuity of the derivative of $N^b$ with respect to the second argument) will not be proved here because the proof is only of technical nature and the derivatives evaluated at the limit points are not needed any further in the proof. We note them here only for clarification of the terms.

(i) A marginal increase in $\varepsilon > 0$ leads to a potential mass of

$$\int_{r_A - \varepsilon}^{r_A} \frac{\pi r}{q}(1 - y(r))\, dr > 0$$

of road users who must now undergo the bottleneck. Since $V > 0$, a positive fraction $k(\varepsilon) > 0$ of this mass will no longer use the road. Obviously, $k$ is a monotonically non-increasing function in $\varepsilon$. We get

$$N^b(r_A - \varepsilon, \varepsilon) = N(r_A - \varepsilon) - k(\varepsilon) \ \cdot \int_{r_A - \varepsilon}^{r_A} \frac{\pi r}{q}(1 - y(r))\, dr.$$

Since $y(r)$ is monotonically increasing we have

$$N^b(r_A - \varepsilon, \varepsilon) \leq N(r_A - \varepsilon) - k(\varepsilon) \ \cdot \varepsilon \cdot \frac{\pi(r_A - \varepsilon)}{q}(1 - y(r_A - \varepsilon)).$$

This leads to

$$\begin{aligned}
\lim_{r\to r_A-\varepsilon-}\ \lim_{\varepsilon\to 0}\frac{\partial N^b(r,\varepsilon)}{\partial\varepsilon} &= \lim_{\varepsilon\to 0}\tfrac{N^b(r_A-\varepsilon,\varepsilon)-N^b(r_A-\varepsilon,0)}{\varepsilon}\\
&= \lim_{\varepsilon\to 0}\tfrac{N^b(r_A-\varepsilon,\varepsilon)-N(r_A-\varepsilon)}{\varepsilon}\\
&\leq \lim_{\varepsilon\to 0}\tfrac{N(r_A-\varepsilon)-k(\varepsilon)\ \cdot\varepsilon\cdot\frac{\pi(r_A-\varepsilon)}{q}(1-y(r_A-\varepsilon))-N(r_A-\varepsilon)}{\varepsilon}\\
&= -k(0)\ \cdot\tfrac{\pi r_A}{q}(1-y(r_A))\\
&< 0.
\end{aligned}$$

(ii) Now suppose that after a marginal increase in $\varepsilon$, there exists an $r_0\in[0,r_A-\varepsilon]$ such that $\lim\limits_{\varepsilon\to 0}\frac{\partial N^b(r_0,\varepsilon)}{\partial\varepsilon}>0$. Then there exists a $\mu>0$ such that

$$N^b(r_0,\mu)>N^b(r_0,0)=N(r_0).$$

Since we have seen in (i) that $\lim\limits_{r\to r_A-\varepsilon-}\lim\limits_{\varepsilon\to 0}\frac{\partial N^b(r,\varepsilon)}{\partial\varepsilon}<0$ and $N(r)$ is continuous, there must exist an $r_1\ \in[r_0,r_A-\varepsilon]$ such that

$$N^b(r_1,\mu)=N^b(r_1,0)=N(r_1).$$

Define

$$r^z:=\min\{r\in[0,r_A-\varepsilon]|\ N^b(r,\mu)=N(r)\text{ and }N^{b'}(r,\mu)\neq N'(r)\}.$$

Without loss of generality assume $N^b(r,\mu)<N(r)$ for all $r\in[0,r^z]$. Reasoning as in the very last part of Proposition 2.4.1's proof analogously leads to a contradiction. q.e.d.

Let us consider the mass of agents who live beyond the bottleneck (at distance higher than $r_A-\varepsilon$), and use the congested line. In the undistorted equilibrium their contribution to total welfare is

$$\Delta W(\varepsilon)=\int_{r_A-\varepsilon}^{r_A}\underbrace{\frac{\pi k}{q}(1-y(k))}_{\text{number of road users}}\cdot\underbrace{\left[l-m-k-\int_0^k c(N(r))\,dr\right]}_{\text{utility they add to welfare}}dk.$$

By installing a bottleneck at $r_A-\varepsilon$, we lose less than $\Delta W(\varepsilon)$ from total welfare because of the distortion. It is *less* than $\Delta W(\varepsilon)$, because some road users behind $r_A-\varepsilon$

will switch to the uncongested line, or even stay on the road such that they will still contribute positively to the welfare function in the new equilibrium.

On the other hand, the welfare gain due to reduced negative externalities is more than

$$\Delta E(\varepsilon) = \int_0^{r_A-\varepsilon} \underbrace{\frac{\pi k}{q}(1-y(k))}_{\text{number of road users}} \cdot \underbrace{\left[\int_0^k c(N(r)) - c(N^b(r,\varepsilon))\, dr\right]}_{\text{cost difference to new equilibrium}} dk.$$

(In the case of a bottleneck it is even *more* than $\Delta E(\varepsilon)$ because we are only considering a part of all road users who profit from the bottleneck.)

Obviously, we have $\lim_{\varepsilon\to 0} \Delta W(\varepsilon) = \lim_{\varepsilon\to 0} \Delta E(\varepsilon) = 0$. Let us define

$$\Delta X(\varepsilon) = \frac{\partial(\Delta E(\varepsilon) - \Delta W(\varepsilon))}{\partial \varepsilon}.$$

If we are able to show that $\lim_{\varepsilon\to 0} \Delta X(\varepsilon) > 0$ we will have proved that the marginal road user at distance $r_A$ is exerting higher negative externalities on other road users than he adds to total welfare. Therefore, the equilibrium is not welfare maximizing.

By using Leibnitz's Rule for differentiation of integrals we get

$$\begin{aligned}\frac{\partial \Delta E(\varepsilon)}{\partial \varepsilon} &= \int_0^{r_A-\varepsilon} \frac{\pi k}{q}(1-y(k)) \cdot \left[\int_0^k -c'(\ldots) \cdot \frac{\partial N^b(r,\varepsilon)}{\partial \varepsilon}\, dr\right] dk + \\ &\quad \frac{\pi(r_A-\varepsilon)}{q}(1-y(r_A-\varepsilon)) \cdot \left[\int_0^{r_A-\varepsilon} c(N(r)) - c(N^b(r,\varepsilon))\, dr\right] \cdot (-1)\end{aligned}$$

If we let $\varepsilon \to 0$, the second summand becomes 0, since

$$\lim_{\varepsilon\to 0} \left[c(N(r)) - c(N^b(r,\varepsilon))\right] = 0.$$

Given that $c' > 0$ and because of Lemma (2.7.2) the first summand is strictly positive for $\varepsilon \to 0$. Altogether this leads to

$$\lim_{\varepsilon\to 0} \frac{\partial \Delta E(\varepsilon)}{\partial \varepsilon} > 0.$$

Now consider the derivative of the contribution to total welfare:

$$\frac{\partial \Delta W(\varepsilon)}{\partial \varepsilon} = \frac{\pi(r_A-\varepsilon)}{q}(1-y(r_A-\varepsilon)) \cdot \left[l - m - (r_A-\varepsilon) - \int_0^{r_A-\varepsilon} c(N(r))\, dr\right]$$

Because of the definition of point A and continuity, if we let $\varepsilon \to 0$, the term in squared brackets becomes 0. Therefore we get

$$\lim_{\varepsilon \to 0} \frac{\partial \Delta W(\varepsilon)}{\partial \varepsilon} = 0$$

and altogether

$$\lim_{\varepsilon \to 0} \Delta X(\varepsilon) > 0.$$

By setting $\varepsilon > 0$ sufficiently small in order to guarantee that $\Delta X(\varepsilon) > 0$ and by setting $d := r_A - \varepsilon$ we get desired result.
q.e.d.

**Proof of 2.5.6.** Analogously to Proposition 2.5.2.
q.e.d.

# Chapter 3

# Dynamics in the Rental Real Estate Market

## 3.1 Introduction

In rental housing markets a large heterogeneity of both lessors and lessees prevails. Landlords of rental real estate[1] differ, among other criteria, in the costs associated with the process of switching renters which we will call processing costs.[2] Large and professional lessors tend to have much lower processing costs than private households with only one unit of real estate to let which can be explained easily by a decreasing average cost function in the number of houses let. Renters differ, among other criteria, with respect to their income variance, their expected length of stay in the dwelling and their probability of defaulting. The parameters described above exhibit a strong influence on the prevailing supply and demand curves. In this chapter, we study the effects of the two-sided-heterogeneity explained above on the distribution of vacant houses, the distribution of rental prices and the distribution of renter quality. Furthermore, we study the dynamics of rental housing markets after exogenous shocks to processing costs.

[1] In what follows we denote any property to rent by "house".

[2] Processing costs include for example eviction costs, renovation costs, legal fees, advertising costs, and costs for the real estate agency.

The economic and legal relationship between lessors and lessees is crucial when studying rental real estate markets. Especially in highly regulated rental markets with strict legal regulations[3], typical moral hazard problems occur. The incentives of the renter to maintain the house properly and to pay the rent on time may be distorted because of legal reasons. The renter may anticipate that it is costly and time-consuming for the landlord to undergo legal proceedings which leads to inefficient behavior, such as not properly maintaining the house or defaulting on rental payments.

On the other side, landlords try to mitigate this problem by selection procedures: Potential candidates are screened with respect to their income and the respective variance. The landlord also has an incentive to minimize the probability that the renter defaults or moves out soon which leads to discrimination against young renters[4] and against renters with a high probability of defaulting or moving out.[5] By setting an appropriate rental price landlords influence the probability of finding an appropriate renter: Lowering the rental price creates a larger set of potential lessees willing to pay the imposed rent, but also lowers the rent income while a high rental price lowers the set of potential tenants.

Because of heterogenous processing costs on the side of the landlords, different preferences with respect to renter quality will prevail in the market: Landlords who face high processing costs have an incentive to select rental candidates more rigorously than landlords with low processing costs. In order to find a renter of appropriate quality, they have to accept a lower rental price than landlords with low processing costs. The resulting cost-dependent equilibrium rent is representing the efficiency rent. This concept of efficiency rents is similar to the concept of efficiency wages in the economic theory of labour markets (Stiglitz, 1974) or the theory of efficiency interest rates (e.g. Stiglitz and Weiss, 1981).

[3]We do not necessarily refer to a situation where the legal procedures of eviction are complicated and costly, our model also incorporates markets where defaulting is (nearly) impossible and the landlord only has to bear the costs when renters move out and do not present the next tenant.

[4]Among others, Guiso, Jappelli and Pistaferri (2002) provide empirical evidence that the variance of income and the employment risk is negatively correlated with age.

[5]In the following we integrate these renter properties in the term "quality".

We model the rental market for real estate in a stylized, but intuitive manner. In order to abstract from substitution or income effects, in our model all dwellings are of the same size and quality. Dependent on the equilibrium distribution of renter qualities, a certain fraction of houses is vacant. Landlords differ only in their processing costs. In each period, a certain number of potential renters who differ only in quality, enter the market and create a standard demand function decreasing in the rental price. The quality of a renter is represented in terms of defaulting probability on rental payments or more generally by the probability of moving out.[6] We assume that the willingness to pay in terms of the rental price for a house is negatively correlated with renter quality.[7] Furthermore, from landlords' point of view, the renter quality associated with a certain rental price positively depends on the market power of landlords which is represented by a function of the number of houses on the market.

The higher the processing costs, the stricter the landlords' requirements regarding renter quality have to be. Landlords with low processing costs can afford to let their houses to renters with relatively high probability of moving out or defaulting, whereas landlords with high processing costs only accommodate high quality renters, but they also accept a lower stream of rental payments. The prevailing vacancy rate is determined by the qualities of actual renters. We show that higher average processing costs lead to a lower vacancy rate in the long run.

Furthermore, we study the dynamic short time effects of an increase of market frictions due to processing costs.[8]

Holding everything else constant, we find that a change in these costs will lead to an overreaction of the average rental price: For example, after an increase in processing costs, the average rental price will overshoot its new equilibrium value in the first periods after the shock.

[6]More formally, the quality is the probability of not terminating the rental contract in one period of time.

[7]A simple rationale for this assumption is the existence of market power on the side of the renters.

[8]An example for an exogenous shock to processing costs is the unexpected imposition of restrictive legal regulations in the rental market such as administrative obligations or legal fees regarding eviction procedures.

The intuition is simple: We distinguish between two effects influencing the average rent. First, an equally distributed increase in processing costs leads to lower average rents because all landlords want to choose higher renter qualities on average. The second effect concerns the market power of landlords. Due to relatively low processing costs in the old steady state, landlords have chosen high rental prices and relatively low renter qualities. This translates into many vacant units of housing on the market and to low market power of landlords. Directly after the shock to processing costs, the low market power of landlords is not altered. In contrast, in the new steady state, the market power of landlords is relatively high due to few vacant units of housing on the market.

In sum, the shock to processing costs leads to a strongly reduced average rent. In the long run, this negative effect on rents is mitigated by the effect of increased market power of the landlord.

We will proceed as follows. In section 2, we describe some approaches that explain the structure and the consequences of the lessor-lessee relation in the rental market and compare them with our approach. The formal model is developed in section 3. The resulting equilibrium distributions of rental prices and qualities of renters are studied in section 4. In this section we also present the core part of this work: the dynamic short-time effects of an unanticipated shocks to processing costs. Finally, in the last part, we conclude.

## 3.2 Related Literature

One of the first articles describing the special features of the rental housing market is Olsen (1969). In this article the author studies rather informally to what extent and under which conditions the real estate market differs from a standard competitive market. A general overview of the early theory of rental housing markets is given by Arnott (1987). Among others, Lind (2001) gives an overview on the legal regulations in the rental housing market and categorizes different types of regulation.

Arnott (1989) examines the vacancy rate in rental housing markets. Landlords exert monopoly power over tenants, which leads to a market rent above costs. Nevertheless, in this model free entry and exit leads to zero profits of landlords. The equilibrium is created by the mechanism of vacancies. Although this model features heterogeneity of vacant housing units, this approach shares an important feature with our model: a high vacancy rate is useful to potential renters, because they have better choice possibilities between vacant units. Therefore their market power is higher.

A paper which shares some intuitions with our work is Wasmer (2004). There the author makes use of search models usually applied in the theory of labor markets to explain some characteristics of the housing market. In a continuous time setting, a matching and screening mechanism is introduced. This model aims to study the effects of changes in the costs associated with legal eviction procedures to the market outcome. It is shown that higher legal costs lead to more selective landlords, higher rents and to lower vacancy rates.

Hubert (1995) studies the so-called tenure discount in rental housing markets. In this paper an adverse selection problem arises due to the asymmetric information in the market. It is explained why long-standing tenants pay lower rents than average tenants. This paper focuses on "service costs", which are the costs caused by the renter during his stay in the dwelling, such as maintenance costs and administrative costs. It is shown that landlords charge bad tenants strictly higher rents than good tenants in the case of a renewal of the rental contract.

Similarly, Basu and Emerson (2003) consider a model of tenancy rent control under asymmetric information. If there is any inflation in the economy, landlords tend to let their houses to short-staying tenants. Furthermore, they show that under certain circumstances it may be optimal for landlords not to increase the rent, even when housing demand exceeds supply, which they call efficiency pricing.

Deng, Gabriel and Nothaft (2002) study the duration of residence in rental housing. They provide empirical evidence that this duration is dependent on structural characteristics and on renter characteristics. We use their results as a foundation of our assumption that the vacancy rate and the turnover per period are determined by the quality distribution of renters in the market. Furthermore, they show that changes in the legal environment, e.g. the imposition of rent control mechanisms, have strong influence on the average duration of residence, which indicates a change in the distribution of renters.

Wheaton (1990) constructs a search model of the housing market which explains a strong negative relationship between market prices and vacancy rate. His approach serves as an explanation of structural vacancy in housing markets.

A further model which highlights the matching aspect in rental housing markets is Anas (1997). He shows that under certain circumstances, rent regulation and central matching mechanisms in the housing market can lead to welfare improvements. Interestingly, in his model the regulated rent is set above the market rent in order to create more supply of housing, which in turn improves the match.

Read (1991, 1993) studies the quality dispersion of houses and vacancies in the rental market characterized by imperfect information on the part of the tenants. In these models the rental price of houses is fixed and there are no asymmetries between landlords.

Read (1997) builds a stochastic search model of imperfect information where the vacancy rate is induced by free entry of landlords in the rental market. He examines the influence of exogenous parameters on the vacancy rate. In particular, he shows that the vacancy rate falls with increases in the cost of housing provision, which is in line with our results. In this model, potential tenants can be completely described by an exogenously given demand curve, a feature which we will adopt in similar form.

In a theoretical approach, Read (1988) shows that the vacancy rate in the rental housing market is dependent on the renter's behavior when moving out and on the advertising by landlords. In this model a searching mechanism is introduced such that a natural vacancy rate occurs due to landlords who do not find an appropriate renter. The choice variable of landlords in the model is not the price, but the degree of advertising, which endogenously determines the vacancy rate.

All these approaches do either not consider double heterogeneity on both sides of the market or do not analyse the transitional dynamics of exogenous shocks to the landlord's cost characteristics. With this work, we want to fill this gap in the literature.

## 3.3 The Model

In the following the formal model is developed. In order to capture the effects of landlords' cost-dependent tenant selection, the dynamic setting must be rich enough to capture heterogeneity on the side of the landlords and on the side of the tenants as well.

**Timing.** Time is discrete. The game is played infinitely many periods. Landlords discount the future by $0 < \delta < 1$ per period. In each period a rented house becomes vacant with probability $q$ which is equal to the renter's quality. If a house is vacant the landlord sets a rent $R$ and attracts a renter of quality $q(R, m)$ non-stochastically, where $m$ is the prevailing vacancy rate.

### 3.3.1 Supply Side of the Market

**Housing Stock.** For simplicity we assume that all dwellings offer the same housing services. There is a fixed stock of housing of mass 1. Each dwelling is owned by exactly one risk-neutral landlord.

**Landlords.** Landlords differ in their processing costs $c$. They are uniformly distributed on $[c_0, c_0 + \Delta] \times [0, \frac{1}{\Delta}]$, where $0 \leq c_0$ and $\Delta > 0$. The state of a house can either be "vacant" or "rented". In case of a rented house, landlords get a payoff of $R$ per period, where $R$ is the contractually determined rent. Each landlord can be identified[9] by $(c, i)$, where $c \in [c_0, c_0 + \Delta]$ and $i \in [0, \frac{1}{\Delta}]$. Landlord $(c, i)$ quotes the rental price $R(c)$. The maximum imposable rent is set to $R_{\max}$.[10]

[9]The index $i$ serves only for notational convenience. Landlords' decisions are only dependent on their costs $c$, but not on $i$. However, for $i_1 \neq i_2$ the state of the house of landlord $(c, i_1)$ can differ from the corresponding state of landlord $(c, i_2)$.

