# Housing density and rising prices

The impact of real estate and land prices upon the density of urban dwellings in the housing market of Utrecht

Gerjan Heijkoop - 2749769
Supervisor: Jan Rouwendal

August 10, 2022

**Abstract**

*This study contributes to the literature on urban economics. It provides evidence about floor space used per hectare over time. A system-generalised method of moments (sys-GMM) technique is used on a dataset of 9922 grid cells of 100 x 100 meter from the city of Utrecht. Major findings are: existing housing has modest adjustment to an optimal floor space and newly developed sites are built close to the optimal floor area per hectare. Next to this, newly built hectares contain more housing units compared to earlier developed sites.*

# Content

# 1. Introduction

Recent decade showed an extraordinary increase in real estate price levels. Due to this development, also (urban) land prices were rising. Extremely low interest rates and economic prosperity resulted in all time high prices within real estate (related) markets. Small countries with large population densities, like the Netherlands, experienced large tension in their housing market because of supply inelasticity. This is a consequence of strict zoning regulation. However, empirical evidence about the consequences for the size of urban dwellings and newly built apartments are still few. According to housing market models, urban structures change by different factors: the development of prices, demographics and economic activities. This thesis takes the Muth housing market model as a starting point to investigate the development of dwelling size and housing price over time.

Several housing market models take into account the land market, the Muth model is the best-known. Most (empirical) housing market studies are about the demand side of housing, however the supply side is investigated less. It is of large importance to study land market dynamics or land as an intermediate input, in order to know how land markets affect both the supply of new housing and the housing market. Fuller et al. (2020) show that housing prices and markets strongly influence wealth and income distributions. They mention that contemporary capitalism causes housing capital and value of land to become increasingly dominant. Therefore, they propose more empirical work to study the inequality of intergenerational wealth and its increase in political sensitivity. David Ricardo already feared that a population growth would result in rent-seeking behavior among landowning elites claiming larger shares of national income. Recent studies show that this phenomenon might be of larger impact nowadays. Rural land parcels are subject to speculation and incentives not to sell land, therefore they can be seen as having a certain option value (Buitelaar, 2021; Van der Krabben, 2021). It can be seen as an ongoing process of landowners who have an incentive to wait for higher prices, leading to lower housing supply, followed by further waiting or delay to sell land. Dutch authorities and housing construction companies are struggling with the development and profitability of low and middle segment type of dwellings. The soaring prices led to lower amounts of affordable houses for first-time home-buyers, also because the profitability of these types of dwellings is lower compared to more expensive detached houses (Doodeman, 2021). The dataset used for this research contains information of 100 by 100 grid cells of the city of Utrecht. Aim is to obtain insights about the dynamics of housing stock and its division over space and over time.

# 2. Theoretical framework

A framework of literature and characteristics of the studied topic is made to clarify the background of the research. This framework consists of empirical work regarding the structure and dynamics of land and housing markets and literature of the Dutch housing and land market. Firstly, the characteristics of housing markets are described. Secondly, the empirical work regarding land and housing markets are reported. Thirdly, the Dutch housing market and its institutions, shaping its background, are described.

## 3.1 Characteristics of housing markets

In many economies, housing markets and their connected credit markets play an important role in the macro economy (Muellbauer & Murphy, 2008). The demand side of housing is driven by factors like: credit availability, income, interest rates, demography and expected appreciation. The supply side is driven by factors like: spatial planning structure, tax systems and the structure of local governments. According to Muellbauer & Murphy housing wealth plays “a potentially very important role in macroeconomic fluctuations and the distribution of its welfare”. Housing markets differ largely from other markets in their dynamics and characteristics, despite their connection within the macroeconomic context. Durability, heterogeneity and restrictions to objects and land impact the degree of market clearance and price adjustments. These characteristics make that the housing market can be considered as a market with an adjusting stock of housing. The existing housing stock can be adjusted or improved, this makes the adjustment of the stock downward rigid (Boelhouwer, 2020). In other words, adjustment is not fully met by the supply of new housing. The adjustment towards a long run equilibrium takes time. This has several origins. Costs involved with stock adjustment are smoothed over time by construction industries. Doing so, their adjustments are not made at once. Next to that, negotiations and strict planning regulations imposed by governments makes housing supplies come with delay. These characteristics shape the dynamics of housing markets. Besides that, housing markets are connected to financial markets. Housing finance systems and institutions interact, thereby impacting the functioning of the housing markets. Also monetary policy has large influence on asset prices and their volatility (Voigtländer, 2014). Vermeulen & Rouwendal (2007) argue that as a consequence of the zoning system, the supply of residential land is more to be seen as a governmental decision rather than an outcome of market supply.