[10]In many countries, the legal rent control system sets factual maximal rents. Another explanation for this assumption is that $R_{\max}$ is larger than the maximal willingness to pay in the rental market.

### 3.3.2 Demand Side of the Market

**Tenants.** Potential tenants differ with respect to their quality $q \in [0,1]$. $q$ is the per-period probability of quitting the house. Renters' willingness to pay for a house in terms of the maximal accepted rental price is decreasing in their quality.[11] We model the demand side of the model as a function[12] $q(R,m)$ which determines the quality of the tenant a landlord attracts when quoting the rent $R$ at a prevailing vacancy rate of $m$. The market is perfect in the sense that we abstract from any problems arising in search models, such as uncertainty in the matching process. In particular, landlords are always matched to tenants. The function $q$ is assumed to be twice continuously differentiable in both arguments. Furthermore, for all $m \in ]0,1[$ and $R \in [0, R_{\max}]$, we impose the following conditions on $q$:

1. $q(0,m) < 1$
2. $\frac{\partial q(R,m)}{\partial R} < 0$
3. $\frac{\partial q(R,m)}{\partial m} \leq 0$
4. $\frac{\partial^2 q(R,m)}{\partial m \partial R} \leq 0$
5. $\frac{\partial^2 q(R,m)}{(\partial R)^2} < 0$

**Comments on the Conditions.**
Condition (1) guarantees that tenants of quality 1 do not exist in the market. Even if the vacancy rate is very small and the quoted rent is zero, the attracted tenant terminates the rental contract with strictly positive probability in each period.

Condition (2) states that the market power of tenants is positive. The maximal rental

[11] This structure can be easily created by a search model in which tenants are matched randomly to landlords, but face a certain discount rate at each matching process. However, modelling the tenant's decision explicitly would not add further insights to the model.

[12] Tenants take the prevailing rental prices as given. See [42] for a similar setting, where the market demand function is derived by exogenously given individual demand functions.

price lessees accept is decreasing in their quality. From the landlord's perspective this means that the higher the quoted rent, the lower is the corresponding quality of the tenant.

Condition (3) states that from the lessor's point of view, tenant quality is decreasing in the vacancy rate. Intuitively, $m$ is a measure of competition between landlords: The higher $m$, the more landlords are on the market and the stronger is the competition between them. Therefore, at high values of $m$ the relative market power of lessees is strong which leads to relatively low qualities of tenants on average.

Condition (4) states that the cross derivative of the quality function is non-positive. Concerning the market side of the tenants, this assumption states that the relative market power of tenants and the rental price itself act as substitutes when determining the market quality. With increasing market power of tenants (increasing $m$), tenants react less sensitively to changes in the rental price.

The last condition ensures that the landlord faces a concave maximization problem. Due to this assumption the first order conditions describe the landlords' decision problem.

### 3.3.3 Vacancy Rate

Let $v_t(c,i)$ be a function determining the state of the house of landlord $(c,i)$ in period $t$:

$$v_t(c,i) = \begin{cases} 1 & \text{if the dwelling of landlord } (c,i) \text{ is vacant in period } t \\ 0 & \text{otherwise} \end{cases}$$

The market vacancy rate $m_t$ in period $t$ is defined as the fraction of houses actually on the market and not occupied by a renter:

$$m_t = \int_{c_0}^{c_0+\Delta} \int_0^{\frac{1}{\Delta}} v_t(c,i)\; di\, dc.$$

Let $\hat{q}_t(c,i)$ denote the tenant quality of landlord $(c,i)$ in period $t$. The quality of the tenant determines the probability of the termination of a rental contract. The expectation of the state of the house is determined by the quality of the tenant in the preceding period:

$$\mathbb{E}(v_t(c,i)) = 1 - \hat{q}_{t-1}(c,i).$$

Since we are in the case of a continuum of independently distributed random variables, we get

$$m_t = \int_{c_0}^{c_0+\Delta} \int_0^{\frac{1}{\Delta}} \big(1 - \hat{q}_{t-1}(c,i)\big)\, di\, dc. \tag{3.1}$$

### 3.3.4 Problem of the Landlord

The discounted time value of a vacant house shall be denoted by $z$. By setting a rent an appendant tenant quality $q(R,m)$ is obtained. In any following period the lessee quits with probability $1-q$. In the case of a quitting lessee the landlord gets $\delta(z-c)$, where $c$ are the processing costs of the landlord. The following depiction shows the recursive income structure for a landlord with a vacant house who sets the rent $R$ and gets a tenant of quality $q = q(R,m)$, where $m$ is the prevailing vacancy rate:

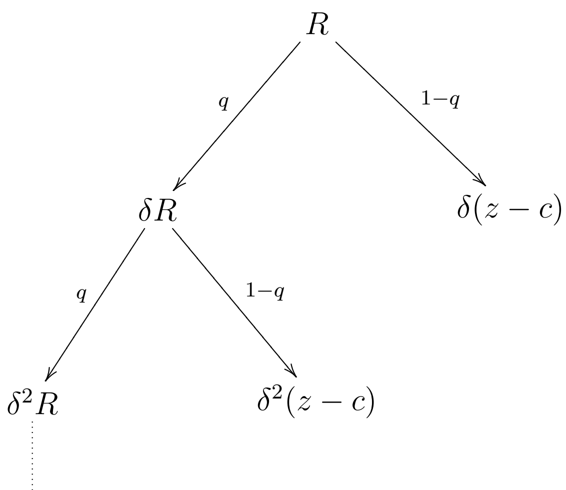


Figure 3.1: Recursive Income Structure for a Landlord

For $\delta < 1$, this recursive structure leads to the following maximization problem of the landlord

$$\max_{R \in [0, R_{\max}]} \frac{R - \delta c(1 - q(R, m))}{1 - \delta} \tag{3.2}$$

**Proof.**

For $\delta < 1$ we get:

$$\begin{aligned} z &= R + \delta q R + \delta(1-q)(z-c) + \delta^2 q^2 R + \delta^2 q(1-q)(z-c) + \\ &\quad \delta^3 q^3 R + \delta^3 q^2 (1-q)(z-c) + \ldots \\ &= R \sum_{k=0}^{\infty} (\delta q)^k + (1-q)(z-c)\delta \sum_{k=0}^{\infty} (\delta q)^k \\ &= \frac{R}{1-\delta q} + \frac{(1-q)(z-c)\delta}{1-\delta q}. \end{aligned}$$

Rearranging terms leads to

$$z = \frac{R - c(1-q)\delta}{1-\delta}.$$

## 3.4 Equilibrium

In this section we define the equilibrium concept and study the resulting distributions of rental prices and appendant tenant qualities. We show how rental prices and the vacancy rate react in response to exogenous shocks to the cost parameters.

### 3.4.1 Equilibrium Definition

First, let us fix the vacancy rate and define the appropriate equilibrium concept.

**Definition 3.4.1 (Static Solution)** *A static equilibrium*[13] $(R^*(c), m)$ *is defined as a function*

$$\begin{aligned} R^* : [c_0, c_0 + \Delta] &\to [0, R_{\max}] \\ c &\mapsto R^*(c) \end{aligned}$$

[13] Although we do not explicitly model tenants' strategies, we implicitly define their strategic interaction by the function $q$. In our setting, tenants take the prevailing rental prices as given. Nevertheless, we call the optimal solution to the problem of the landlord an equilibrium.

*together with a vacancy rate $m \in [0,1]$ such that $R^*(c)$ solves the optimization problem (3.2) for all $c \in [c_0, c_0 + \Delta]$.*

**Law of Motion.**

The law of motion says that in period $t$, landlord $(c, i)$ must attract a new tenant with probability $1 - \hat{q}_{t-1}(c, i)$. The quality of the new tenant is dependent on the rent and on the vacancy rate in period $t$, which is determined by equation (3.1). In steady state, the vacancy rate and the distribution of tenant qualities stay constant.

**Definition 3.4.2 (Steady State Equilibrium)** *A steady state equilibrium is defined as a static equilibrium $(R^*(c), m^*)$ such that the vacancy rate $m^*$ is determined by the equilibrium distribution of qualities:*

$$m^* = \frac{1}{\Delta} \int_{c_0}^{c_0+\Delta} \big(1 - q(R^*(c), m^*)\big)\, dc.$$

### 3.4.2 Equilibrium Properties

**Proposition 3.4.3** *Given a vacancy rate $m > 0$ and $\delta < 1$, each static equilibrium is unique. Inner solutions are characterized by the following set of first order conditions:*

$$1 + \delta c \frac{\partial q(R, m)}{\partial R} = 0 \quad \forall\ c \in [c_0, c_0 + \Delta]. \tag{3.3}$$

*A sufficient condition for the existence of an inner solution is the following condition:*

$$-\frac{1}{\delta(c_0 + \Delta)} < \frac{\partial q(0, m)}{\partial R} \ \textit{and}\ -\frac{1}{\delta c_0} > \frac{\partial q(R_{\max}, m)}{\partial R}. \tag{3.4}$$

**Proof.** Uniqueness: For $\delta < 1$, uniqueness follows from the fact that $\frac{\partial q(R,m)}{\partial R}$ is strictly decreasing in $R$.

Existence of inner solutions: The mean value theorem of differential calculus states that $\frac{\partial q(R,m)}{\partial R}$ runs through all values between $\frac{\partial q(R_{\max},m)}{\partial R}$ and $\frac{\partial q(0,m)}{\partial R}$. By assumption, we get $[-\frac{1}{\delta c_0}, -\frac{1}{\delta(c_0+\Delta)}] \subset [\frac{\partial q(R_{\max},m)}{\partial R}, \frac{\partial q(0,m)}{\partial R}]$. Hence for all $c \in [c_0, c_0 + \Delta]$, an $R^*$ in $[0, R_{\max}]$ exists such that $-\frac{1}{c\delta} = \frac{\partial q(R^*,m)}{\partial R}$, which is equivalent to the first order conditions.

q.e.d.

The first order conditions state that at the optimum, the marginal benefit of a rent increase (equal to 1) must equal the marginal costs expressed as the discounted increase in costs due to the corresponding quality reduction of the tenant.

**Definition 3.4.4 (Average Rent)** *The average rent in the market is defined as*

$$\bar{R} = \frac{1}{\Delta} \int_{c_0}^{c_0+\Delta} R(c)\, dc.$$

We are now ready to describe the distribution of rental prices and tenant qualities:

**Proposition 3.4.5** *In a static equilibrium, $R^*(c)$ is decreasing in $c$, whereas the equilibrium quality $q^*(c)$ is increasing in $c$.*

**Proof.** The Implicit Function Theorem applied to (3.3) shows that

$$\frac{\partial R^*(c)}{\partial c} = -\frac{\frac{\partial q}{\partial R}}{c \frac{\partial^2 q}{(\partial R)^2}} < 0$$

Since $q^*(c) = q(R^*(c), m)$ and $\frac{\partial m}{\partial c} = 0$, we get

$$\frac{\partial q^*(c)}{\partial c} = \frac{\partial q}{\partial R} \cdot \frac{\partial R^*}{\partial c} + \frac{\partial q}{\partial m} \cdot \frac{\partial m}{\partial c} > 0.$$

q.e.d.

The Proposition above describes the behaviour of landlords in the market: Landlords with high processing costs are more picky than those with low processing costs and set the rent lower in order to get a better tenant quality.

**Proposition 3.4.6 (Comparative Statics)** *In a static equilibrium, the average rent is decreasing in the discount factor $\delta$ and in the vacancy rate $m$.*

**Proof.** Exactly the same reasoning as in (3.4.5) shows that $\frac{\partial R^*(c)}{\partial \delta} < 0$ for all $c \in [c_0, c_0 + \Delta]$. Therefore we get $\frac{\partial \bar{R}}{\partial \delta} = \frac{1}{\Delta} \int_{c_0}^{c_0+\Delta} \frac{\partial R^*(c)}{\partial \delta}\, dc < 0$.

Application of the Implicit Function Theorem and the use of assumptions (4) and (5) shows that

$$\frac{\partial R^*}{\partial m} = -\frac{\frac{\partial^2 q}{\partial R \partial m}}{\frac{\partial^2 q}{(\partial R)^2}} < 0.$$

Hence we get

$$\frac{\partial \bar{R}}{\partial m} = \frac{1}{\Delta} \int_{c_0}^{c_0+\Delta} \frac{\partial R^*(c)}{\partial m} \, dc < 0.$$

q.e.d.

Intuitively spoken, this proposition describes the following effects: Increasing the discount factor means the future becomes more important and therefore landlords tend to avoid the cost $c$ associated with the tenant switching process. Hence they want to transfer a higher tenant quality into the next period, which is achieved by a lower rent.[14] A higher vacancy rate means less market power for the landlords, therefore they must lower the rental price in order to attract an tenant of appropriate quality.

### 3.4.3 Steady State Equilibrium

**Proposition 3.4.7 (Steady State)** *Under the conditions (3.4) and if* $|\frac{\partial q(R,m)}{\partial m}| < 1$ *for all* $m \in [\,0,1]$ *and* $R \in [\,0, R_{\max}]$*, all quoted rental price functions* $R : [c_0, c_0 + \Delta] \rightarrow [\,0, R_{\max}]$ *lead to a uniquely*[15] *defined vacancy rate* $m^*$*.*

**Proof.** We prove that for all functions $R : [c_0, c_0 + \Delta] \rightarrow [0, R_{\max}]$ the equation

$$m = \frac{1}{\Delta} \int_{c_0}^{c_0+\Delta} \big(1 - q(R(c), m)\big) \, dc$$

has a unique solution $m^*$. This can be shown by the fixed point theorem of Banach: Define the operator $T : [0,1] \rightarrow [0,1]$ as

$$T(m) = \frac{1}{\Delta} \int_{c_0}^{c_0+\Delta} \big(1 - q(R(c), m)\big) \, dc.$$

[14]Note that we focus on the costs associated with renter switching. In our setting the probability of rental payments is not influenced by the tenant quality.

[15]This uniqueness condition, together with the fact that for all vacancy rates $m^* \in ]0,1[$ a uniquely determined quoted rental price function $R^* : [c_0, c_0 + \Delta] \rightarrow [\,0, R_{\max}]$ exists, should give rise to a unique steady state equilibrium.

Since the operating space $[0,1]$ is compact and $|\frac{\partial q}{\partial m}| < 1$, we get

$$\begin{aligned}|T(m) - T(m')| &= |\frac{1}{\Delta}\int_{c_0}^{c_0+\Delta} \big(q(R(c), m') - q(R(c), m)\big)\, dc| \\ &\leq \frac{1}{\Delta}\int_{c_0}^{c_0+\Delta} |q(R(c), m') - q(R(c), m)|\, dc \\ &< \frac{1}{\Delta}\int_{c_0}^{c_0+\Delta} |m - m'|\, dc \\ &= |m - m'|\end{aligned}$$

Hence $T$ is a contraction mapping. By the Fixed Point Theorem of Banach, we see that $T$ has a unique solution $m^*$.

q.e.d.

### 3.4.4 Dynamics of Exogenous Shocks

Suppose processing costs undergo a uniform shock of $s > 0$, that is, all landlords face the same increase.[16] The processing costs after the shock are distributed uniformly on $[c_0', c_0' + \Delta]$, where $c_0' = c_0 + s$. Let us consider the case of small shocks to processing costs, such that $s$ is sufficiently small. Then the conditions (3.4) continue to guarantee the existence of a unique steady state equilibrium. All variables referring to the new steady state equilibrium are denoted by superscript $**$, while the values referring to the equilibrium in the first period after the shock[17] are denoted by $+$.

**Proposition 3.4.8 (Effect on Vacancy Rate and Prices)** *If the vacancy rate has sufficiently low influence*[18] *on the tenant quality, that is* $|\frac{\partial q(R,m)}{\partial m}|$ *and* $|\frac{\partial^2 q(R,m)}{\partial R \partial m}|$ *both are*

---

[16]For concreteness, we only consider the case of $s > 0$. The case of $s < 0$ is symmetric.

[17]Our setting is general enough to cover both extreme cases of agents' future expectations after a shock. Assuming perfect rational foresight or myopic behaviour both lead to the same qualitative effects. Nevertheless, for simplicity we can assume that agents act myopically and build their optimal decisions only on the current state of the economy.

[18]This situation occurs for example in markets which are close to monopoly. See Basu and Emerson (2003) for an argument that this is the case in most rental housing markets.

*sufficiently small for all* $R \in [0, R_{\max}]$ *and* $m \in ]0,1[$*, the steady state vacancy rate* $m^*$ *and the steady state average rent* $\bar{R}^*$ *are decreasing in* $c_0$*.*

**Proof.** See Appendix.

This proposition confirms the main intuitions: After a uniform increase in costs, landlords become more selective with respect to the tenant quality. Therefore, they must charge a lower average rental price. An increase in the average quality of tenants means less tenants moving in the market and therefore a lower vacancy rate.

Under the conditions of (3.4.8), we get the following overreaction pattern of rental prices.

**Proposition 3.4.9 (Adverse Shock)** *After a uniform shock* $s > 0$ *to processing costs, in the first periods after the shock the average rent* $\bar{R}^+$ *overshoots its new steady state value* $\bar{R}^{**}$*, that is*

$$\bar{R}^+ < \bar{R}^{**} < \bar{R}^*$$

**Proof.** As we have seen, the vacancy rate in the new steady state is lower than in the old steady state. However, the vacancy rate is not adjusting immediately, but only with rate $m$ per period, since only a fraction $m$ of tenants is replaced per period. In the static equilibrium of period $+$, landlord $(c, i)$ faces costs of $c^{**} = c + s$ and a prevailing vacancy rate of $m^*$. As we have seen in (3.4.6), in a static equilibrium, the effect of the vacancy rate on the average rent is negative. The rent of landlord $(c, i)$ in period $+$ is therefore lower than in the new steady state.

q.e.d.

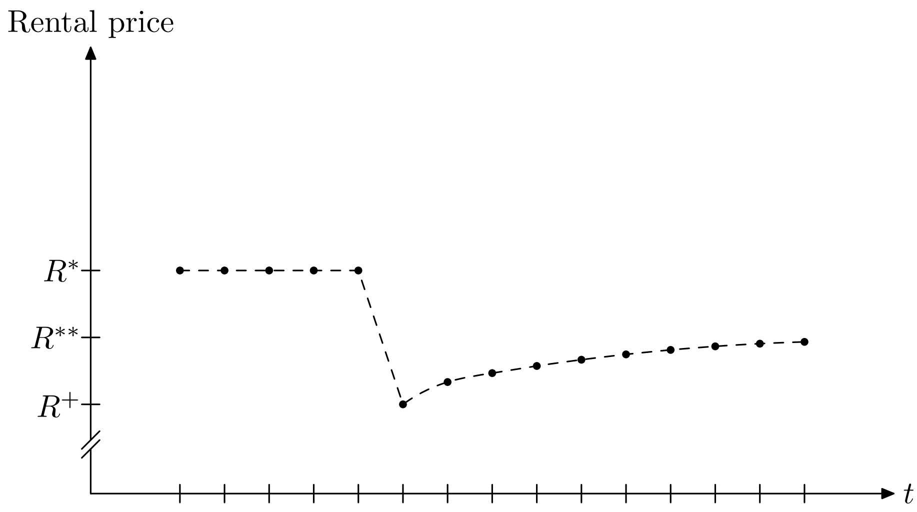


Figure 3.2: Overshooting of Rental Prices

The intuition is the following: Suppose again we are in the situation of an adverse shock $s > 0$ to $c$. Landlords react by increasing the rent, because the quality of the tenant has become more important. Since the quality of tenants before the shock was relatively low, directly after the shock there is still a larger number of vacant houses on the market than in the new steady state. This large number of vacant houses directly after the shock adds further down pressure on rental prices. However, this effect is mitigated successively by the fact that the vacancy rate decreases steadily by the entry of new high quality tenants. In the long run, relatively few vacant houses are on the market due to the increased average quality of tenants.

## 3.5 Summary and Conclusion

In this chapter we formalize the idea of efficiency rents in a simple way: Landlords choose the rental price according to their processing costs which include for example agency costs, costs for advertising and possible legal fees. In the model equilibrium, low renter qualities are matched to landlords with low processing costs, while landlords with high processing costs choose high renter qualities. In steady state, the vacancy rate is determined by the distribution of renter qualities on the market. In the model, the vacancy rate is positively correlated to the market power of renters. We study the dynamics of a uniform shock to processing costs. Two effects are crucial in the transition process of average prices:

First, the effect on processing costs for all landlords translates into a corresponding change in the desired renter qualities and therefore in the rents.

The second effect concerns the vacancy rate: Since after the shock, the average quality of new renters is raised, the vacancy rate adapts as well which means that less real estate is vacant. However, this second effect is not immediate because only a fraction equal to the vacancy rate of all actual renters are changed each period. Both effects work in opposite direction. It is found that the average rental price in the market exhibits an overshooting pattern. Directly after the shock, the average price is sinking below its new steady state level. This effect is caused by the slow adaption of the vacancy rate: In the first period after the shock to processing costs, the prevailing vacancy rate is still the same, and therefore cannot mitigate the effect of changed processing costs of landlords. In the long run, the vacancy rate adapts to its steady state level and moderates the first effect.

In this model we make use of a generalized tenant demand structure and do not model the renter's decision problem explicitly. In order to calibrate and simulate the model with real world data, our model can be used as a foundation to construct a search model which admits for the same behaviour of the demand side as we impose in the model and is amenable to empirical verification.

## 3.6 Appendix

**Proof of Proposition** 3.4.8.

Consider the first order conditions

$$1 + \delta c \frac{\partial q(R^*(c), m^*)}{\partial R} = 0 \quad \forall\, c \in [c_0, c_0 + \Delta], \tag{3.5}$$

together with the equation determining the steady state vacancy rate,

$$m^* - 1 + \frac{1}{\Delta} \int_{c_0}^{c_0+\Delta} q(R^*(c), m^*)\, dc = 0, \tag{3.6}$$

and the definition of the average steady state rent

$$\bar{R}^* = \frac{1}{\Delta} \int_{c_0}^{c_0+\Delta} R^*(c)\, dc.$$

Let us now consider a fixed $c \in [c_0', c_0' + \Delta] \cap [c_0, c_0 + \Delta]$. (Since we only consider the case where the shock $s$ is sufficiently small, the set $[c_0', c_0' + \Delta] \cap [c_0, c_0 + \Delta]$ must be non-empty.) By plugging (3.6) in (3.5), we get

$$1 + \delta c \frac{\partial q(R^*(c), 1 - \frac{1}{\Delta} \int_{c_0}^{c_0+\Delta} q(R^*(c), m^*)\, dc)}{\partial R} = 0.$$

Applying the Implicit Function Theorem and Leibnitz's Rule on this equation for fixed $c$ and using the conditions for $q$, we get

$$\frac{\partial R^*(c)}{\partial c_0} = - \frac{\overbrace{\frac{\partial^2 q}{\partial R \partial m}}^{\leq 0} \overbrace{(-\frac{1}{\Delta})}^{<0} \overbrace{\Big(q(R^*(c_0 + \Delta), m^*) - q(R^*(c_0), m^*)\Big)}^{>0}}{\underbrace{\frac{\partial^2 q}{(\partial R)^2}}_{<0} + \underbrace{\frac{\partial^2 q}{\partial R \partial m}}_{\leq 0} \underbrace{\Big( - \frac{1}{\Delta} \int_{c_0}^{c_0+\Delta} \frac{\partial q}{\partial R}\, dc)\Big)}_{>0}} \geq 0$$

This equation states that for **fixed** $c$, the steady state equilibrium rent is increasing in $c$. If we fix $c$, the only effect which comes into play is the change in the vacancy rate. An increase in $c_0$ will increase average renter qualities in the market. Hence the vacancy rate decreases which leads to more market power for landlords and therefore to higher rents.

Leibnitz's rule for the differentiation of integrals leads to

$$\frac{\partial \bar{R}^*}{\partial c_0} = \frac{1}{\Delta}\Big(\underbrace{\int_{c_0}^{c_0+\Delta} \frac{\partial R^*(c)}{\partial c_0}\, dc}_{\geq 0} + \underbrace{R^*(c_0+\Delta) - R^*(c_0)}_{<0}\Big). \tag{3.7}$$

It is easy to see that if $\frac{\partial^2 q(R,m)}{\partial m \partial R}$ is sufficiently small for all $R \in [0, R_{\max}]$ and $m \in ]0,1[$, the term $\frac{\partial R^*(c)}{\partial c_0}$ becomes arbitrarily small. Therefore the difference $R^*(c_0+\Delta) - R^*(c_0)$ (which is not dependent on the cross derivative $\frac{\partial^2 q}{\partial m \partial R}$) outweighs the integral in equation (3.7) and we get

$$\frac{\partial \bar{R}^*}{\partial c_0} < 0.$$

In order to compute $\frac{\partial m^*}{\partial c_0}$ we consider equation (3.6). Again, by application of the Implicit Function Theorem and Leibnitz's Rule, we get

$$\frac{\partial m^*}{\partial c_0} = -\frac{\frac{1}{\Delta}\Big(\overbrace{\int_{c_0}^{c_0+\Delta} \frac{\partial q}{\partial R}\frac{\partial R^*(c)}{\partial c_0}\, dc}^{<0} + \overbrace{\big(q(R^*(c_0+\Delta), m^*) - q(R^*(c_0), m^*)\big)}^{>0}\Big)}{\underbrace{1 + \frac{1}{\Delta}\int_{c_0}^{c_0+\Delta} \frac{\partial q}{\partial m}\, dc}_{>0}}. \tag{3.8}$$

The denominator is positive because of the assumption that $|\frac{\partial q}{\partial m}|$ is sufficiently small. To determine the sign of the numerator, consider the term

$$|\int_{c_0}^{c_0+\Delta} \frac{\partial q}{\partial R}\frac{\partial R^*(c)}{\partial c_0}\, dc|.$$

Plugging in the formula for $\frac{\partial R^*(c)}{\partial c_0}$, we get

$$\begin{aligned}
&\Big|\int_{c_0}^{c_0+\Delta} \frac{\partial q}{\partial R} \cdot \frac{\partial R^*(c)}{\partial c_0}\, dc\Big| \\
&= \Big|\int_{c_0}^{c_0+\Delta} \frac{\partial q}{\partial R}\Big( - \frac{\frac{\partial^2 q}{\partial R \partial m}\big(-\frac{1}{\Delta}\big)\big(q(R^*(c_0+\Delta), m^*) - q(R^*(c_0), m^*)\big)}{\frac{\partial^2 q}{(\partial R)^2} + \frac{\partial^2 q}{\partial R \partial m}\big(-\frac{1}{\Delta}\int_{c_0}^{c_0+\Delta}\frac{\partial q}{\partial R}\, dc\big)}\Big)\, dc\Big| \\
&\leq \int_{c_0}^{c_0+\Delta} \Big|\frac{\partial q}{\partial R} \cdot \Big(\frac{\frac{\partial^2 q}{\partial R \partial m}\big(\frac{1}{\Delta}\big)\big(q(R^*(c_0+\Delta), m^*) - q(R^*(c_0), m^*)\big)}{\frac{\partial^2 q}{(\partial R)^2} - \frac{\partial^2 q}{\partial R \partial m}\big(\frac{1}{\Delta}\int_{c_0}^{c_0+\Delta}\frac{\partial q}{\partial R}\, dc\big)}\Big)\Big|\, dc \\
&< \int_{c_0}^{c_0+\Delta} \Big|\frac{\partial q}{\partial R} \cdot \Big(\frac{\frac{\partial^2 q}{\partial R \partial m}\big(\frac{1}{\Delta}\big)\big(q(R^*(c_0+\Delta), m^*) - q(R^*(c_0), m^*)\big)}{\frac{\partial^2 q}{\partial R \partial m}\big(\frac{1}{\Delta}\int_{c_0}^{c_0+\Delta}\frac{\partial q}{\partial R}\, dc\big)}\Big)\Big|\, dc \\
&= |q(R^*(c_0+\Delta), m^*) - q(R^*(c_0), m^*)|.
\end{aligned}$$

Hence we see that the positive difference $|q(R^*(c_0+\Delta), m^*) - q(R^*(c_0), m^*)|$ outweighs the negative integral term in the numerator of equation (3.8). Altogether we get

$$\frac{\partial m^*}{\partial c_0} < 0.$$

q.e.d.

# Chapter 4

# A Generalization of the Ortalo-Magné-Rady Model on Credit Constraints and House Price Dynamics

## 4.1 Introduction

Empirical studies[1] show that the market for privately owned houses is relatively volatile. It reacts sensitively to exogenous income shocks, especially to income shocks of the youngest cohort in the market. This is puzzling because of three reasons: First, housing is an extremely durable good with a low depreciation rate. Therefore one could expect the housing price to be relatively stable as well. Secondly, the planning horizon of house buyers should be longer as compared to a standard consumption good. Exogenous income shocks, even if permanent, should have a minor impact in comparison to the steady rise in average income when growing older. Third, due to financial constraints and mobility considerations, young buyers of real estate typically buy relatively small dwellings. Thereby, the price for expensive real estate should only be influenced by

[1] See for example Sutton (2003) or Andrew and Meen (2003).

substitution effects which are considered to be relatively low.

Ortalo-Magné and Rady (2006)[2] offer a rationale that is able to explain the effects outlined above. In their model, they account for the fact that many house buyers are credit constrained, and many young and middle-aged households would like to own a more expensive property but cannot afford the required down payment to get the credit from a bank. For concreteness, let us consider a positive income shock to the whole economy. Then, the reasoning is the following: the income of the young influences the price of starter homes and thereby the wealth of households that already own a starter home. These households - typically between the young and the old cohort - profit in two ways from the shock. First, their income has risen, and secondly, they profit from the investment feature of real estate: their total wealth increases tremendously by the fact that the value of the main part of their wealth - the investment in the property - has risen. Thus, their demand for houses is strongly increased which leads to a strongly increased price. The price in the period directly after the shock is even higher than in the new steady state equilibrium because the main driver of the effect is the large price-induced change in wealth of those households that already own a smaller property: in all subsequent periods, these households do not face such a radical change in their wealth because the property they own was bought when prices were already high.

Furthermore, there is empirical evidence that transaction volume in the market is positively correlated with prices. This is puzzling because standard theories of mean reverting asset values predict exactly the opposite: If the price is relatively high, fewer agents will buy, which in turn leads to a decrease in prices again. Moreover, if we assume that the market was already in equilibrium, in a simple static demand-supply model, rising prices would lead to fewer potential buyers whose willingness to pay is at least as high as the price. Therefore, transaction volume should decrease.

[2] Although we will briefly present the assumptions, intuitions and the main results of the model of Ortalo-Magné and Rady (2006), for readers who are not familiar with the special features of their dynamic overlapping-generations setting, it may be convenient to consult their paper during or before the reading of this chapter.

This positive correlation of transaction volume and prices can be explained by the model of Ortalo-Magné and Rady (2006) as well: Again, suppose for concreteness that prices rise. Then, the first effect is that more and more households (especially those who already own a property) become unconstrained in their property decision. Many agents now can afford a (more expensive) dwelling. This leads to the price overshooting pattern described above and to a strongly increased transaction volume for expensive dwellings. Simultaneously a second effect comes into play: Some old households that are not constrained in their decision anyway, trade down from an expensive dwelling to a cheaper one, compared to the situation where prices do not rise. The decisions of these old households are assumed to be driven entirely by preference reasons such that the marginal old house buyer is indifferent between the housing utility of a flat minus its user cost and the housing utility of an expensive dwelling minus its user cost. Since the first effect is stronger than the second one, transaction volume for all kinds of dwellings rises together with prices.

However, Ortalo-Magné and Rady (2006) build their model on quite unrealistic assumptions. In order to keep the model analytically tractable, they abstract from any consumption smoothing behaviour by imposing linear consumption utility. This seems to contradict reality in two ways: First, empirical studies show that agents act in highly risk averse ways, especially when income is low. Secondly, the linear utility specification, together with the imposed positive interest rate, implies that in all periods of life, except for the last one, numeraire consumption of all households, even the richest, equals zero. This feature keeps the model mathematically solvable, but makes it impossible to embed the model in a richer structure[3], where for example goods and labour markets are integrated.

Moreover, for analytical convenience, and in order to have the strongest price effects when reselling a dwelling, the model assumes transaction costs of zero.

In this chapter, we generalize the model outlined above in two ways: First, we introduce transaction costs, and secondly, we soften the restrictive assumption of linear

[3]A simple structure of this type could be a computational general equilibrium model, where not only the market for housing is in equilibrium, but the goods market as well.

utility for non-housing consumption. By introducing transaction costs and a logarithmic utility function, we study the robustness of the model in the presence of more realistic assumptions. In the generalized version, the model is no longer fully analytically tractable. Therefore, numerical simulations are used to derive the results.

We begin in the theoretical part by describing the generalized steady state equilibrium and provide conditions which give rise to so-called full consumption behaviour, which means that all remaining wealth is immediately consumed after a potential housing transaction. Furthermore, we explain why the generalized setting leads to property prices such that for age 1 and age 2 agents the qualitative distribution of dwellings is not altered. However, in contrast to the benchmark model of linear consumption utility, the decision of age 3 households does not only depend on their preferences, but as well on their wealth. Moreover, we give some approximations of property prices.

In the numerical part we study the evolution of consumption over the life cycle in the extreme case of full consumption.

Furthermore, we study the overshooting pattern of prices. It is shown that the logarithmic specification allows for the same qualitative overshooting pattern, but the effects are quantitatively smaller. We identify two effects governing the change in prices. First, due to consumption smoothing, the relative willingness of households to buy real estate is increased after income shocks. This effect is present in any period after the shock. The second effect comes only into play directly after the shock, when capital gains occur. Compared to the linear specification, not all house purchases possible are actually accomplished because a certain fraction of the flat owners prefer to increase their consumption instead.

We find that the introduction of transaction costs does not have any allocative consequences in steady state. Interestingly, after the introduction of transaction costs, the flat price changes proportionally to transaction costs, whereas the house price is overproportionally reduced. This is due to the fact that repeat purchases become less attractive.

The chapter is structured as follows: In the next section, we give a short overview of the literature concerning the effects treated in this work. Building on the model in

Ortalo-Magné and Rady (2006), Section 3 describes the generalized setting, while in Section 4, the most important results of the benchmark case of linear utility and no transaction costs are presented. In the subsequent part, we derive some formal results about the steady state in the generalized case. In Section 6, the economy is simulated for both linear and logarithmic consumption utility and the consequences of exogenous income shocks are studied. Moreover, we present the effects of transaction costs in the model. The last section summarizes the findings and concludes.

## 4.2 Related Literature

In the following, we constrain ourselves to literature which we consider to be either most important[4] or very closely related to our work. Some important contributions not mentioned here can be found in Ortalo-Magné and Rady (2006).

### 4.2.1 High Volatility and Sensitivity to Income of the Young

In his seminal paper analysing the role of credit constraints on the dynamics of house prices, Stein (1995) builds a stylized theoretical framework in which he is able to partly explain the high volatility in the real estate market. The crucial point of this model are the self-reinforcing effects from shocks to house prices. As in the model of Ortalo-Magné and Rady (2006), a multiplier effect occurs when a household willing to buy a new property has incurred losses or gains by selling its old property. Secondly, these self-reinforcing effects give rise to multiple equilibria. Both the existence of multipliers and multiple equilibria serve as an explanation for the observed high volatility in the housing market. Ortalo-Magné and Rady (1999) show that changes in the ability to get a credit and the income of the youngest cohort in the market explain most of the variance of house prices. Andrew and Meen (2003) attribute the decline in house prices in Britain at the beginning of the 1990s to the income of the young first time buyers.

[4]Cho (1996) provides a good general survey about several topics in the literature on house price dynamics.

### 4.2.2 Correlation between Prices and Transaction Volume

Stein's model (1995) also serves as an explanation of the positive correlation of prices and transaction volume. For instance, when house prices rise, some households are no longer credit constrained and attempt to move, inducing a surge in transaction volume. Genesove and Mayer (2001) apply prospect theory to explain this effect in the case of falling prices. If all decisions are made in relation to a certain reference point (like the price of the house at which one bought it) and if a nominal loss aversion is presumed (which easily can be modelled by a concave value function for gains in the trading process, a convex function for losses and a kink at the reference point) households tend to avoid losses when reselling their dwelling and real estate becomes an illiquid asset. Ortalo-Magné and Rady (2004) build on their basic theoretical framework which focuses on vertical transactions along the property ladder. They add empirical evidence for the positive correlation of prices and volume for Wales and England. Krainer, Spiegel and Yamori (2005) develop an overlapping generations model with search frictions and heterogeneity of agents' valuations for houses. Moreover, they provide empirical evidence using data from Japan to show that after negative income shocks, prices are sticky and houses become an illiquid asset which leads to a lower transaction volume. In their model, search theory and financial constraints are combined in order to rationalize the phenomenon of loss aversion. There, the obligation to pay back the debt associated with a previous house purchase is the main driver of the distorted trading decision in the real estate market.