## 3.2 Empirical work regarding housing and land supply

Regarding housing supply, major empirical work is based upon the Muth-Mills housing model, the monocentric city model (Brueckner, 1987). This stylized static model, also influenced by Alonso, assumes a city with a fixed population and income level around a central business district. The city center is assumed to be zoned for work. Households prefer to live closer to the city center, because they commute. The price and density adjust to clear the market for housing. This structure implies that the center has higher and more densely built areas. The model simplifies reality in several ways. It assumes all households to have the same income, preferences and it assumes that houses are similar instead of heterogeneous. Next to this, the city is assumed to be monocentric, commuters are assumed to travel only to the city center. Thirdly, factors like interest rates or wages are considered to be strictly exogenous. One aspect worth studying regarding the housing and land market can be derived from Muth’s condition of housing production. Muth’s condition stems from the monocentric city model (MCM) developed by Muth and Mills. Muth’s condition states that the housing supply (h) is a function of capital and land (k and a). With capital having a decreasing marginal productivity the supply function will be concave. Muth derives the marginal product of capital and land regarding housing supply. If land is more expensive, more capital will be used to develop one housing unit. Vice versa, if land is cheap, more land will be used in the development of one housing unit.

While many papers are written about the demand of housing, DiPasquale (1999) mentions that the empirical evidence on the supply of housing is far less convincing. Evidence of a causal relation between housing prices and output of new housing is very hard to find. Supply seems to be (almost) perfectly inelastic. Empirical work on housing supply is even more scarce outside the US. Vermeulen & Rouwendal (2007) estimate that the short-run housing supply in the Netherlands is almost fully inelastic. Structural analyses of housing supply consider either residential investment or new construction in units. It appears that the market for housing is far from efficient. Most models consider the adjustments of stock to jump to its long run equilibrium in steps. DiPasquale & Wheaton (1994) argue that a slow clearance of the housing market can be related to different search frictions. Their results strongly suggest that models with price levels work better when lagged stock or other measures of stock disequilibrium are included. This may be a reason for underestimating the responsiveness of new construction to market conditions.

Regarding land supply, the empirical evidence about the relationship between urban land price and parcel size is very strong. Colwell & Munneke (1999) show that the relationship between total price of land and size of land is first a convex function altering to an increasing concave function. Ahlfeldt & McMillen (2018) investigated the relationship between land values and skyscrapers in Chicago. Their results suggest that there is a "supply-side mechanism that promotes the typical land-use segregation observed within cities". The cost of building tall buildings increases by an exponential rate, they argue that the affordability of housing will not be reached by the development of tall buildings. McDonald (1981) reviewed the elasticity of capital-land substitution in urban housing. All of the studies in that paper show that the intensity of land use (for newly constructed housing) is determined by the value of land.

## 3.3 The Dutch housing market and its institutions

The Netherlands has a more restrictive land use policy compared to its neighboring countries, resulting in one of the lowest supply elasticities in the Western countries (Boelhouwer & Hoekstra, 2009). There are different possible explanations for this. The Dutch Housing Act (Woningwet) of 1901 was the foundation of its current strict land use regulation. Its initial purpose was to provide clean water and sewerage. Nowadays, the country is densely populated resulting in even more strict policy towards spatial planning. Many argue that the Dutch way of organizing adds to a large stagnation in housing production. The Dutch urban land market is more or less separated from the agricultural land market. Almost the entire cultivated area of the Netherlands has been organized publicly when it is developed (Needham et al., 2018). The providence of residential land has been very limited in the past decades, or it has been taxed implicitly, or permissions were only granted conditionally (Vermeulen & Rouwendal, 2007). This background shows that the Dutch supply of housing can be regarded to be driven by the responsiveness of the institutional setting, rather than by the 'free market driven' reactiveness of construction companies. The restrictiveness can be seen as having a twofold impact upon housing supply. First, the amounts of space for to be developed sites are more scarce compared to an unregulated market setting. Second, the provision of housing in existing areas is bounded to the regulations and decision made in the past. An old monumental canal house does not adjust in size and surrounding sites are often restricted to be built with similar height.