### 4.2.3 Overshooting of Prices after Exogenous Shocks

Lamont and Stein (1999) and Malpezzi (1999) examine empirically the dynamics of the real estate market in the United States after income shocks and show that house prices tend to overshoot. Capozza et al. (2002) explore the dynamics of real house prices by measuring serial correlation and mean reversion coefficients for metropolitan areas. They find significant overshooting patterns, especially for regions with high real construction costs and low mean reversion. Aitken, Grimes and Kerr (2004) empirically test the efficiency of the housing market by estimating some factors that determine the dynamic path of house prices. They find that house price dynamics are influenced by past transaction volumes. However, they attribute the observed overshooting pattern mainly to extrapolative price expectations when new information about the market is gained.

## 4.3 The Generalized Model

In this section, we present the model in generalized form where the utility function for consumption is not necessarily linear and buyers of real estate have to pay transaction costs. The decision problems of agents become one dimension richer, since they have to control for optimal consumption plans and optimal housing plans simultaneously. In order to compare with the model developed in Ortalo-Magné and Rady (2006), most of the remaining economic environment is left unchanged.

### 4.3.1 Economic Setup

**Households and Income.** We consider an infinite overlapping-generations model with four generations. Households are heterogenous with respect to their income $i$ and their preferences $m$. Both $i$ and $m$ are distributed uniformly on $[0, 1]$. The preference index $m$ is not known to the household until the beginning of age 3. A household of age $j \in \{1, 2, 3, 4\}$ and index $i$ receives income $e_j(i)$, where $e_j$ is assumed to be strictly increasing and continuous in the income index. All households are perfectly rational and able to anticipate the future (except for the unforeseen exogenous shock).

**Consumption and Housing.** There are only two goods in the economy, namely the numeraire good and real estate. An agent can either own no dwelling ($\varnothing$), a flat (F) or a house (H). In each period, the consumption of the numeraire good must be non-negative.

**Utility Function.** Households' preferences are described by the following additively separable utility function:

$$\mathcal{U}(c_1, c_2, c_3, c_4, h_2, h_3, h_4) = U(h_2, \frac{1}{2}) + U(h_3, \frac{1}{2}) + U(h_4, m) + \sum_{i=0}^{4} u_c(c_i),$$

where $U(h, m)$ represents the housing utility of agents with preference index $m$. In the first two periods $m$ is set to $\frac{1}{2}$. At age 3, agents learn their preference index. Numeraire consumption in period $j$ is denoted by $c_j$. $u_c$ is the general utility function for consumption which we call c-utility function. The only regularity assumptions we impose on this c-utility function are concavity, continuity and strict positive monotonicity. We allow for specifications of $u_c$ where $\lim_{x \to 0} u(x) = -\infty$, such as $u_c(x) = ln(x)$. In this case, we set $u_c(0) := -\infty$.

The type of housing at age $j$ is denoted by $h_j$. Housing utility is defined as follows:

$$U(h, m) = \begin{cases} -\Delta & \text{if } h = \varnothing \\ 0 & \text{if } h = F \\ u(m) & \text{if } h = H \end{cases}$$

where $\Delta$ is strictly positive and $u$ is strictly increasing and continuous in $m$.

**Market, Supply and Demand.** The market for real estate is assumed to be competitive. Per definition, the price of "no dwelling" is set to zero. The market price of a property is denoted by $p_h,\ h \in \{F, H\}$. We assume that the buyer of the new property has to pay the total price $p_h^T = (1+T)p_h,\ h \in \{F, H\}$, where $0 \leq T$ represents the rate of transaction costs[5], whereas the seller receives the price $p_h$. Supply is fixed. There exists a mass of $S_F$ flats and a mass of $S_H$ houses. Demand is generated for flats and houses separately by aggregating the demands of the first three age cohorts.

**Small Open Economy with Credit Constraints.** Interest rates for saving and for borrowing are equal, denoted by $r \geq 0$.[6] In order to get a credit for the purchase

[5] Trivially, transaction costs must be paid only if the household downgrades or upgrades its property, but not when staying in the old property.

[6] In the model with linear c-utility, a strictly positive interest rate is assumed. Due to this assumption, all consumption is delayed to the last period of life. However, in order to prove some formal

of a property, households need to pay at least the fraction $0 < \gamma < 1$ of the total price from their own savings. It is assumed that in any period when holding the property, the end-of-period total wealth (housing and non-housing wealth) must be equal to or higher than $\gamma$ times the total price of the property.[7]

**Timing.** The timing within each period is as follows:

1. Households derive utility from housing.
2. Households obtain their period income $e_j(i)$.
3. Households trade on the real estate market.
4. The numeraire good is consumed.

This implies that a buyer of a property in period $t$ has to pay the prevailing total price in period $t$, but derives utility from living in this property in period $t+1$.

Households of different ages differ in the information they can use to derive their optimal decision. All personal information is gathered in the so-called individual state. The following table shows the beginning-of-period information available to households, dependent on their age:

| Age | Information about the Personal State |
|---|---|
| 1 | income index $i$ |
| 2 | income index $i$, non-housing wealth $w$, property $h$ |
| 3 | income index $i$, non-housing wealth $w$, property $h$, preference index $m$ |
| 4 | income index $i$, non-housing wealth $w$, property $h$, preference index $m$ |

results in the generalized case, it is convenient to allow for $r = 0$.

[7]This setting is more restrictive than in reality, where it would be sufficient that total wealth in the period of buying the property is equal to or higher than $\gamma$ times the price of the dwelling. Nonetheless, in order to integrate the decision problem of a household in a dynamic optimization setting, it is convenient to impose the down payment condition even in periods after the purchase, since this allows a simpler specification of the feasibility constraint.

In steady state, age 1 households anticipate their non-housing wealth $w$ and their property $h$ at ages 2 and 3.

### 4.3.2 Parameter Assumptions

We use the notation $e(i) := (1+r)e_1(i) + e_2(i)$. As we will see later for age cohorts 1 and 2, the basic parameter assumptions are sufficient to generate the same qualitative housing allocations as in the linear model, whereas the housing allocation for the age 3 cohort differs notably from the linear case. In the following, we present the generalized parameter assumptions[8] and describe their consequences in the case of constant prices.

$$2\frac{1}{2} < S_F + S_H < 3 \text{ and } \frac{1}{2} < S_H < 1 \tag{4.1}$$

$$\gamma > \frac{T}{(1+r)(1+T)} + \frac{r}{1+r} \tag{4.2}$$

$$e_1(0) = 0 \tag{4.3}$$

$$e_2(0) > e_1(3 - S_F - S_H) \tag{4.4}$$

$$e_2(i) > e_1(i) \text{ for all } i \in [0,1] \tag{4.5}$$

$$e_3(0) > e(1) \tag{4.6}$$

$$e_2(1) + \big(\gamma(1+r)(1+T) - T - r(1+T)\big)p_F^* > \max\{e_1(1), e(3 - S_F - S_H)\} \tag{4.7}$$

$$u(1 - S_H) < 0 \tag{4.8}$$

$$\Delta > u(\frac{1}{2}) > 0 \tag{4.9}$$

Roughly speaking, these assumptions state a positive relationship of income and age to the affordability of real estate. The older the household grows and the more income it gets, the less binding becomes the credit constraint:

1. Per definition, age 1 households are born without any property.
2. Assumptions (4.2), (4.5) and (4.6) imply that if a household bought a property in period $t$, the household can afford that property in period $t+1$ if the price

[8] In the generalized model, $p_F^*$ cannot be expressed in a closed form of exogenous variables. In order to avoid this problem, assumption (4.7) could be replaced by the (stronger) condition $e_2(1) > \max\{e_1(1), e(3 - S_F - S_H)\}$.

remains the same. Furthermore, the household has no incentive to downgrade at age 2.

3. Assumption (4.3) ensures that a positive mass of age 1 households is not able to buy any property.

4. Assumptions (4.4) and (4.9) guarantee that every age 2 household is able to buy at least a flat and has an incentive to do so.

5. According to assumptions (4.7), (4.8) and (4.9) a positive mass of age 2 households is able to buy a house and has an incentive to do so.

6. Assumptions (4.6) and (4.9) ensure that every age 3 household is able to buy any dwelling and has an incentive to buy at least a flat.

7. Condition (4.8) ensures that dependent on their preferences, some age 4 households will own a flat, although they could afford a house. Together with $u(\frac{1}{2}) > 0$ and assumption (4.6), this condition guarantees that at least a mass of $\frac{1}{2}$ of all houses is purchased by age 3 households.

8. Assumption (4.9) ensures that a house is more desirable than a flat which is more desirable than no property without taking into account the holding costs of a property.

Depending on the c-utility function, some households might choose not to purchase a property, even if they could afford it because the sum of c-utility in the period of buying and the resulting value function in the subsequent period (including housing wealth) is larger without the property than in the case of the purchase of the property. In this case, the marginal buyer of a property is indifferent between buying the dwelling and staying in the old dwelling. For concreteness, consider the case where housing prices are constant and marginal age 1 flat buyers always consume all remaining wealth after the (potential) housing transaction.[9] When buying a (bigger) property, the household

[9]Due to conditions which we will derive later, marginal age 1 flat buyers consume any remaining non-housing wealth after the (potential) real estate purchase. However, note that this condition is

consumes less in the period of the purchase, but will consume *more* in the periods after the purchase. This is due to the fact that the household transfers more total wealth into the future by the property transaction than in the case of no purchase. To make clear the tradeoff, the following table shows the countervailing effects for a marginal age 1 flat buyer.

| Marginal Age 1 Flat Buyer ($i_F^* = 3 - S_F - S_H$) | buy a flat | do not buy a flat |
|---|---|---|
| consumption in period 1 | low | high |
| total wealth transfer in next period | relatively high | relatively low |
| housing at age 2 and age 3 | $F, F$ | $\varnothing, F$ |

Concerning the total wealth transfer, consider the case of a flat purchase. The total wealth[10] at age 2 is given by the following expression:

$$e_2(i) + \underbrace{(1+r) \cdot [e_1(i) - c_1(i) - p_F^T] + p_F}_{\text{total wealth transfer}},$$

where $c_1(i)$ is the consumption at age 1. In the case of minimal[11] wealth transfer into period 2 all remaining wealth is consumed after the flat purchase at age 1 and we get

$$c_1(i) = e_1(i) - \gamma p_F^T.$$

Plugging this expression for consumption in the term above, we get the following expression for the total wealth transfer:

$$\gamma(1+r)(1+T)p_F - Tp_F - r(1+T)p_F.$$

In order to understand the economics behind this term rearrange it:

$$p_F - (1+T)(1-\gamma)p_F - r(1+T)(1-\gamma)p_F$$

general enough to allow for parameter combinations which lead to an equilibrium such that this must not necessarily hold for high $i$. This is due to the incentive to smooth consumption over the life cycle for age 1 households that will be house buyers at age 2. The consequences of this effect shall be captured by our setting.

[10]Total wealth includes housing wealth at actual net prices and (possibly negative) non-housing wealth after getting the per-period income and before a potential trade on the property market.

[11]This is the case that we call full consumption.

The first summand, $p_F$, represents the actual net housing wealth in period 2. However, total wealth also includes the loan of the agent which is given by the second summand, $-(1+T)(1-\gamma)p_F$. Furthermore, the interest payments from period 1 to period 2 are $-r(1+T)(1-\gamma)p_F$.

Let us define the factor of minimal wealth transfer as

$$F_{(T,r)} := \gamma(1+r)(1+T) - T - r(1+T).$$

Simple calculations show that $F_{(T,r)} > 0$ due to assumption (4.2). Hence the total wealth transfer into period 2 is positive.

In case of no property purchase, the corresponding total wealth transfer is zero, since, per assumption, all wealth at age 1 is consumed immediately.

### 4.3.3 Equilibrium Definition

The definition of a recursive equilibrium in the general case only slightly differs from the linear case. Nevertheless, in the generalized case, the decision rules and the appendant value functions now include terms for consumption.

The beginning-of-period state of the economy is given by a collection
$x = (x_2^w, x_3^w, x_4^w, x_2^h, x_3^h, x_4^h, )$ of measurable functions

$$x_2^w, x_3^w : [0,1] \to \mathbb{R}, \quad x_4^w : [0,1] \times [0,1] \to \mathbb{R}$$
$$x_2^h, x_3^h : [0,1] \to \mathcal{H}, \quad x_4^h : [0,1] \times [0,1] \to \mathcal{H},$$

where $x_2^w(i)$, $x_3^w(i)$ specify the non-housing wealth of age 2 and age 3 agents. $x_4^w(i,m)$ is the corresponding age 4 non-housing wealth after learning the preference index. Analogously, $x_2^h(i), x_3^h(i)$ and $x_4^h(i,m)$ specify the property the agent owns.[12] Households' decisions depend only on their state and on current and expected future prices. In the generalized case, it is also possible that agents save money from one period to the subsequent one in order to smooth numeraire consumption: Under certain conditions

[12]States that differ only on a null set of agents are considered as identical.

agents might reduce their current consumption in favor of future consumption when they anticipate scarce non-housing wealth in future periods of expensive property purchases.

The feasibility set for housing and consumption at age $j$ is given by

$$\Gamma_j(i,w,h,x) = \{(h',c) \in \mathcal{H} \times [0,\infty] \mid e_j(i) + w + p_h(x) \geq \gamma(1+T(h,h'))p_{h'} + c\}$$

where the transaction costs are defined as follows:

$$T(h,h') = \begin{cases} 0 & \text{if } h = h' \\ T & \text{otherwise} \end{cases}$$

The corresponding beginning-of-period non-housing wealth at age $j = 2,3,4$ is given by

$$W_{j-1}(i,w,h,h',c,x) = (1+r)[e_{j-1}(i) + w + p_h(x) - (1+T(h,h'))p_{h'}(x) - c]$$

Note that the non-housing wealth does not involve the down payment constant $\gamma$ since the resale value of a property is independent of the down payment made one period before.

**Definition 4.3.1** *A recursive competitive equilibrium consists of*

1. *the set $X$ of all possible states $x$ of the economy,*
2. *decision functions $h_1(i,x), h_2(i,w,h,x)$ and $h_3(i,m,w,h,x)$ for housing,*
3. *decision functions $c_1(i,x), c_2(i,w,h,x)$ and $c_3(i,m,w,h,x)$ for consumption,*
4. *value functions $v_1(i,x), v_2(i,w,h,x)$, $v_3(i,m,w,h,x)$ and $v_4(i,m,w,h,x)$,*
5. *property price functions $p_F(x)$ and $p_H(x)$, and*
6. *a law of motion $\Phi : X \to X$*

*such that the following conditions hold:*

(a) *Under the property price functions $p_F$ and $p_H$ and the law of motion $\Phi$, the households' decision functions solve their maximization problems and generate the corresponding value functions for all $x \in X$.*

(a.1) *$h' = h_1(i,x)$ and $c = c_1(i,x)$ solve*

$$v_1(i,x) = \max_{(h',c)\in\Gamma_1(i,0,\varnothing,x)} \left\{u_c(c) + v_2(i, W_1(i,0,\varnothing,h',c,x), h', \Phi(x))\right\};$$

(a.2) *$h' = h_2(i,w,h,x)$ and $c = c_2(i,w,h,x)$ solve*

$$\begin{aligned} v_2(i,w,h,x) = \quad & U(h,\frac{1}{2}) \\ & + \max_{(h',c)\in\Gamma_2(i,w,h,x)} \int_0^1 \left[u_c(c) + v_3(i, W_2(i,w,h,h',c,x), h', \Phi(x))\right] dm; \end{aligned}$$

(a.3) *$h' = h_3(i,m,w,h,x)$ and $c = c_3(i,m,w,h,x)$ solve*

$$\begin{aligned} v_3(i,m,w,h,x) = \quad & U(h,\frac{1}{2}) \\ & + \max_{(h',c)\in\Gamma_3(i,w,h,x)} \left[u_c(c) + v_4(i,m,W_3(i,w,h,h',c,x), h', \Phi(x))\right]; \end{aligned}$$

(a.4) *$v_4(i,m,w,h,x) = U(h,m) + u_c(e_4(i) + w + p_h(x))$ if $e_4(i) + w + p_h(x) \geq 0$ and $-\infty$ otherwise. If $v_4(i,m,w,h,x) > -\infty$, $c_4(i,m,w,h,x) = e_4(i) + w + p_h(x)$.*

(b) *Housing markets clear. That is, the measure of all agents who acquire a flat equals $S_F$ and the measure of all agents who acquire a house equals $S_H$.*

(c) *Agent's decision rules are consistent with the law of motion.*[13]

## 4.4 The Benchmark Case: Linear c-Utility and No Transaction Costs

For completeness, and to have a benchmark case, we present the results derived in Ortalo-Magné and Rady (2006) in strongly abbreviated form. For a more elaborate description, further details and all appendant proofs, see the original paper.

[13] The formalizations of (b) and (c) are trivial, but tedious. See Ortalo-Magné and Rady (2006) for their analogon in the case of linear c-utility.

Due to the linear c-utility function and the fact that future consumption utility is not discounted, the positive interest rate leads to zero consumption in the first three periods of life. All remaining wealth (housing and non-housing wealth) is consumed in period four.

In the linear case, condition (4.9) is replaced by the following two assumptions:

$$(1+r)^2 r\gamma^{-1}[e(1) - e_1(3 - S_F - S_h)] < u(\frac{1}{2}) \tag{4.10}$$

$$u(\frac{1}{2}) + (1+r)^2 r\gamma^{-1} e_1(3 - S_F - S_h) < \Delta \tag{4.11}$$

These assumptions are stronger than (4.9).[14] They ensure that per-period housing utility is always larger than the appendant maximal consumption loss. Therefore, whensoever a household is able to buy a property, the property will be purchased. Furthermore, assumption (4.10) guarantees that if both properties are affordable, the trade-off concerning utility and consumption loss is always better for a house than for a flat. Together with the competitive market assumption, this implies that only the wealthiest agents in cohorts one and two can afford a house. Since we set transaction costs to zero in the benchmark case, assumption (4.2) reduces to

$$\gamma > \frac{r}{1+r}.$$

In the case of zero transaction costs, condition (4.7) reduces to

$$e_2(1) + (\gamma(1+r) - r)p_F^* > \max\{e_1(1), e(3 - S_F - S_H)\}$$

As we will see later, the price of flats equals $\gamma^{-1}e_1(3 - S_F - S_H)$. Thus we get

$$e_2(1) + (\gamma(1+r) - r)\gamma^{-1}e_1(3 - S_F - S_H) > \max\{e_1(1), e(3 - S_F - S_H)\}.$$

This is equivalent to

$$e(1) > \max\{e_1(1), e(3 - S_F - S_H)\} + r\gamma^{-1}e_1(3 - S_F - S_H).$$

[14] Assumption (4.11) facilitates the proofs in the model with linear c-utility, since it ensures that the steady state price $p_F^*$ of a flat equals $\gamma^{-1}e_1(3 - S_F - S_H)$. If only assumption (4.9) were imposed, the market mechanism would only lead to the inequality $0 < p_F^* \leq \gamma^{-1}e_1(3 - S_F - S_H)$. However, the qualitative housing allocation for age 1 and age 2 cohorts would stay the same.