Several fiscal and institutional factors impact the demand of housing. 95 percent of the rented homes in the Netherlands are subjected to rent regulation, this makes that being 'inside' the social renting system is very attractive. Social housing tenants are protected when renting below the deregulation limit of € 763,47 with certain regulation. The government determines the height of these rents and social renting is only applicable below a certain income level. Next to this, owner occupiers are subsidized. The Dutch mortgage interest tax relief system allows homeowners to deduct part of their mortgage interest payments from their income tax. In combination with a national mortgage guarantee, this led to one of the largest amount of mortgage debt per capita. In 2003 it peaked to an average loan-to-value of 112 per cent, far beyond the European average. Most European countries are below 100 per cent. It can be concluded that the Dutch housing market is far from competitive. The inelastic housing supply makes that the Dutch housing market is characterized as a demand-oriented market driven by factors like income, mortgage interest rates, rent levels, inflation, GDP, unemployment and population growth. Boelhouwer & Hoekstra concluded in 2006 that the Dutch housing market policy in combination with its spatial planning policy is inconsistent. Demand is stimulated and regulated, whereas the supply is subjected to very tight regulations. Boelhouwer & Hoekstra (2009) mentioned these problematic issues in the Dutch housing market in 2009 which remained roughly unsolved until present. They highlighted the disadvantage of the generous tax regimes for mortgage interest deduction and plead for more tenure neutrality. The combination of this demand stimulation (lack of tenure neutrality) and restrictive characteristics of spatial planning created a gap between tenure types. It is argued that the gap would grow as a consequence of maintaining fiscal relief for homeowners, regulated rents (resulting in waiting lists) and restrictive spatial planning.

The causes of the current Dutch housing shortage can be argued from various factors. Low interest rates resulting in easier access to capital is obviously a very important factor of rising house prices. According to Groenemeijer (2021) the main cause of the housing shortage is that the Dutch administration responded too late to a higher population growth. Governmental organizations assumed the size of the Dutch population would fall in the foreseeable future. In the period 2010-2013 the stop on large-scale housing plans was a significant consequence of this view. It is important to keep in mind that demographic forecasts are always surrounded by uncertainties. Simplification of procedures must contribute to a faster supply. However, time-saving opportunities are limited, since the interests of local residents, nature, economy, archaeology, etc. are taken into account. A buffer of well specified planning capacity to absorb fluctuations seems like a better strategy. Currently, the government set its goal to provide for 130% planning capacity in tense housing market areas.

The Dutch housing market not only has challenges towards the supply of housing. Next to that, big challenges towards sustainability have to be overcome. By 2030 every Dutch house should be gas-free. Socio-economic trends and innovation steer the housing sector making an energy transition. It is important to know what kind of (future) buildings are to be made sustainable. Economic developments, demographical developments and preferences influence the type of housing. In this transition a large focus upon detached single family homes might be misleading (Ebrahimigharehbaghi et al., 2019). It is clear that the owner occupier sector has a large share in energy usage. Ebrahimigharehbaghi et al. write about the importance of focus upon specific household types for renovating buildings. It is wise to focus more upon techniques for sustainability improvements for densely developed city when the future developments are heading towards such an environment. Next to this, it is important to take in mind that development at a certain point in time comes with an option value. Building now means future developments have to adjust to the realized environment. This might come with certain costs.