### 4.4.1 Steady State

In this section, the resulting steady state equilibrium is described. Variables which denote the steady state values are denoted by an *.

**Proposition 4.4.1** *There is a unique steady state equilibrium. The steady state allocation of properties at the beginning of each period is determined by the critical endowment indices*

$$\begin{aligned} i_F^* &= 3 - S_F - S_H \\ i_H^* &= e_1^{-1}(\gamma p_H^*) \\ i_{\varnothing H}^* &= e^{-1}(\gamma p_H^*) \\ i_{FH}^* &= e^{-1}(r p_F^* + \gamma p_H^*) \end{aligned}$$

*and the critical preference index*

$$m_H^* = u^{-1}(r[p_H^* - p_F^*]),$$

*where the prices $p_F^*$ and $p_H^*$ of flats and houses are determined by the following two equations:*

$$\begin{aligned} p_F^* &= \gamma^{-1} e_1(3 - S_F - S_H) \\ 3 - S_H &= e_1^{-1}(\gamma p_H^*) + \min\{e_1^{-1}(\gamma p_F^*), e^{-1}(\gamma p_H^*)\} \\ &\quad + e^{-1}(r p_F^* + \gamma p_H^*) - e_1^{-1}(\gamma p_F^*) + u^{-1}(r[p_H^* - p_F^*]) \end{aligned}$$

*The indices satisfy the following conditions:*

$$0 < i_F^* < \frac{1}{2} < i_{FH}^* < i_H^* \leq 1, \quad 0 < i_{\varnothing H}^* < i_{FH}^* \text{ and } 0 < m_H^* < \frac{1}{2}.$$

- *Age 2 agents with endowment index $i < i_F^*$ hold no property, those with $i_F^* < i < i_H^*$ hold a flat; and those with $i > i_H^*$ hold a house.*
- *Age 3 agents with endowment index $i < \min\{i_F^*, i_{\varnothing H}^*\}$ or $i_F^* < i < i_{FH}^*$ hold a flat; and those with $\min\{i_F^*, i_{\varnothing H}^*\} < i < i_F^*$ or $i > i_{FH}^*$ hold a house.*

- *Age 4 households with preference index $m < m_H^*$ hold a flat, those with $m > m_H^*$ hold a house.*

*The per period transaction volume of flats $\eta_F^*$ and of houses $\eta_H^*$ are*

$$\eta_F^* = i_H^* - i_F^* + \min\{i_{\varnothing H}^*, i_F^*\} + m_H^*(i_F^* - \min\{i_{\varnothing H}^*, i_F^*\} + 1 - i_{FH}^*)$$
$$\eta_H^* = 1 + i_F^* - \min\{i_{\varnothing H}^*, i_F^*\} - i_{FH}^* + (1 - m_H^*)(\min\{i_{\varnothing H}^*, i_F^*\} + i_{FH}^* - i_F^*)$$

The following figure from Ortalo-Magné and Rady (2006) shows the housing allocation graphically.

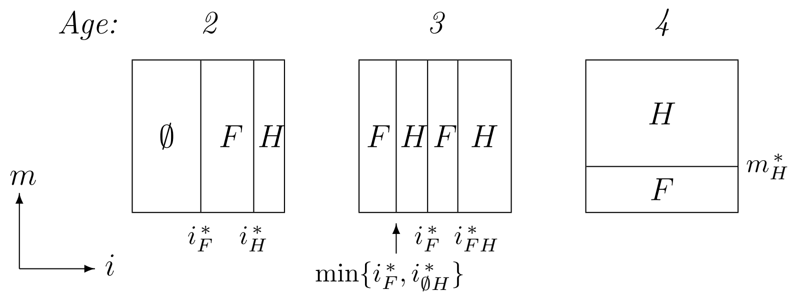


Figure 4.1: Steady-State Equilibrium Allocation of Properties

### 4.4.2 Exogenous Income Shocks

An exogenous unforeseen income shock is modelled by the multiplication of the factor $z > 0$, $z \neq 1$ to the income functions for all ages: The new income functions are set to $ze_j(i)$. If $z > 1$, the shock is positive, otherwise negative. If $|z-1|$ is sufficiently small, the parameter assumptions continue to hold and Proposition 4.4.1 applies again: The new steady state equilibrium shares all qualitative properties of the equilibrium without shock. Let us denote the steady state after the shock with a double superscript **. From Proposition 4.4.1, we know that in steady state we have $p_F^* = \gamma^{-1}e_1(3 - S_F - S_H)$, therefore in the new steady state we get:

$$p_F^{**} = zp_F^*$$

The price of houses, $p_H$, changes less than proportionally with endowments, but more, in absolute terms, than the price of flats:

$$
\begin{aligned}
p_H^* + (z-1)p_F^* < p_H^{**} < zp_H^* \quad &\text{if} \quad z > 1 \\
zp_H^* < p_H^{**} < p_H^* + (z-1)p_F^* \quad &\text{if} \quad z < 1
\end{aligned}
$$

Concerning the transition to the new steady state, it can be shown that if the shock is sufficiently small, that is, $z$ is close to 1, the transition to the next steady state is finished within five periods for the endowment profiles, within two periods for the housing allocation and the house price, and immediately for flat prices. As we will explain later, the length of the transition is independent of the c-utility specification. The allocation in the period right after the shock shall be denoted by a superscript +.

In the following, let us restrict ourselves to the special case where all house buyers owned a property (flat or house) before the purchase of the house. This can be achieved by the following assumption:

$$
e(\frac{3}{2} - S_H) \geq \max\{e_1(1), e(3 - S_F - S_H)\} + r\gamma^{-1}e_1(3 - S_F - S_H)
$$

Then we have $\min\{i_F^{**}, i_{\emptyset H}^{**}\} = i_F^{**}$ and $i_H^{**} = 1$. In this situation, a change in the price of flats has the strongest effect possible on the demand for houses, since all age 2 house buyers face a large change in wealth.

### Overshooting and Correlation of Prices and Transaction Volume

**Proposition 4.4.2** *If $ze_1(i_{FH}^{**}) < p_F^{**}$, in the first period after the shock, the house price $p_H^+$ overshoots its new steady state level:*

$$
\begin{aligned}
p_H^+ > p_H^{**} \quad &\text{if} \quad z > 1 \\
p_H^+ < p_H^{**} \quad &\text{if} \quad z < 1
\end{aligned}
$$

*Furthermore, if house prices overshoot, prices of flats and houses move with transaction volume.*

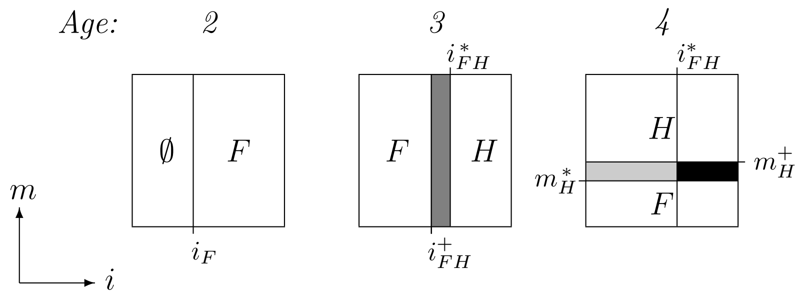


Figure 4.2: Allocations of Dwellings and Transactions

Figure 4.2 is from Ortalo-Magné and Rady (2006) and shows the housing allocation right after the shock compared to the old steady state. Proposition 4.4.2 represents the core part of Ortalo-Magné and Rady (2006). Using Figure 4.2, we can make clear the mechanism at work: Let us consider a positive shock to income. The credit constraint of age 2 house buyers is loosened in two ways: First, their income increased and secondly, they have become wealthier by a rise in the price of the flat they own. More of them contribute to the demand for houses. The grey bar in the housing allocation for age 2 households shows their increased demand for houses. Age 3 households are unconstrained in their decision anyway. Their willingness to pay for a house is the same in the first period after the shock and in the new steady state since their willingness to pay depends (besides their preference index $m$) only on the current flat price and on future property prices. However, the flat price adjusts immediately after the shock and the future house price is the new steady state price in both instances. Altogether, this implies that the age 3 demand function for houses is the same right after the shock as in the new steady state. Since, per assumption, all potential age 2 house buyers will purchase a house if they can afford it, the total demand for houses rises strongly, driven only by the change in demand of age 2 house buyers. This leads to a higher house price in the period right after the shock. It is even higher than in the new steady state because the loosening effect on the credit constraint for age 2 house buyers is weaker if they bought their flat at the new steady state price. At the same time, for age 3 households, houses become more unattractive because of the higher price. More of them trade down to a flat. The sum of the black and the grey bar in the housing allocation box for age 3 households represents the decrease in the demand for houses of age 3 households. Now consider the transaction volume for houses: It has risen strongly by age 2 house buyers (grey bar on the left side), but is also partly reduced by the decreased housing transactions of age 3 house buyers (grey bar on the right side). In sum, the transaction volume of houses has risen. The change in transaction volume for flats is given by the black bar in the right panel. Obviously, it is positive as well.

## 4.5 Steady State in the Generalized Model

### 4.5.1 Steady State Equilibrium

In the following, we will see that for age cohorts 1 and 2, the economic environment and the parameter assumptions give rise to the same qualitative housing allocation[15] as in the benchmark model. The propositions which do not directly follow from the text are relegated to the appendix. Without formal proof, we begin by noting that robust numerical results confirm the standard intuition of uniquely defined prices:

**Numerical Result 4.5.1** *In the generalized model, the steady state equilibrium property prices $p_H^*$ and $p_F^*$ are uniquely defined.*[16]

#### Parameter Assumptions and Desirability

Concerning the *affordability* of dwellings, it is clear that the parameter assumptions allow for the same housing allocation as in the case of linear c-utility: A consumption plan $(c_1, c_2, c_3, c_4)$, where $c_i \to 0, (i = 1, 2, 3)$ will lead to total property prices $p_h^{T*},\ h \in \{F, H\}$ in such a way that the same qualitative housing allocation is achieved as in the case of linear c-utility and no transaction costs. However, in the generalized case,

[15] Our notation of the critical indices is as in the benchmark model.

[16] It is not possible to present a closed form solution or a simple equation defining the equilibrium property prices in the generalized case since they depend on optimal consumption in each period. Nevertheless, these optimal decisions are the main drivers of uniqueness.

the question is whether a housing plan is *desirable*. The trade-off in simplified form[17] can be described as follows: The purchase of a (bigger) property increases future housing utility, but simultaneously reduces temporary consumption because of the down payment. In the following we comment on parameter assumptions governing the desirability of dwellings. Our aim is not to provide formal proofs, but to explain the mechanisms at work.

- It is never desirable to trade down[18] in the property market at age 2 because even if the household consumes all remaining wealth after the property purchase at age 1, the total beginning-of-period wealth (including housing wealth) at age 2 is larger than at age 1:

  $$e_2(i) + \underbrace{F_{(T,r)} \cdot p_h}_{>0 \text{ by } (4.2)} > e_1(i), \quad h \in \{F, H\}$$

  Since $u_c$ is concave, the higher the total wealth, the lower is the potential utility loss because of lower temporary consumption. Therefore, if it was optimal to purchase a property $h$ at age 1, it cannot be optimal to downgrade the property at age 2.

- Assumption (4.4) states that the poorest agents of measure $3 - S_F - S_H$ are always of age 1. Due to assumption (4.9), the market mechanism leads to a price $p_F^{T*}$ smaller or equal to $\gamma^{-1} e_1(3 - S_F - S_H)$, so that only age 1 agents with an income index smaller than $3 - S_F - S_H$ do either not have an incentive or do not have the financial possibilities to buy a flat. Since property markets must clear, all other agents have an incentive to buy at least a flat.[19]

- Assumptions (4.7), (4.8) and (4.9) ensure that some flat owners of age 2 have an incentive to acquire a house: Market clearing and assumption (4.9) lead to

[17] The wealth transfer into the future has to be considered as well.

[18] Remember that households derive the housing utility of a property in the period after the purchase.

[19] Due to assumption (4.9), the housing plan $(h_1, h_2) = (\varnothing, H)$ is strictly dominated by $(F, F)$ as far as housing utility is concerned: $0 > -\Delta + u(\frac{1}{2})$. Numerical simulations show that property prices adjust in such a way that this is true even when taking the corresponding holding costs and transaction costs into account.

a house price so that in equilibrium no house is vacant. Because of (4.8), not all houses are held by age 3 households. Assumption (4.7) ensures that at least some of the remaining houses (those properties not held by age 3 households) are bought by age 2 flat owners with high income index. To see this, note that $e_2(1) + F_{(T,r)} \cdot p_F^*$ is a lower bound[20] for the total wealth of the age 2 flat owner with the highest income index, which is larger than the maximally possible wealth of age 1 house buyers, $e_1(1)$, and larger than the maximally possible wealth of age-2 first-time buyers who acquire a house, $e(3 - S_F - S_H)$. By continuity, we see that at least some of the remaining houses must be purchased by age 2 flat owners.

Altogether, we see that in steady state, the generalized model leads to the same qualitative housing distribution up to age 3 cohorts as in the benchmark model. In particular, the critical index for the marginal age 1 flat buyer $i_F^* = 3 - S_F - S_H$ is independent of the c-utility specification. All other critical indices of households of age 1 and 2 might differ quantitatively from the benchmark model since the demand for houses of age 3 households differs slightly from the benchmark case. The driving force behind this result is the market clearing condition.

We summarize these findings in the following proposition.

**Proposition 4.5.2** *Assumptions (4.1) to (4.9) give rise to a housing allocation which is qualitatively the same as in the benchmark case for age 2 and age 3 households.*

However, in the general case, the qualitative housing distribution of the age 4 cohort differs from the linear case. The housing decision of age 3 agents is dependent on their income index. The intuition is the following: The larger $i$, the more agents will opt for a house, since for rich agents (high $i$), the marginal c-utility is low. Therefore, the payoff of more housing consumption is relatively high. In contrast, poor agents consume at a relatively high marginal c-utility level. For them, increasing consumption is advantageous in comparison to the purchase of a house. The following proposition

[20]This is the total wealth in the case when households consume all remaining wealth at age 1 after the (potential) property purchase.

shows that the critical preference index $m$ is a decreasing function of the income index $i$.

**Proposition 4.5.3** *If $u_c$ is strictly concave and $e_3(i)$ is sufficiently large for all $i \in [0,1]$, the critical preference index $m^*(i)$ decreases in the income index $i$.*

### 4.5.2 Equilibrium Prices

In the generalized model, it is no longer possible to derive simple analytical conditions which determine property prices.

**Proposition 4.5.4 (Property Prices)** *If all remaining non-housing wealth is consumed after a (potential) transaction on the housing market at age 1, the following inequalities hold:*

$$0 < p_F^{T*} \leq \gamma^{-1} e_1(3 - S_F - S_H)$$

*and*

$$p_F^{T*} < p_H^{T*} < \gamma^{-1} e_2(1) + \gamma^{-2} F_{(T,r)} e_1(3 - S_F - S_H).$$

*Without the above assumption, we only get the following inequality*

$$p_F^{T*} < p_H^{T*} < \gamma^{-1} e(1).$$

### 4.5.3 Determinants of Agents' Decisions

In the case of a general utility function for consumption, we have to distinguish two cases:

**Case 1.** All critical indices are determined by indifference conditions between housing and consumption.

**Case 2.** Some critical indices are determined by credit constraints.

In the first case, all marginal property buyers are indifferent between two different properties $h$ and $h'$ ($h \neq h'$ and $h, h' \in \{\varnothing, F, H\}$), and their numeraire consumption assumes a value where the sum of housing utility of property $h$, current c-utility and the appendant next-period-value function equals the corresponding sum for $h'$. Obviously, a sufficient condition for the first case is $\lim_{x\to 0} u_c(x) = -\infty$, which ensures that consumption is always strictly positive.

In the second case, some or all critical indices at ages 1 and 2 are determined by credit constraints. Then, the sum of housing utility of property $h$, current c-utility and the appendant next-period value function is strictly higher than the corresponding sum for $h'$. A sufficient condition[21] for case 2 is

$$\Delta > u_c((1+r)^2 r\gamma^{-1} e_1(3 - S_F - S_H)) - \lim_{x\to 0} u_c(x), \tag{4.12}$$

which ensures that the marginal age 1 flat buyer strictly prefers a flat to no property. This condition states that the gain of housing utility ($\Delta$) is always strictly larger than the appendant maximal consumption loss.

Note that even if we assume a c-utility function where the marginal utility tends to infinity at some point, we can still be in the second case because absolute utility values are crucial.[22]

### 4.5.4 Extreme Cases of Consumption Patterns

In the following we present a condition that if fulfilled results in young and poor agents not building any savings. The intuition is that the raise in income is strong enough over the life cycle, such that consumption smoothing implies that after a potential property purchase, all remaining wealth is consumed. This condition can be interpreted as

[21] This condition is a generalization of (4.11) of the benchmark model and has the same interpretation. Note that $\Delta > u(\frac{1}{2})$ is anyway fulfilled in the general case.

[22] If one sets $u_c(x) = \sqrt{x}$, we have $\lim_{x\to 0} u_c'(x) = \infty$, but there exist parameter conditions which give rise to the case where all critical indices at age 1 and 2 are determined by credit constraints.

the extreme case compared to the benchmark model, where consumption is zero and savings are maximal in the first three periods of life.

**Proposition 4.5.5** *If the following conditions hold, age 2 flat buyers*[23] *with income index $i$ will not build any savings in the first two periods of life.*

1. *$e_3(i)$ is sufficiently high.*
2. *$u_c'(e_1(i)) > (1+r)u_c'\big(e_2(i) - e_1(3 - S_F - S_H)\big)$*
3. *$u_c(0) + u_c\big(e(i) - e_1(3 - S_F - S_H)\big) < u_c(e_1(i)) + u_c\big(e_2(i) - e_1(3 - S_F - S_H)\big)$*

The intuition behind this result is that agents try to smooth consumption over the life cycle. Since their income is increasing fast with respect to age, savings would lead to an even stronger inequality of consumption during the life cycle.

In the following we will call the consumption behaviour induced by Proposition 4.5.5 "full consumption".

If the conditions of Proposition 4.5.5 are met, we are able to present an equation which determines the steady state flat price.

**Proposition 4.5.6 (Flat Prices)** *If the three conditions given in Proposition 4.5.5 are fulfilled for all $i \in [0,1]$, the steady state price of flats $p_F^*$ is determined by the following condition.*

$$\begin{aligned} & u_c(e_1(i_F^*)) + u_c\big(e_2(i_F^*) - \gamma(1+T)p_F^*\big) - \Delta \\ = \; & u_c\big(e_1(i_F^*) - \gamma(1+T)p_F^*\big) + u_c\big(e_2(i_F^*) + \big(F_{(T,r)} - \gamma(1+T)\big)p_F^*\big), \end{aligned}$$

*where $i_F^* = 3 - S_F - S_H$.*

[23] If we replaced the second inequality by the stronger condition $u_c'(e_1(i)) > (1+r)u_c'(e_2(i) - \gamma p_H^*)$ for all $i \in [0,1]$, we would be able to guarantee that **no** age one agent will build any savings. Nonetheless, our (weaker) assumption allows for temporary consumption reduction due to expected future property purchases.

In the extreme case of full consumption this proposition shows how to compute the flat price.[24] If we are not in one of these extreme cases of full consumption (as stated in Proposition 4.5.5) or zero consumption (as in the benchmark case), the equations determining flat prices become extremely tedious.

## 4.6 Exogenous Income Shocks in the Generalized Model

### 4.6.1 Functional Specifications and Parameter Assumptions

In the following, we use numerical simulations to study the qualitative and quantitative effects of an exogenous shock to income. Our aim is to compare the resulting effects in case 1 (marginal property buyers are determined by indifference conditions between consumption and housing utility) and case 2 (only credit constraints determine the marginal income indices). In order to isolate the effect of consumption smoothing (only observable in case 1) we begin by setting transaction costs to zero. In the second part, we separately analyse the effects of transaction costs.

#### Functional Specifications

In order to derive results in the generalized case, it is necessary to specify functional terms for the c-utility function, the function determining the utility of a house and the income functions. Besides the standard linear c-utility function $u_c(c) = c$, we use

$$u_c(x) = \ln(x),$$

which guarantees strictly positive numeraire consumption for all households.[25] As $\lim_{x\to 0} \ln(x) = -\infty$, this utility function is a representative of case 1. We set the

[24] Concerning the house price, it is not possible to give such a simple equation because of the complicated expression of age 3 demand.

[25] Households with income index i=0 have mass zero, therefore we can neglect them.

function determining the utility of a house to

$$u(m) = 10 \cdot m - 3.$$

The following table shows the specification of the income functions.

| Age | Functional Form |
|---|---|
| 1 | $e_1(i) = i$ |
| 2 | $e_2(i) = 1 + 0.2i$ |
| 3 | $e_3(i) = 3 + i$ |
| 4 | $e_4(i) = i$ |

**Parameter Specifications**

The model parameters are chosen as follows.

| Parameter | Variable | Value |
|---|---|---|
| Interest rate | $r$ | 0.05 |
| Rate of down payment | $\gamma$ | 0.10 |
| Supply of houses | $S_H$ | 0.80 |
| Supply of flats | $S_F$ | 2.00 |
| Rate of transaction costs | $T$ | 0.00 |
| Factor of income shock | $z$ | 1.04 |

The value which determines the utility of a flat, $\Delta$, is set to 2.11025. Note that we only consider positive and uniform[26] shocks to the income structure. Besides fulfilling conditions (4.1) to (4.9) of the generalized model, the specifications given above also fulfill the (stronger) conditions given in the benchmark model. Furthermore, the parameter and functional specifications

- guarantee that house buyers are repeat buyers, that is, each house buyer already owns a property.

[26]Income functions $e_j$, $j = 1, \ldots, 4$ are multiplied by the same factor $z > 1$.

- ensure that young and poor households consume all remaining wealth after a potential housing transaction in each period (full consumption).
- trigger an overreaction[27] pattern in house prices after income shocks in the linear case.

See the appendix for explanations and formal details.

In summary, the functional specifications and parameter assumptions fulfill all conditions in the generalized setting and the benchmark case as well. Regarding the logarithmic c-utility function, they give rise to case 1. Hence, all critical indices are determined by indifference conditions rather than credit constraints. Moreover, the environment is chosen such that it should generate overshooting in house prices.

### 4.6.2 Numerical Analysis

#### Linear C-Utility

In order to compare our setting with the benchmark model we begin by presenting the numerical results in the case of linear consumption utility. The following table shows the values of the corresponding equilibria after an exogenous shock to income with factor $z = 1.04$. The per-period transaction volume is denoted by $n_F$ and $n_H$, respectively.

| | old steady state * | directly after shock + | new steady state ** |
|---|---|---|---|
| $p_F$ | 2.000 | 2.080 | 2.080 |
| $p_H$ | 19.176 | 20.256 | 19.902 |
| $i_F$ | 0.200 | 0.200 | 0.200 |
| $i_{FH}$ | 0.814 | 0.774 | 0.810 |
| $m$ | 0.386 | 0.426 | 0.389 |
| $n_F$ | 1.071 | 1.079 | 1.072 |
| $n_H$ | 0.685 | 0.693 | 0.684 |

[27] Overreaction of prices is defined as $|p_H^+ - p_h^*| > |z-1|p_H^*$.

Since we do not have any first-time house buyers in the economy, the exact values of $i_{\varnothing H}$ and $i_H$ are not relevant.

Note that these results are in line with the theoretical work in Ortalo-Magné and Rady (2006): The price of flats adjusts immediately to its steady state level and the change is proportional to the change in endowments ($z = 1.04$).

House prices show the overshooting pattern ($p_H^+ > p_H^{**}$) and the absolute change in steady state house prices is higher than the corresponding change in flat prices.

$$0.72 \approx p_H^{**} - p_H^* > p_F^{**} - p_F^* = 0.08$$

However, the steady state price of houses changes less than proportionally with endowments.

$$\frac{p_H^{**}}{p_H^*} \approx 1.037 < 1.040 \approx \frac{p_F^{**}}{p_F^*}$$

In the period directly after the shock we identify the appendant increase in trading volume.

$$n_F^+ > n_F^* \text{ and } n_H^+ > n_H^*$$

Hence we have a positive correlation of prices and transaction volume after an exogenous shock to income.

Figure 4.3 illustrates the numerical results in the linear case.

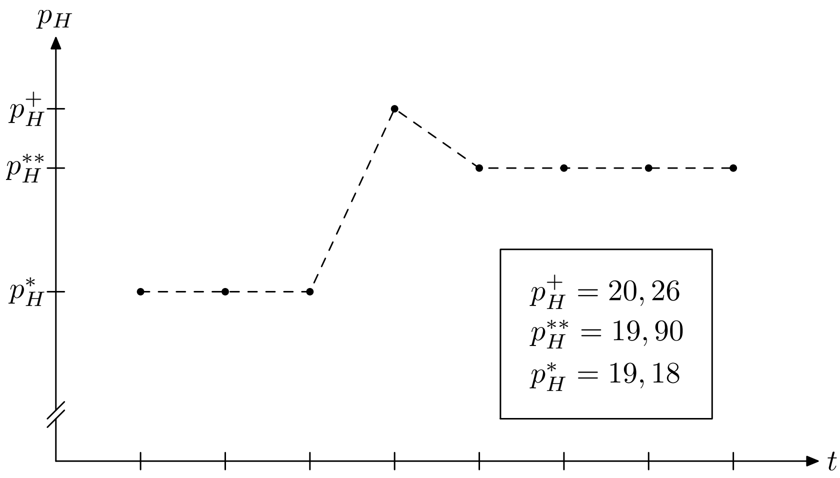


Figure 4.3: House Price Dynamics with Linear C-Utility

**Logarithmic C-Utility**

Now we turn to the main results of this work. Using the same parameter specifications, we simulate the model using the logarithmic c-utility function $u_c(c) = \ln(c)$. The following table summarizes the numerical results.

| | old steady state * | directly after shock + | new steady state ** |
|---|---|---|---|
| $p_F$ | 1.782 | 1.855 | 1.855 |
| $p_H$ | 14.336 | 14.931 | 14.910 |
| $i_F$ | 0.200 | 0.200 | 0.200 |
| $i_{FH}$ | 0.866 | 0.866 | 0.866 |
| $m$ average | 0.334 | 0.334 | 0.334 |
| $n_F$ | 1.045 | 1.045 | 1.045 |
| $n_H$ | 0.711 | 0.711 | 0.711 |

Once again, we realise that the flat price adjusts immediately to its steady state level. We also observe an overshooting pattern in prices, but much weaker than in the linear case. The results also indicate that the change in the house price is more in absolute terms and less in proportional terms than the change in the flat price.

The weak overshooting pattern goes in line with a small effect on the critical indices. Hence, the corresponding effect on transaction volume is also small. As outlined in the appendix, our numerical approach heavily exploits the mathematical structure of the model and the fact that the steady state prices adjust within two periods.[28]

Figure 4.4 illustrates these numerical results in the case of a logarithmic utility function.

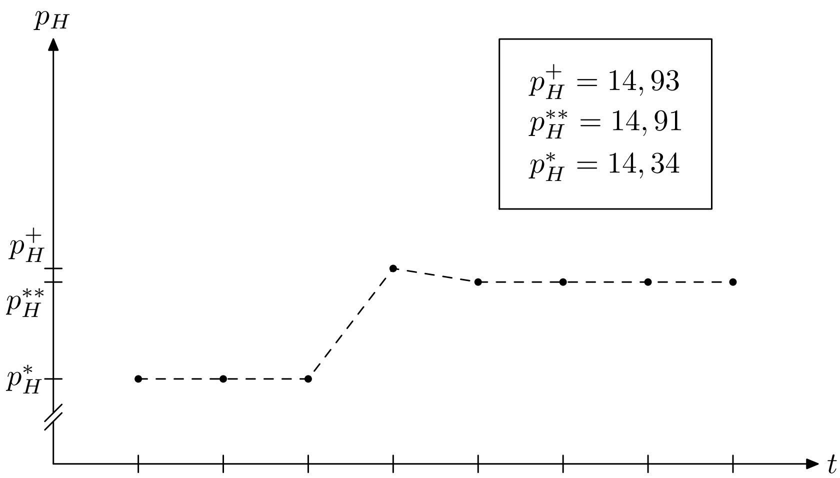


Figure 4.4: House Price Dynamics with Logarithmic C-Utility

As proved in Proposition 4.5.3, the critical preference index $m^*$ is decreasing in the income index. In the table above, only an average value is given, whereas Figure 4.5 shows the dependence of the critical preference index on the income index.

[28]In order to calculate the exact values in the first period after the shock, we need the consumption and property distributions of the old steady state and the new steady state. Since we only use discrete approximations for these distributions, our calculating accuracy is about 1e-3 to 1e-4. Therefore the effect on transaction volume is too small to capture it numerically.

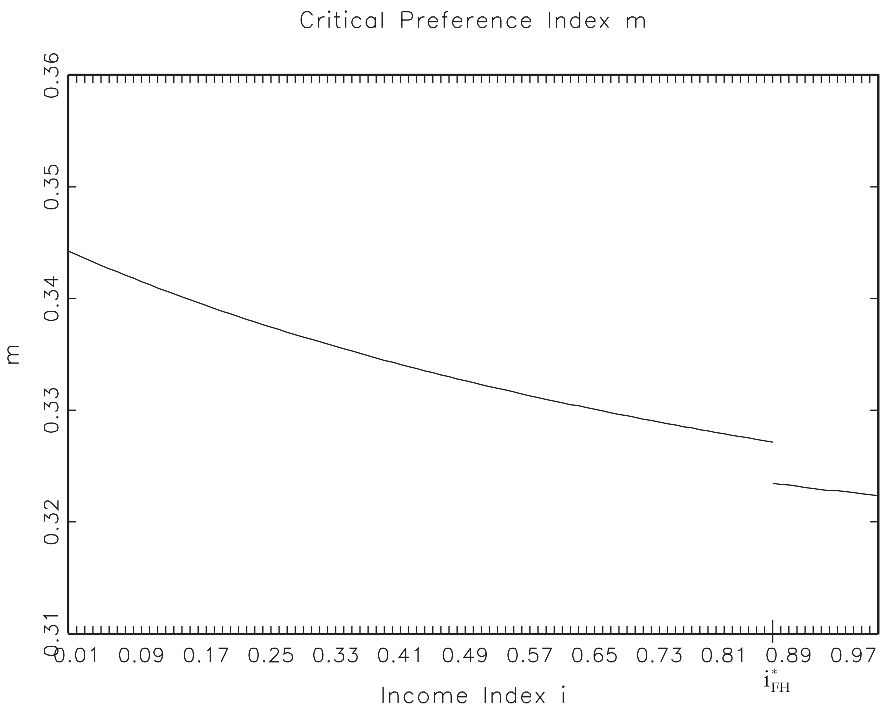


Figure 4.5: The Critical Preference Index $m^*$

Note that the discrete jump at $i^*_{FH} \approx 0.866$ stems from the discrete jump in age 3 wealth of those households that bought a property at age 2 and thereby transferred more money into age 3.

The logarithmic c-utility specification gives rise to strictly positive consumption in each period. The following figures show the optimal steady state consumption path dependent on the income index.

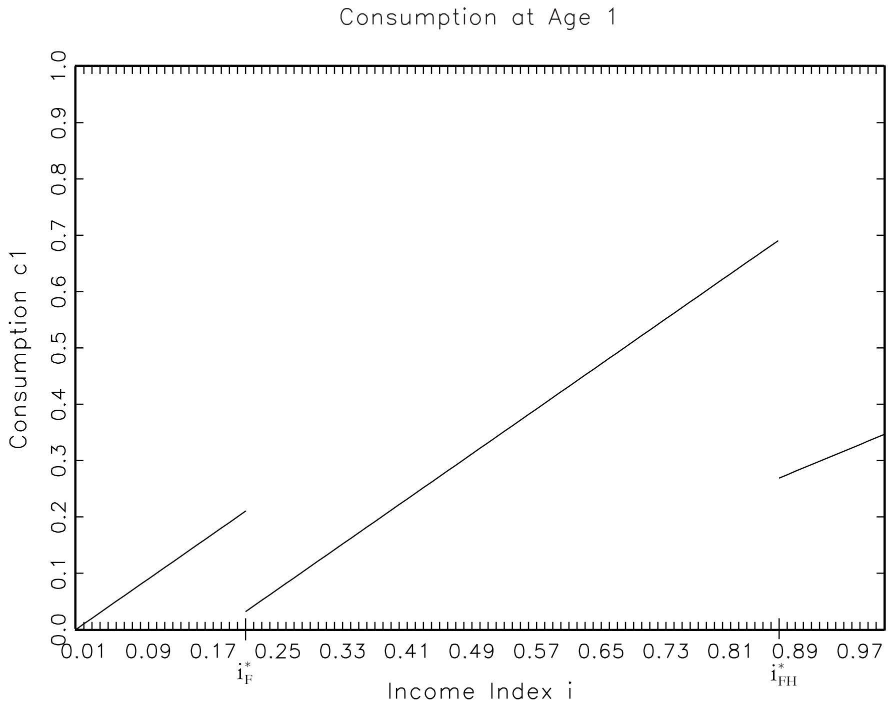


Figure 4.6: Age 1 Consumption with Logarithmic C-Utility

The discrete jumps in the consumption structure are due to the fact that all age 1 agents with income index higher than $i_F^* = 3 - S_F - S_H = 0.2$ buy a flat at age 1. As proved earlier, the conditions for full consumption are met. Thus, all age 1 agents, except for the wealthiest, consume all remaining wealth after a potential housing transaction.

Those with income index higher than $i_{FH}^*$ do not show full consumption at age 1. They prefer to save money in order to smooth consumption because they anticipate the purchase of a house at age 2. Hence, they divide their wealth optimally between the first two periods of life.

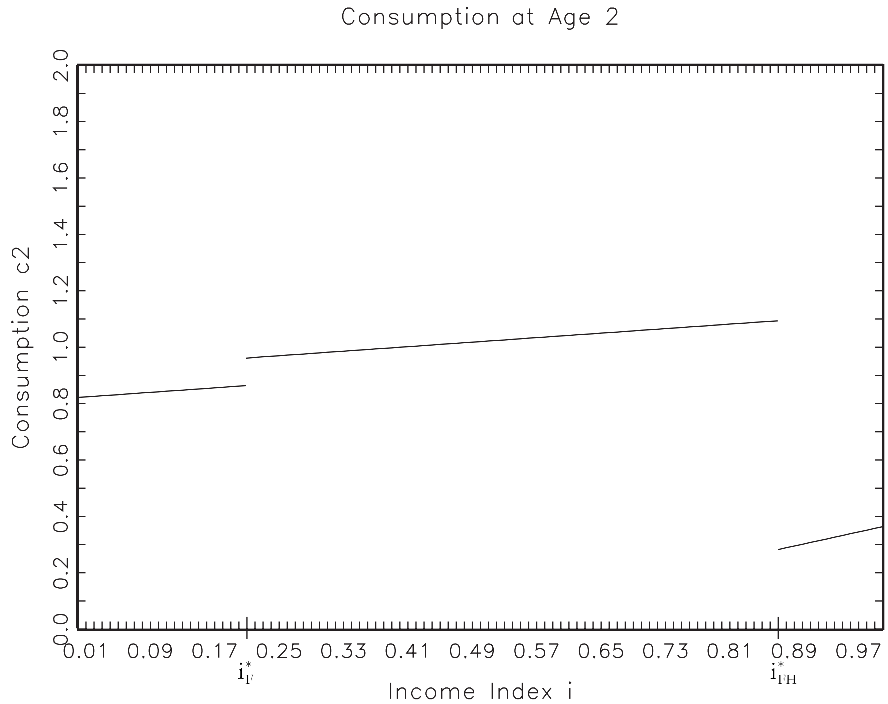


Figure 4.7: Age 2 Consumption with Logarithmic C-Utility

As noted in the beginning, there is a discrete jump in consumption for age 2 households at $i_F^*$. This discontinuity is due to the fact that flat buyers have consumed less at age 1 and therefore transferred more total wealth into age 2 which they now use to increase consumption. For all age 2 agents we observe the full consumption behaviour. In our specification, age 2 households never build any savings because they anticipate that their age 3 income is sufficiently high such that in all cases savings would even increase the imparity of consumption over the life cycle. For $i > i_{FH}^*$, their consumption is reduced tremendously because of the down payment for the house.