## 3.4 The characteristics of Utrecht

Utrecht is one of the oldest cities in the Netherlands. Before the 17$^{th}$ century it was the biggest city of the country. The Utrecht region is one of the most competitive economic regions in Europe. Next to this, the city is the road and rail junction of the Netherlands. Partly for this reason, many companies and institutions have their headquarters there. By 2035, Utrecht will have more than 20 percent more inhabitants, according to the Planbureau voor de leefomgeving (PBL, 2019). This makes Utrecht the fastest growing city of the four largest cities in the Netherlands. Utrecht has a historic center that is completely surrounded by a canal. The well-known Oudegracht and the Nieuwegracht are paved from south to north. The city's population grew from 256.404 to 361.742 people. The external validity of this research is assumed to be good. Utrecht can be regarded as a 'standard' Dutch city, results may differ when more rural parts will be part of the analysis.

## 3.5 Knowledge gap

The knowledge gap in this research is to know to what extent parcel sizes are impacted by a surge in land or home prices. Using the monocentric city model (MCM) it can be argued that sizes of land used for the development of housing units are adjusted when the price of land rises. This means constructing companies shift towards a larger share of capital as input in one unit of housing. The size of a parcel will shrink when the cost of land rises. In practice, housing price levels mainly determine the residual value of land below the realized building. This is also implicated by the theory of Muth.

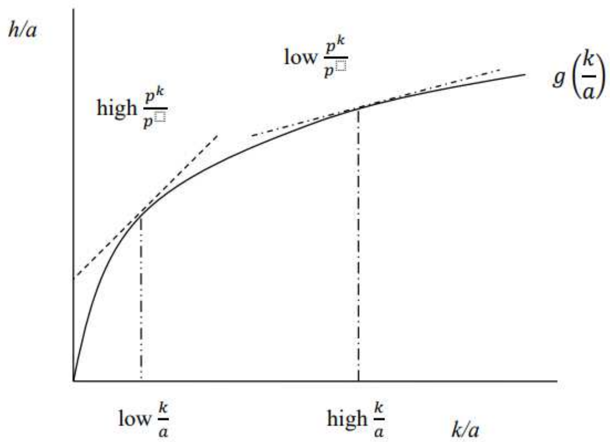


*Figure 1: (Rouwendal, 2019)*

$P^k$= price of capital
$P^{\square}$= price of land
k = the amount capital per unit of land
a = the amount of land used
h = one unit of housing

The marginal productivity of capital decreases if land prices increase, leading to the concave shape of the curve. A lower pk/pa will be attained for newly developed buildings. In other words, more capital is used per developed unit instead of land, which is a consequence of land being more expensive. This means a shift towards the right on the concave function of g (k/a). This research investigates how fast adjustment in price or land size takes place. Several characteristics of housing markets come into play when applying this theory in practice. The optimal floor space of a certain unit newly developed will be attained sooner than with earlier developed dwellings. Restrictions and delays resulting from spatial planning come into play, meaning only part of the floor space optimum will be reached.

This research examines to what extent and with what pace the gap in optimal floor space is filled over time. It is an empirical study investigating whether this theory holds within the Dutch context of the city Utrecht in the period of 2012 until 2022. Aim is to gain insight into the impact of housing and land prices upon the developed type of dwelling or the amount of land/capital used. Building the right type of dwellings is a point of concern in the debate about the housing crisis. Next to this, the Dutch building environment has future challenges with respect to nature inclusiveness and biodiverse residential areas (van Haaster-de Winter et al., 2020). This challenge might be impacted by developments in prices and sizes.

# 3. Descriptive data and methodology

## 3.1 Descriptive data and interpretation

The data used for the analysis in this thesis is a dataset containing information about parcels with the size of one hectare. Each parcel is indicated in which neighborhood it is situated (gemeente_code, stadsdeel_naam, wijk_code and pc4). For every parcel the amount of dwellings, the living space, the amount of buildings, the building year, the amount of dwellings per period of ten years and the ratio of surface that can be built upon is available. The dataset is very rich, containing 9,922 grid cells of 100x100. A table containing detailed descriptive statistics is in the appendix. The average number of houses per grid cell in 2012 was 13.6. In 2022, this amount grew to 16.2 houses per grid cell. This is an increase of 19.7 %.

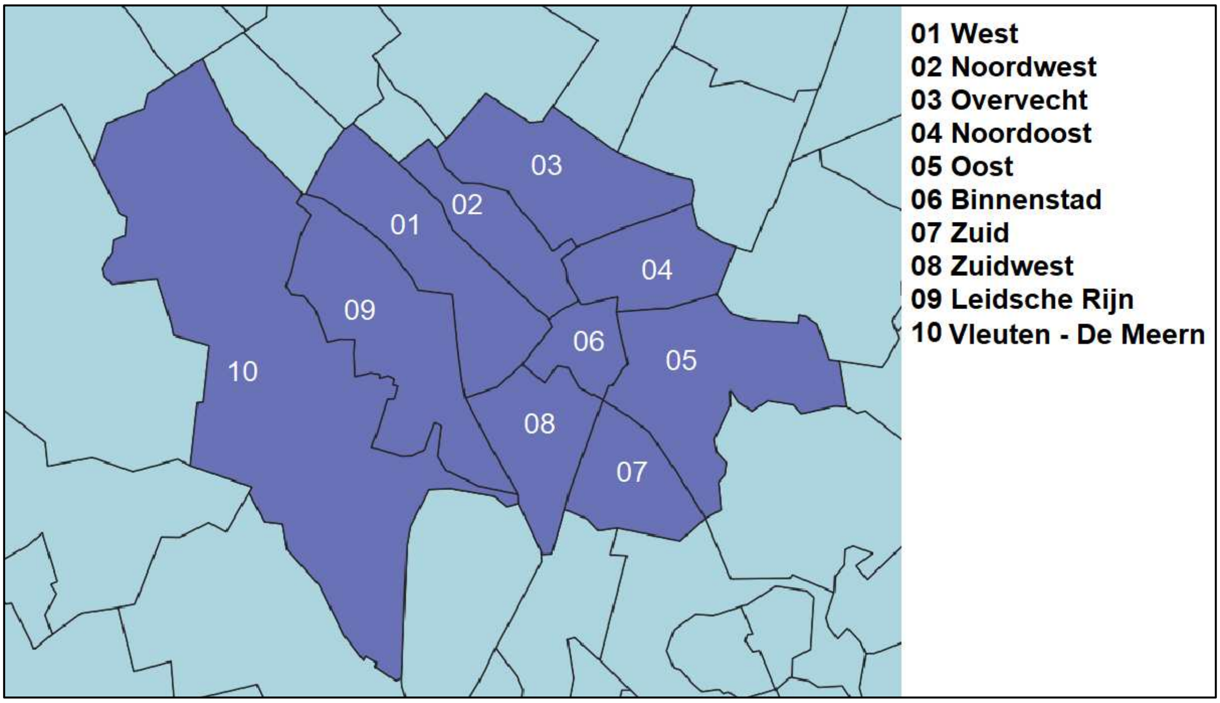


*Figure 2: Neighborhoods*
*Source: Authors' computations*

Three of the areas (wijken) have an average number of houses per grid cell of 30 or higher; these are the city center (06), the north-east area (04) and the north-west area. 3,334 grid cells contained at least one house in 2012. 5,729 grid cells contain no home throughout the period 2012-2022. 874 grid cells contain zero houses in 2012 and at least one house in 2022. These grid cells contain an average count of houses of 19.8 in 2022. Meaning that empty grid cells in 2012, but containing houses in 2022 got an average of 19.8 new houses. 3,334 grid cells already had houses in 2012. The average number of houses in 2012 of these (already built) grid cells was 40.4 houses in 2022 this increased to 42.9 houses, meaning a growth in houses of 6.4% is realized in these areas. The average floor space per house declined every year from 381.3 square meter in 2013 to 177.9 in 2022. This already indicates that the surging housing prices might have impacted the amount of dwellings on a grid cell.