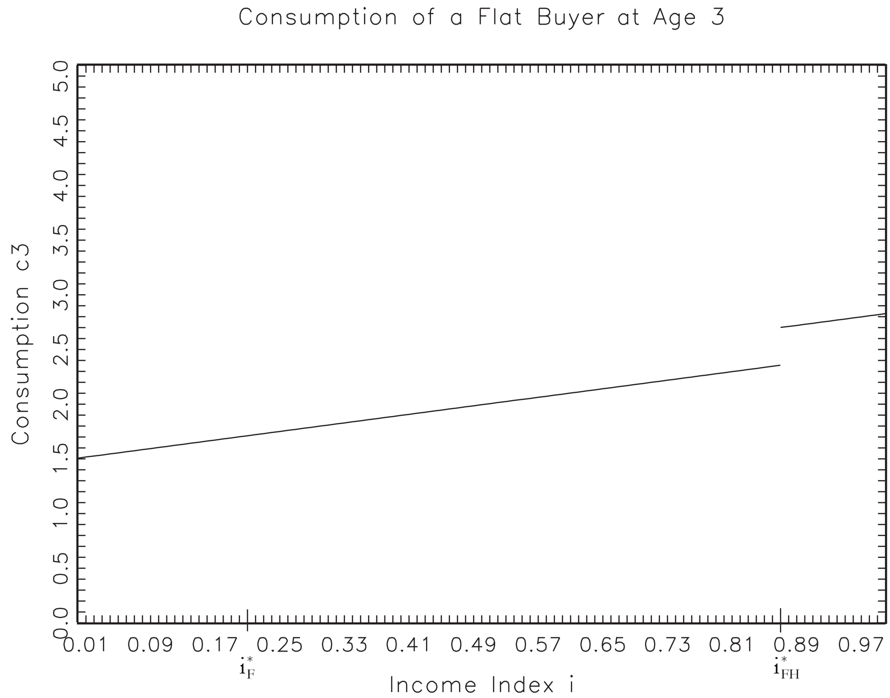


Figure 4.8: Age 3 Consumption with Logarithmic C-Utility for Age 3 Flat Buyers

In contrast to age 1 and age 2 agents, age 3 households do not show full consumption because they want to smooth consumption over the last two periods of life. They anticipate that their age 4 income is much lower than at age 3.

Again, the effect of increased wealth transfer into the future by holding a property can be observed: age 3 agents with income index larger than $i^*_{FH}$ consume discretely more than those who did not buy a house in period 2.

Note that age 3 and age 4 consumption are only shown for age 3 flat buyers ($m < m^*$). The corresponding graphs for house buyers ($m > m^*$) are qualitatively the same. However, due to the purchase of a house, their consumption is shifted downwards in periods 3 and 4.

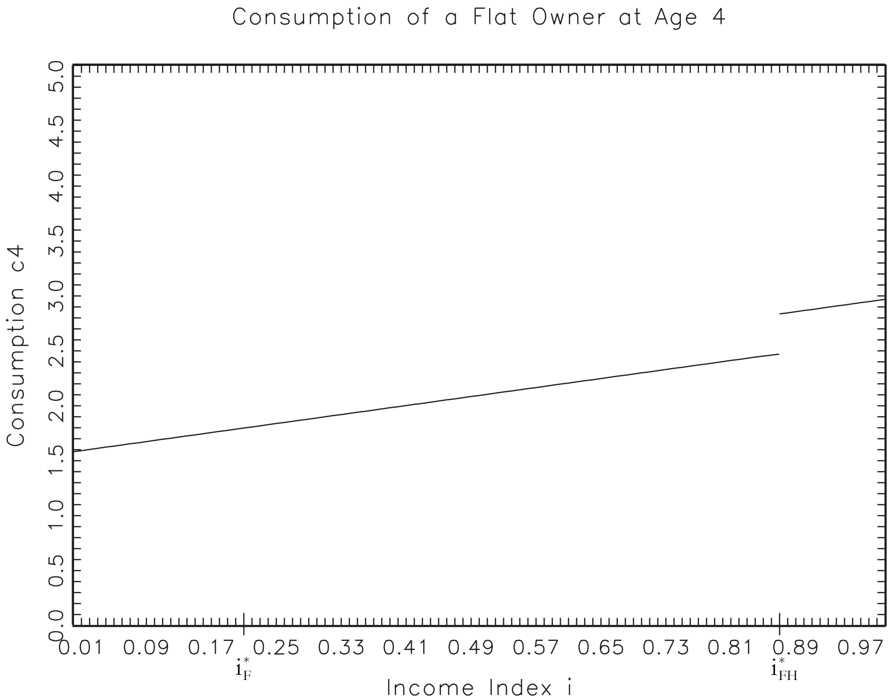


Figure 4.9: Age 4 Consumption with Logarithmic C-Utility for Age 3 Flat Buyers

The age 4 consumption function is very similar to the consumption function at age 3, except for the fact that it is slightly larger. This is due to the positive interest rate. It is not optimal to consume the same amount in both periods because forgoing a unit of c-consumption at age 3 leads to $1 + r$ units of consumption at age 4.

**Relative Effects**

Let us consider the short and the long time change in price after an exogenous shock.

First, turn to the long time effects. We consider the change in steady state prices

$$f_h^l = \frac{p_h^{**}}{p_h^*}, \quad h \in \{F, H\}.$$

Using the numerical values given above, we get

$$f_H^l(\text{linear}) = 1.037 \text{ and } f_H^l(\ln) = 1.040,$$

$$f_F^l(\text{linear}) = 1.040 \text{ and } f_F^l(\ln) = 1.041,$$

Hence we can see the relative long-time effect on property prices is larger in the case of logarithmic c-utility than in the linear model. The reason is the following: In case of the logarithmic utility function, all agents act on a flatter part of the c-utility function after the income shock. Thus, consumption utility becomes less important and the willingness to acquire a property is increased. This translates directly into a stronger effect on property prices.

Regarding the short time effects, we only consider the change in house prices

$$f_H^s = \frac{p_H^+}{p_H^*}$$

Using the numerical values given above, we get

$$f_H^s(\text{linear}) = 1.056 \text{ and } f_H^s(\ln) = 1.045,$$

which indicates a stronger effect on house prices in the linear model than in case of logarithmic c-utility. Besides the first effect, a second effect comes into play. Consider the case of logarithmic c-utility. By the increase in income, all households become richer. Furthermore, capital gains on flats increase the wealth of flat owners more than proportionally. A certain number of them is now better off when acquiring a house. However, some of the poorest households of that number are still better off when they invest their additional wealth in consumption. Hence the demand for houses does not increase as much as in the case of linear c-utility where all agents who have the possibility to buy a house will do so. Because of this appendant increase in consumption, the price of houses does not increase as strongly as in the linear case.

### 4.6.3 Transaction Costs

Let us turn to the situation of prevailing positive transaction costs $T$ in the economy. For simplicity, consider the case of linear c-utility. As we have seen, the total steady state price of flats continues to be determined by the equation

$$\gamma p_F^{T*} = e_1(3 - S_F - S_H),$$

hence we get

$$p_F^* = \gamma^{-1} \frac{e_1(3 - S_F - S_H)}{1+T}.$$

For the numerical exercise we set the parameter $T = 0.05$, and leave all other specifications unchanged. The following table summarizes the results.

| | old steady state * | directly after shock + | new steady state ** |
|---|---|---|---|
| $p_F$ | 1.905 | 1.981 | 1.981 |
| $p_H$ | 17.278 | 18.304 | 17.590 |
| $i_F$ | 0.200 | 0.200 | 0.200 |
| $i_{FH}$ | 0.808 | 0.770 | 0.777 |
| $m$ | 0.392 | 0.334 | 0.334 |
| $n_F$ | 1.076 | 1.083 | 1.094 |
| $n_H$ | 0.683 | 0.690 | 0.671 |

Compared to the situation without transaction costs, the steady state house price changes more than proportionally.

$$p_H^{*T>0} < \frac{1}{1+T} \cdot p_H^{*T=0}.$$

The intuition is simple. Since the flat price adjusts as given in the equation above, only the house price can capture the effect of an increased incentive to stay in the old home. Compared to the case of zero transaction costs, agents now have less incentive to climb up the property ladder successively and to change their dwelling due to preference reasons. Thus, their willingness to pay for houses, once they have acquired a flat, is much lower than without transaction costs.

Note further that the relative overshooting effect is even higher than in the case without transaction costs.

$$\frac{p_H^{T+}}{p_H^{T**}} \approx 1.041 > 1.018 \approx \frac{p_H^+}{p_H^{**}}.$$

The numerical simulations do not confirm any further the positive correlation of prices and transaction volume.

Transaction costs in case of logarithmic c-utility do not add further insights in the understanding of consumption smoothing effects in the model. Therefore, the consideration of this combination is neglected.

## 4.7 Conclusion

We generalized the model of Ortalo-Magné and Rady (2006) by introducing a general concave utility function for non-housing wealth. Furthermore, we studied the consequences of transaction costs for the model equilibrium. This enabled us to test the robustness of the model when more realistic assumptions are made.

We found that under appropriate conditions the housing allocation for age 1 and 2 households is robust to the introduction of a general c-utility function. The curve which describes the marginal house buyer at age 3 is dependent on the preference index and on the income index as well. We presented conditions which give rise to an equilibrium where consumption is non-zero for almost all households. We also determined conditions by which young and poor households immediately consume all remaining wealth left after potential housing transactions, so-called full consumption. We derived numerically the optimal consumption paths for the c-utility function $u_c(x) = \ln(x)$ and compared them to the benchmark case of linear consumption utility. We found that overshooting does not differ qualitatively from the linear case under certain parameter assumptions.

However, we identified two countervailing effects. Under the logarithmic specification, the relative difference of steady state prices is larger than in the benchmark case, whereas the overshooting effect is smaller. The steady state effect on relative prices is easily explained by a decreased relative willingness to pay for consumption in the case of a strictly concave utility function. The overshooting effect is smaller because not all capital gains on real estate are reinvested in the housing market under a concave utility function.

In the last part we introduced transaction costs and found that they have a direct proportional influence on the flat price, but tend to lower the price of houses overproportionally. This is due to the fact that households tend to avoid repeat buys in the real estate market in order to save on transaction costs. In this model, both the absolute and the relative price dispersion of dwellings is significantly reduced by the introduction of transaction costs.

## 4.8 Appendix

### 4.8.1 Proofs

**Proof of Proposition 4.5.3.** Note first that age 4 consumption always will be equal to or higher than consumption at age 3 for all $i \in [0,1]$ because after the property purchase at age 3, households divide their remaining wealth between periods three and four such that they smooth their consumption optimally. They take into account that there will be no property purchase at age 4. If $e_4(i)$ is sufficiently large, age 4 consumption is even strictly higher than age 3 consumption.

Define $c_j^h(i)$ as the optimal consumption at age $j, (j = 3, 4)$ of household $i$ when his age 3 housing decision is $p_h, (h = F, H)$. Note that the housing decision at age 3 defines the housing utility at age 4 and that $c_4^h(i)$ is uniquely determined by the choice of $c_3^h(i)$. The marginal age 3 house buyer with income index $i$ is indifferent between a flat and a house:

$$U(F, m^*(i)) + u_c(c_3^F(i)) + u_c(c_4^F(i)) = U(H, m^*(i)) + u_c(c_3^H(i)) + u_c(c_4^H(i)) \qquad (4.13)$$

$$\Leftrightarrow 0 + u_c(c_3^F(i)) + u_c(c_4^F(i)) = u(m^*(i)) + u_c(c_3^H(i)) + u_c(c_4^H(i))$$

$$\Leftrightarrow m^*(i) = u^{-1}[u_c(c_3^F(i)) - u_c(c_3^H(i)) + u_c(c_4^F(i)) - u_c(c_4^H(i))]$$

The equation above considers the marginal age 3 house buyer and states that the housing utility derived at age 4 must equal the sum of foregone consumption utility[29] at age 3 and 4. With increasing total wealth, the foregone consumption looses importance because of the concavity of the utility function. Since $e_3(i)$ is assumed to be sufficiently large for all $i \in [0,1]$, the total wealth of age 3 agents is increasing in the income index $i$. Hence, the right hand side of equation (4.13) is decreasing in $i$ and so is $m^*$.

q.e.d.

**Proof of Proposition 4.5.4.** The price of flats is strictly positive because $S_F < 3$

[29]In fact, foregone consumption at age 4 can be negative due to a positive wealth transfer by the property purchase.

and $U(\varnothing, m) < U(F, m)$ for all $m \in [0,1]$. Since only age 1 agents with income index lower than $i = 3 - S_F - S_H$ cannot afford or do not want to buy a flat, we get

$$\gamma p_F^{T*} \leq e_1(3 - S_F - S_H) - c_1.$$

Since $c_1$ represents non-negative consumption at age 1, the inequality concerning the flat price $p_F^{T*}$ follows immediately.

We already know that age 2 households with the highest income index are house buyers. Without the assumption that all remaining non-housing wealth is consumed at age 1, an upper bound for their maximal wealth at age 2 is $e(1)$ and we get

$$\gamma^{-1} e(1) > p_H^{T*}.$$

With the assumption that all remaining non-housing wealth is consumed at age 1, their maximal wealth is $e_2(1) + F_{(T,r)} p_F^*$. Hence we get

$$e_2(1) + F_{(T,r)} p_F^* > \gamma p_H^{T*}.$$

Together with the inequality $p_F^* \leq \gamma^{-1} e_1(3 - S_F - S_H)$ we see

$$p_F^{T*} < p_H^{T*} < \gamma^{-1} e_2(1) + \gamma^{-2} F_{(T,r)} . e_1(3 - S_F - S_H).$$

Finally, suppose $p_H^{T*} \leq p_F^{T*}$. Then, because of assumption (4.9), the demand for flats will be zero for age 1 and age 2 households. Moreover, again due to assumption (4.9), the demand of age 3 households is less than $\frac{1}{2}$. From assumption 4.1, we easily get $S_F > 1\frac{1}{2}$, such that the flat market cannot be in equilibrium. Contradiction.
q.e.d.

**Proof of Proposition 4.5.5.** Consider households that are flat buyers at age 2. First, note that $e_2(i) - \gamma p_F^{T*}$ is their minimal remaining wealth left for consumption at age 2. Due to Proposition 4.5.4, a lower bound for this expression is given by

$$e_2(i) - e_1(3 - S_F - S_H).$$

On the other side, $e_1(i)$ is an upper bound for their remaining wealth left for consumption at age 1. Consider the following maximization problem on savings:

$$\max_{\varepsilon\in[0,e_1(i)]} u_c(e_1(i)-\varepsilon)+u_c\big(e_2(i)-e_1(3-S_F-S_H)+(1+r)\varepsilon\big)$$

Due to the third condition, $e_1(i)$ cannot be a solution to this maximization problem. Therefore a strictly positive solution must fulfill the first order condition:[30]

$$u_c'(e_1(i)-\varepsilon)=(1+r)u_c'\big(e_2(i)-e_1(3-S_F-S_H)+(1+r)\varepsilon\big)$$

Since $u_c$ is concave, the left hand side is weakly increasing in $\varepsilon$, whereas the right hand side is weakly decreasing in $\varepsilon$. Setting $\varepsilon$ to zero, we get

$$u_c'(e_1(i))=(1+r)u_c'\big(e_2(i)-e_1(3-S_F-S_H)\big),$$

which contradicts condition 2. Hence this maximization problem does not have an interior solution. Therefore $\varepsilon=0$ is the unique solution which means that savings at age 1 cannot be optimal. Since we assume that $e_3(0)$ is sufficiently high, there cannot be any incentive to save in the first two periods of life in order to increase consumption at age 3 or age 4.
q.e.d.

**Proof of Proposition 4.5.6.** The marginal age 1 flat buyer is indifferent between no dwelling and a flat:

$$\text{utility}(\varnothing)=\text{utility}(F).$$

Since the conditions in Proposition 4.5.5 ensure full consumption of the marginal age 1 flat buyer at age 2 and his housing decision at age 2 will be a flat anyway, the total wealth transfer into age 3 amounts to $F_{(T,r)}p_F^*$, independently of the property decision at age 1. Hence, the indifference condition can be stated in terms of age 1 and age 2 consumption.

$$\begin{aligned}&u_c(c_1(i_F^*)\,|\,h_1=\varnothing)+u_c(c_2(i_F^*)\,|\,h_1=\varnothing)+U(\varnothing,\frac{1}{2})\\ &=u_c(c_1(i_F^*)\,|\,h_1=F)+u_c(c_2(i_F^*)\,|\,h_1=F)+U(F,\frac{1}{2})\end{aligned}$$

[30]Note that this condition cannot be fulfilled in the case of linear c-utility.

Again, due to full consumption at age 1, we get the following consumption levels for age 1 households.

$$c_1(i_F^* \,|\, h_1 = \varnothing) = i_F^* = 3 - S_F - S_H \text{ and } c_1(i_F^* \,|\, h_1 = F) = 3 - S_F - S_H - \gamma p_F^{T*}$$

At age 2 the marginal age 1 flat buyers will decide on a flat anyway. Having purchased no flat at age 1, the household's age 2 consumption level is

$$c_2(i_F^* \,|\, h_1 = \varnothing) = e_2(i_F^*) - \gamma(1+T)p_F^*$$

Having bought a flat at age 1, the household's age 2 consumption level is

$$c_2(i_F^* \,|\, h_1 = F) = e_2(i_F^*) + F_{(T,r)}p_F^* - \gamma(1+T)p_F^*$$

Plugging in these values and the corresponding values for housing utility in the equation above, we immediately get the result.
q.e.d.

### 4.8.2 Further Conditions Fulfilled by Specification

#### Repeat Buyers

Concerning a linear c-utility function, by our specification we fulfill the following condition which ensures that all houses are purchased by repeat buyers, that is, each house buyer already owns a property. This creates a situation where flat price changes have the strongest possible effect on transitional dynamics.

$$e(\frac{3}{2} - S_H) \geq \max\{e_1(1), e(3 - S_F - S_H)\} + r\gamma^{-1}e_1(3 - S_F - S_H) \tag{4.14}$$

In fact, since $u(1 - S_H) < 0$ and $u(\frac{1}{2}) > 0$, at least a measure of $\frac{1}{2}$ houses is purchased by age 3 households in each period. The measure of remaining houses is positive, but lower than $S_H - \frac{1}{2}$. Consider now the richest age 2 households of measure $S_H - \frac{1}{2}$. A lower bound for their income index is

$$1 - (S_H - \frac{1}{2}) = \frac{3}{2} - S_H. \tag{4.15}$$

Since we abstract from transaction costs and consumption is zero at age 1, we only have to consider holding costs. Hence, the total wealth of this cohort is at least

$$e(\frac{3}{2} - S_H) - r\gamma^{-1}e_1(3 - S_F - S_H)$$

which is larger than the maximally possible wealth of age 1 house buyers, $e_1(1)$, and larger than the maximally possible wealth of age-2-first-time buyers who acquire a house, $e(3 - S_F - S_H)$. Thus this assumption ensures that age 2 households with a high income index are the remaining house buyers.