For the regression analysis the data is cleaned by dropping plots with no housing in 2022 and leaving out certain industrial sites. Also grid cells with an average of 900 square meters or larger of living space were dropped. This is to get rid of outliers with extraordinarily high values of square meters. These grid cells might contain greenhouses or large barns belonging to farms etc. The graph below shows the average surface of living space per house by year for all areas of Utrecht together from 2013 on. 2012 (first year of measuring) had a very low value for living space, therefore that year is left out. Many data cells from 2012 have a value for houses of zero or lower than subsequent years, while they report a certain amount of surface for that grid. That is why this year is suspected to have a systematic measurement error. This must be due to a change in the registration of data by the Central Bureau for Statistics (CBS). The so-called 'BAG registration' changed in that year. The average surface of living is expected to decline over time as was mentioned in the theoretical part. Surprisingly, the pattern shows that the average amount of living space for Utrecht as a whole is not declining. Some neighborhoods (wijken) have declining amounts of spaces if they are taken separately. The inner city (Binnenstad) shows declining amounts of space over the years. That area appears to have declined from 118 square meter to just below 112 square meter.


*Figure 3: Average floor space per house (in m²) by neighborhood (wijk)*
*Source: Authors' computations*

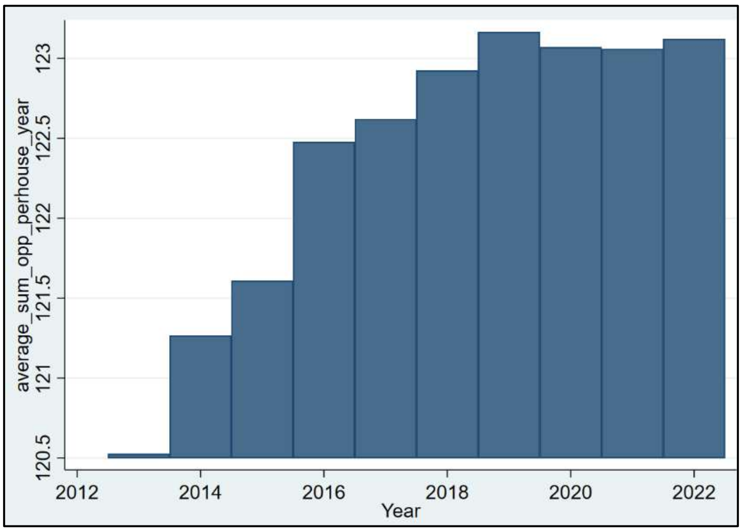


*Figure 4: Average floor space per house (in m²) in Utrecht*
*Source: Authors' computations*

The figures below show the number of houses over time per neighborhood (wijk). The areas IJsselstein and Bunnik (last and first ones) appear to have no data about amounts of houses. It is very clear that the area called Leidsche Rijn (WK034409) is a so-called 'VINEX' location (green field development). It has nearly zero houses in 2001, but ends at an amount of 17,000 in 2022. Also Vleuten-De Meern (WK34410) has a sharp growth in housing numbers. The patterns do not show very uncommon patterns although Noordoost (WK034404) and Oost (WK034405) have a kind of 'block' starting in 2013 and stopping in 2017. This might be due to a structural difference in measurement during that period. Combining figure 3, 4, 5, 6, 7 and 8 and taking into account the background of an old, spatially restricted inner city leads to several insights about Utrecht's urban structure. The newer (outer) parts of Utrecht like Vleuten-De Meern, Leidsche Rijn and Zuid-West grew rapidly during the past twenty years. Thereby, newly built houses were likely to be larger than existing houses (see figure 3), possible reason for this might be the fact that gardens tend to be smaller compared to already built grids. The old inner

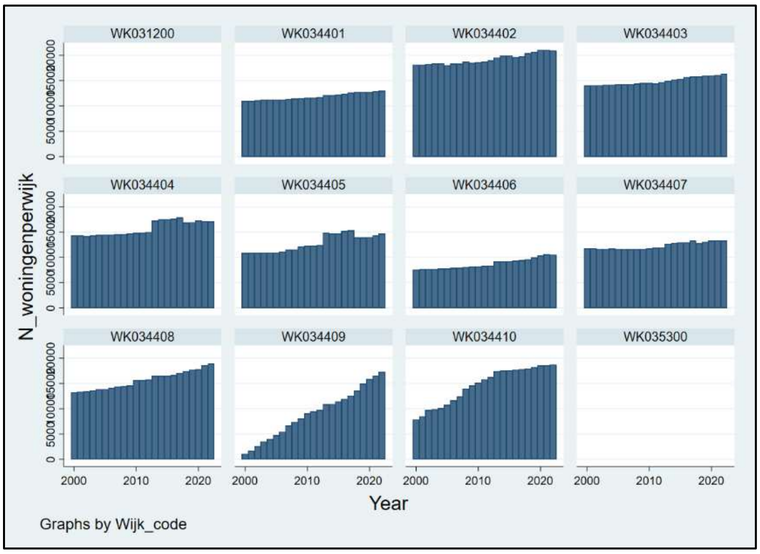


Figure 5: Number of houses by neighborhood
Source: Authors' computations

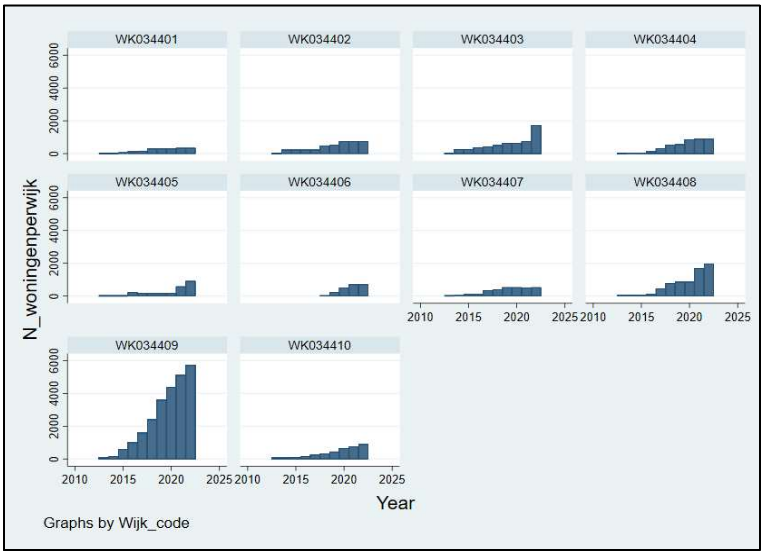


Figure 6: Amount of 'newly built housing' by neighborhood
Source: Authors' computations