Moreover, in the case of logarithmic c-utility and full consumption, condition (4.14) reduces to

$$e_2(\frac{3}{2} - S_H) + F_{(T,r)}p_F^* \geq \max\{e_1(1), e_2(3 - S_F - S_H)\},$$

which is also fulfilled in our setting. By the same reasoning as in the proof of Proposition 4.5.4, we see that in this case, we get the following inequality determining house prices:

$$p_H^{T*} < \gamma^{-1}e_2(\frac{3}{2} - S_H) + \gamma^{-2}F_{(T,r)}e_1(3 - S_F - S_H),$$

which is stronger than in Proposition 4.5.4.

**Full Consumption**

By our specification the conditions of Proposition 4.5.5 are met for all households with income index[31] lower than 0.94, such that age 2 flat buyers show full consumption at age 1 and age 2. That is, all remaining wealth after a potential transaction on the real estate market is consumed immediately.

[31] Consider the inequality $u_c'(e_1(i)) > (1+r)u_c'(e_2(i) - e_1(3 - S_F - S_H))$ and plug in the parameters and functional specifications: the result follows immediately.

**Overreacting Prices**

Ortalo-Magné and Rady (2006) present some further conditions[32] which when met, guarantee an overreaction of house prices in the case of a linear c-utility function. Concerning these conditions, our specification gives rise to the following (in-)equalities.

1. $e_1(i^*_{FH}) < p^*_F$
2. $\min\{i^*_F, i^*_{\varnothing H}\} = i^*_F$
3. $(u^{-1})' \leq \frac{1}{10}$

In the linear case, these conditions are likely to trigger an overreaction pattern of prices.

### 4.8.3 Transition after the Shock

It is crucial to understand the mechanism that determines the length of the transition.

Again consider the generalized setting. Due to the underlying competitive market assumption, the mass of people not able or not willing to buy a home (except for the age 4 cohort) must equal the mass of age 1 agents who have an endowment less or equal the required down payment on a flat. Since this argumentation is independent of past income of any kind the following equation must hold in any period.

$$\gamma p_F = e_1(3 - S_F - S_H)$$

Hence, flat prices adjust immediately and directly proportionally to the endowment of age 1 households.

Let us now consider the dynamics of the house price after a proportional income shock of factor[33] $z > 0$ between period 0 and 1. In period 1 the price overshooting takes place as explained above. This is induced by the increased demand for houses by age 2 households that bought a flat before the shock. They enjoy high capital gains after the

[32] The value $\frac{1}{10}$ in condition 3 represents an arbitrarily chosen upper bound. This condition ensures that the increase of $u^{-1}$ is relatively low around the user cost differential between houses and flats.

[33] The argumentation for $z < 0$ is analogous.

shock because they profit both from higher endowment income and from an increased flat price.

Now consider period 3. Age 1 households get the after-shock income, therefore their willingness to pay for a house is the same as in the new steady state.

Age 2 households were of age 1 in period 1. Those who did not buy a dwelling got the new steady state income at age 1 and age 2. Hence, their contribution to the demand for houses is the same as in the new steady state. Age 2 households that acquired a flat at age 1 also enjoyed the new endowment income in both periods. Moreover, they bought the flat at the new steady state price. Hence their personal state is also the same as in the new steady state equilibrium.

As we have seen in Proposition 4.5.3, the housing decision at age 3 depends on the wealth of age 3 households. In period 1, age 3 households were of age 2 and already got the new steady state income. Only their age 1 income differs from the corresponding steady state income. But since we are in the case of full consumption, age 1 income does not affect the wealth at age 3.

In summary, the willingness to pay for a house is the same as in the steady state for all age cohorts. Since supply is fixed, the equilibrium price of houses in period 2 already must take its steady state value. This argumentation is crucial for our approach since it is independent of numeraire consumption.

### 4.8.4 Numerical Approach

Standard theory of the numerical simulation of models with heterogenous agents suggests three methods to determine the equilibrium. The first makes use of Monte Carlo simulations and tracks a specified sample of households over their life cycle. The second is the standard one: the discretization of the distribution function. The third method would assume a predefined class of functions (as Taylor series or orthogonal polynomials) and solve for the corresponding coefficients. All three methods, used in the standard way for all state variables, lead to prohibitively high processing time for sufficiently smooth results.

Therefore, it is necessary to heavily exploit the mathematical and economic structure. In particular, we make use of the fact that for sufficiently small changes in the endowment index the new steady state house prices are already reached in the second period after a shock.

Our numerical approach is to solve first for both steady states. Then, we take the appendant prices and distributions of wealth and housing as given and search for the equilibrium price of houses in the period after the shock by discretisation of its state space.

# Chapter 5

# Concluding Remarks

In this chapter we provide some concluding remarks on the three models presented in this thesis. We discuss some features of the model settings and limitations of our approaches. Furthermore, we offer possible directions for further research.

Our model of urban transport in chapter two studies the interaction of traffic participants within an urban area. The resulting equilibrium distribution suggests that a star shaped commuting zone emerges, where it pays off for residents to commute to the central district. The edge points of this area are defined by the farthest public transport stations from the centre which serve for commuter traffic. This should have consequences on house prices and rents as well. It could be a fruitful direction of further research to study the consequences of location or distance to a public transport on the real estate market in the presence of congested roads. To our knowledge the existing literature has not developed theoretical models which treat this topic by studying the price and welfare effects of optimal spatial traffic distortion and optimal traffic regulation by tolling.

A simplifying assumption we make in our model of urban traffic is the constant population density over the whole urban area. Yet most cities show the highest population density in housing areas close to the centre and lower densities in the suburbs. Since the effects of bottlenecks and tolls in our model depend on the number of distorted commuters, the optimal positions of bottlenecks and tollgates change as well and the effect of reduced inner-city congestion due to reduced long-distance commuter traffic is weakened. However, introducing a population density of such kind would change our results quantitatively but not qualitatively. In particular, we tested for linearly decreasing densities and found that our model is robust with respect to changes in population density, as long as no extreme distributions are assumed.

In the model, all agents derive the same benefit from a commuting trip. Introducing heterogeneity on the side of the commuters with respect to their willingness to pay for a commuting trip and allowing to choose residence location will lead to a residence distribution. Commuters who derive high utility from a commuting trip or those who value monetary costs less than time costs are then expected to live closer to the centre than others. Introducing these kinds of features would compound our results concerning the welfare implications because in this situation the distortion of long-distance commuters would have the same effect on congestion costs, but the welfare loss due to their distorted decisions would be lowered.

Concerning policy implications our results should be interpreted carefully. Road pricing or capacity reduction without improvement of the public transport system will lead to public resentment and political objections, but public transport investment alone will not counter the attractions of the car. The point made in this chapter is to show that induced by the spatial structure of traffic systems in large cities, local city governments have strong incentives to distort car-commuting traffic negatively.

The model of the rental market in chapter three is extremely stylized and makes use of an exogenously specified function that determines the distribution of tenant qualities associated with a prevailing vacancy rate and rental prices. This setting greatly simplifies the analysis of exogenous shocks to the cost structure of landlords. However, it would be interesting to see how our results would change if we allow the matching process of landlords and renters to be probabilistic. This could be achieved using a search model, where at the beginning of each period landlords set a rental price and tenants set a maximally acceptable rent. When landlords determine the rent they quote, they face uncertainty about tenant quality, while tenants face uncertainty about the rent they have to pay. Furthermore, both sides must consider the risk of not being matched at all. Although this structure considerably complicates the model, it would lend more realism to our analysis. In turn the implications of our model could be tested empirically. Our intuition is that this setting would not change the results. However, we could gain new insights concerning the duration of the searching process and the interrelation between tenant quality and the length of search.

In chapter four, we studied a generalized version of the model of Ortalo-Magné and Rady (2006) by introducing a more general class of utility functions and transaction costs. A further result derived in this thesis is that their model cannot be extended into a multi-period model which is amenable to empirical verification without changing the basic structure and the underlying crucial parameter assumptions. The reason is that the effects of the model depend heavily on the low number of generations in the model and on the fact that the decision problems of the age cohorts are not continuously dependent on the decision problems of other age cohorts. In the basic model, the decision of age cohort 2 is dependent only on their wealth, whereas in age 3, the decision is dependent only on preferences. In contrast, members of the age 4 cohort always sell their home, independently of their wealth or their preferences. This structure keeps the model analytically tractable but also makes it extremely difficult, if not impossible, to calibrate the model. In order to use real world data, the number of generations must be extended tremendously. If the basic model structure remains unaltered, the problem is to determine the age at which the discrete change in the decision problems occurs.

Nevertheless, it would be interesting to construct a model which is amenable to calibration. Our work in chapter four can be seen as an intuitive foundation and a benchmark in order to study the qualitative effects in such a model. However, taking the model of Ortalo-Magné and Rady (2006) as a starting point, it would be necessary to replicate the basic results using a different model structure. The crucial point is to construct the model such that for each age cohort wealth, personal preferences, and the survival probability enter the decision problem continuously. When households grow older the structure of the decision problem must stay the same, but the relative weights on the different housing and wealth decision motives change accordingly. We consider this as a challenge left for future research.

# Bibliography


[1] Aitken, A., Grimes A. and Kerr, S.: *House Price Efficiency: Expectations, Sales, Symmetry*, Motu Working Paper 04-02, May (2004)

[2] Anas, A.: *Rent Control with Matching Economies: A Model of European Housing Market Regulation*, Journal of Real Estate Finance and Economics, 15 (1997), 111-137

[3] Anas, A., Arnott, R. and Small, K.: *Urban Spatial Structure*, Working Paper, University of California Transportation Center, (1997)

[4] Andrew, A. and Meen, G.: *Housing transactions and the changing decisions of young households in Britain: The microeconomic evidence*, Real Estate Economics, 31 (2003), 117-138

[5] Arnott, R.: *Unpriced transport congestion*, Journal of Economic Theory, 21 (1979), 294-316

[6] Arnott, R.: *Economic Theory and Housing*, in: Handbook of Regional and Urban Economics, Vol. 2 (1987), E. S. Mills (ed.), Elsevier Science Pub., 959-988

[7] Arnott, R.: *Housing Vacancies, Thin Markets, and Idiosyncratic Tastes*, Journal of Real Estate Finance and Economics, 2 (1989), 5-30

[8] Arnott, R.: *Time for Revisionism on Rent Control*, Journal of Economic Perspectives, 9 (1995), 99-120

[9] Arnott, R.: *Congestion Tolling and Urban Spatial Structure*, Journal of Regional Science, 38 (1998), 495-504

[10] Arnott, R., Rave, T. and Schöb, R.: *Alleviating Urban Traffic Congestion*, Cesifo Book Series, MIT Press (2005)

[11] Arnott, R. and Yan, A.: *The two-modeproblem: Second best pricing and capacity*, Review of Urban and Regional Development Studies, 12 (2000), 170-199

[12] Bajic, V.: *The Effects of a New Subway Line on Housing Prices in Metropolitan Toronto*, Urban Studies, 20 (1983), 147-158

[13] Bailey, N.: *Deregulated Private Renting: a Decade of Change in Scotland*, Netherlands Journal of Housing and the Built Environment, 14 (2000), 363389

[14] Basu, K. and Emerson, P.: *Efficiency Pricing, Tenancy Rent Control and Monopolistic Landlords*, Economica, 70 (2003), 223-232

[15] de Borger, B., Proost, S., van Dender, K.: *Congestion and tax competition in a parallel network*, European Economic Review, 49 (2005), 2013-2040

[16] Braess, D.: *Über ein Paradoxon aus der Verkehrsplanung*, Unternehmensforschung, 12 (1968), 258-268

[17] Cho, M.: *House price dynamics: A survey of theoretical and empirical issues*, Journal of Housing Research, 7 (1996), 145-172

[18] Capozza, D., Hendershott, P., Mack C. and Mayer, C.: *Determinants of Real House Price Dynamics*, NBER Working Paper No. 9262, (2002)

[19] Deng, Y. , Gabriel, S. and Nothaft, F.: *Duration of Residence in the Rental Housing Market*, USC FBE Working Paper No. 02-3, (2002)

[20] Fujiwara, A., Yamane, K. and Zhang, J.: *Analysis of Travel Behavior Array Pattern From the Perspetive of Transporation Policies*, Journal of the Eastern Asian Society for Transporation Studies, 6 (2005), 91-106

[21] Genesove, D. and Mayer, C.: *Loss aversion and seller behavior: Evidence from the housing market*, Quarterly Journal of Economics, 116 (2001), 255-269

[22] Gleave, S.: *The Case for Rail: Final Report*, prepared for: Transport 2000, Association of Train Operating Companies, RIA, The Railway Forum, RPC and PTEG, (2002)

[23] Guiso, L., Jappelli, T., and Pistaferri L.: *An Empirical Analysis of Earnings and Employment Risk*, Journal of Business and Economic Statistics, 20 (2002), 241-253

[24] Heer, B. and Maussner, A.: *Dynamic General Equilibrium Modelling*, Springer Verlag, (2005)

[25] Henneberry J.: *Transport Investment and House Prices*, Journal of Property Valuation and Investment, 16 (1998), 144-158.

[26] Hubert, F.: *Contracting with costly tenants*, Regional Sciene and Urban Economics, 25 (1995), 631-654

[27] Kanemoto, Y.: *Design of Public Procurement System*, Government Auditing Review, 6 (1999), 3-20

[28] Kanemoto, Y.: *Optimum, market and second-best land use patterns in a von Thunen city with congestion*, Regional Sciene and Urban Economics, 61 (1976), 23-32

[29] Krainer, J., Spiegel, M. and Yamori, N.: *Asset Price Declines and Real Estate Market Illiquidity: Evidence from Japanese Land Values*, Working Paper, Federal Reserve Bank of San Francisco (2005)

[30] Lamont, O. and Stein, J.: *Leverage and house-price dynamics in U.S. Cities*, RAND Journal of Economics, 30 (1999), 498-514

[31] Lind, H.: *Rent Regulation: A Conceptual And Comparative Analysis*, European Journal of Housing Policy, 1 (2001), 41-57

[32] Malpezzi, S.: *A simple error correction model of house prices*, Journal of Housing Economics, 8 (1999), 27-62

[33] Mirlees, J.: *The Optimum Town*, Swedish Journal of Economics, 74 (1972), 114-135

[34] Olsen, E.: *A competitive theory of the housing market*, American Economic Review, 59 (1969), 612-622

[35] Ong, P.: *Auto Insurance Red Lining in the Inner City*, The Access Almanac, 25 (2004), 40-41

[36] Ortalo-Magné F., and Rady, S.: *Boom in, bust out: Young households and the housing price cycle*, European Economic Review, 43 (1999), 755-766

[37] Ortalo-Magné F., and Rady, S.: *Housing transactions and macroeconomic fluctuations: A case study of England and Wales*, Journal of Housing Economics, 13 (2004), 288-304

[38] Ortalo-Magné F., and Rady, S.: *Housing Market Dynamics: On the Contribution of Income Shocks and Credit Constraints*, Review of Economic Studies, 73 (2006), 459-485

[39] Read, C.: *Advertising and Natural Vacancies in Rental Housing Markets*, Real Estate Economics, 16 (1988), 354-363

[40] Read, C.: *A Price Dispersion Equilibrium in a Spatially Differentiated Housing Market with Search Costs*, American Real Estate and Urban Economics Association Journal, 19 (1991), 532-547

[41] Read, C.: *Tenants' Search and Vacancies in the Rental Housing Markets*, Regional Science and Urban Economics, 23 (1993), 171-183

[42] Read, C.: *Vacancies and Rent Dispersion in a Stochastic Search Model with Generalized Tenant Demand*, Journal of Real Estate Finance and Economics, 15 (1997), 223-237

[43] Rouwendal, J. and Verhoef, E.: *Second-best pricing for imperfect substitutes in urban networks*, Discussion paper TI 2003-085/3, Tinbergen Institute, Amsterdam-Rotterdam, (2003)

[44] Schley, F.: *Urban Transport Strategy Review: Experiences from Germany and Zürich*, Deutsche Gesellschaft für technische Zusammenarbeit, (2004)

[45] Solow, R.: *Congestion, density and the use of land in transportation*, Swedish Journal of Economics, 74 (1972), 161-173

[46] Solow, R.: *Land use in a long narrow city*, Journal of Economic Theory, 3 (1971), 430-437

[47] Stein, J.: *Prices and trading volume in the housing market: A model with down-payment effects*, Quarterly Journal of Economics, 110 (1995), 379-406

[48] Stiglitz, J.: *Alternative Theories of Wage Determination and Unemployment in LDC'S: The Labor Turnover Model*, Quarterly Journal of Economics, 88 (1974), 194-227

[49] Stiglitz, J. and Weiss, A.: *Credit Rationing in Markets with Imperfect Information*, American Economic Review, 71 (1981), 393-410

[50] Strotz, R.: *Urban Transportation Parables*, in: The Public Economy of Urban Communities, ed. by Julius Margolis, Whasinton D.C., Ressources for the Future (1965), 127-169

[51] Sutton, G.: *Explaining changes in house prices*, BIS Quarterly Review, September (2002), 46-55

[52] Verhoef, E.: *The demand curve under road pricing and the problem of political feasibility: a comment*, Transportation Research 29 (1995), 459-465.

[53] Verhoef, E.: *Second-best congestion pricing schemes in the monocentric city*, Journal of Urban Economics, 58 (2005), 367-388

[54] Vickrey, W.: *Congestion theory and transport investment*, American Economic Review Proceedings, 59 (1969), 251-260

[55] Wardrop, J.: *Some theoretical aspects of road traffic research*, Proceeding of the Insitute of Civil Engeneers, Part II (1952), 325-378

[56] Wasmer, E.: *Housing market discrimination, housing regulation and intermediaries*, Discussion Paper, prepared for the Annual Meeting of the American Economic Association on January 7, 2005, Philadephia

[57] Wheaton, W.: *Price-induced distortions in urban highway investment*, Bell Journal of Economics, 9 (1978), 622-632

[58] Wheaton, W.: *Vacancy, Search, and Prices in a Housing Market Matching Model*, Journal of Political Economy 61 (1990), 1270-1292

[59] Wilson, J.: *Optimal road capacity in case of unpriced congestion*, Journal of Urban Economics, 13 (1983), 337-357

**Thomas Maier**

**Academic CV**

| | |
|---|---|
| 1996 | Abitur, Ignaz-Taschner-Gymnasium |
| 1999 | Vordiplom Mathematics, University of Munich |
| 2002 | Diplom Mathematics, University of Munich |
| Since October 2002 | Doctoral student and member of the MGSE |
| Summer 2004 | Research Stay at the University of Cambridge, U.K. |
| 2005-2007 | Research Assistant at the Seminar for Dynamic Modelling, Munich |