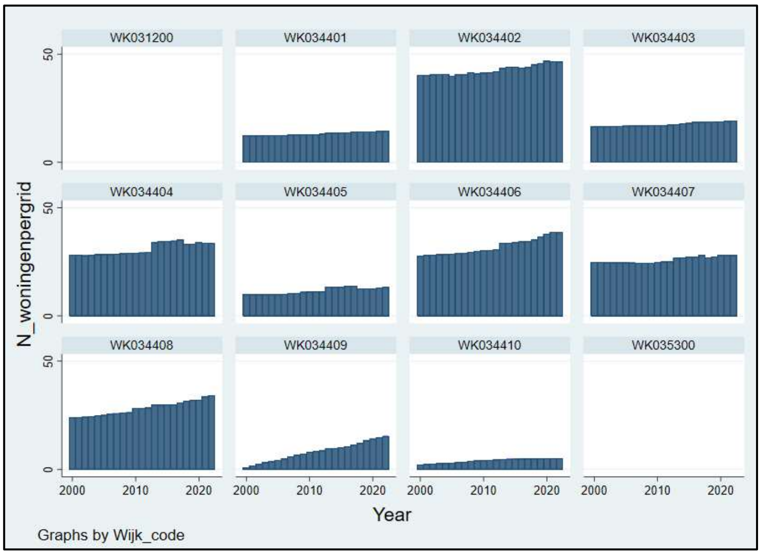


Figure 7: Houses per 100x100 meter (average)
Source: Authors' computations

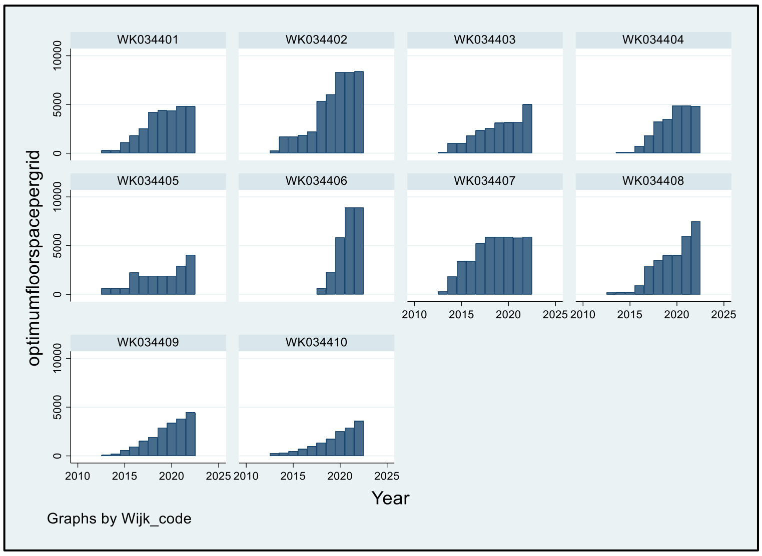


Figure 8: Floor space per 100x100 meter newly built grids
Source: Authors' computations

city grew much slower in the amount of houses, approximately from 8,000 to 9,500. Measuring the average floor space per house without Vleuten-De Meern leads to a decreasing pattern over time as of 2019. Leaving out the three 'richest' neighborhoods Oost, Leidsche Rijn and Vleuten-De Meern (WK034405, WK034409 and WK034410) with the largest amounts of floor space leads to a sharp decrease in floor space in 2022. It can be said that the newer outskirts have larger houses on average, though over time the sizes of newly built houses decline. That might be due to preferences of economic prosperity. These newer houses still contribute to a higher average for Utrecht as a whole. This might also be an explanation for the declining growth in figure 4. Figure 6 provides an overview of grown sites. It counts cells with less than 10 houses in 2011, 2012 or 2013, but more than 15 in 2022. Taking these cells prevents grown city center cells to drop out of the measurement. Leidsche Rijn is the neighborhood with the largest expansion. The average floor space per house within these years was around 130 square meters. Figure 7 indicates the densification pattern of the city. It is very identical to figure 5. The graph shows the average number of houses growth path for every neighborhood in Utrecht. This is an indication of a more densely built city. The neighborhoods with very low values still have a lot of green fields in their territory. Vleuten-De Meern (the last one) is the best example of this.

Figure 6 and 8 show by their steady growth path that the last two graphs contain green field development (mainly detached houses) in the outskirts Vleuten-De Meern and Leidsche Rijn.

To detect an optimum floor space over years it is possible to make estimations of floor spaces reached by newly developed sites in certain time spans. The graphs on the next page (figure 9 to figure 12) show floor spaces of newly developed sites within different time spans. These graphs were made by selecting grid cells with less than 10 houses 5 years before they had at least 15 houses. For example, figure 9 has grid cells containing less than 10 houses in 2001 or 2002. It appeared that housing in 2012 had a larger average floor space per grid cell than 2007. 2017 also had a higher floor space, over 5000 square meters per grid cell. 2022 will most likely have more, although the pattern shows developments might be still under construction. These statistical facts suggest that optimal floor space per grid cell grows throughout a period of 5 years. In a period of ten years the optimum seems to grow by an extra 900 square meters per 100 x 100 grid cell. Figure 13 shows that the average floor space per house of the newly developed grids declined by approximately three meters throughout a decade. Several things can be concluded from these statistical facts. They suggest that construction companies use more space from a grid to allocate to housing (instead of other usage). Next to that, it is surprising to see that the floor space per house of Utrecht as a whole (figure 4) grew, while the floor space of newly built dwellings declined (figure 13). Reasonable explanation for the gain in floor space is the fact that homeowners often reconstruct part of their home. According to Nieuwbouwwijzer (2020) the amount of mortgage applications used for reconstructions is as large as mortgage applications for newly built housing. This number might be even larger when reconstruction without a mortgage is taken into account. Whereas newly built housing requires mortgages more often.

The graph of figure 14 shows the most important variables graphically, floor area and the average of the price indices. Both floor area per hectare and price indices rose to throughout the period of research. This fits into the theory in the way that more floor area is realized on a hectare. In other words, land is used more intensively during the period of price hikes.

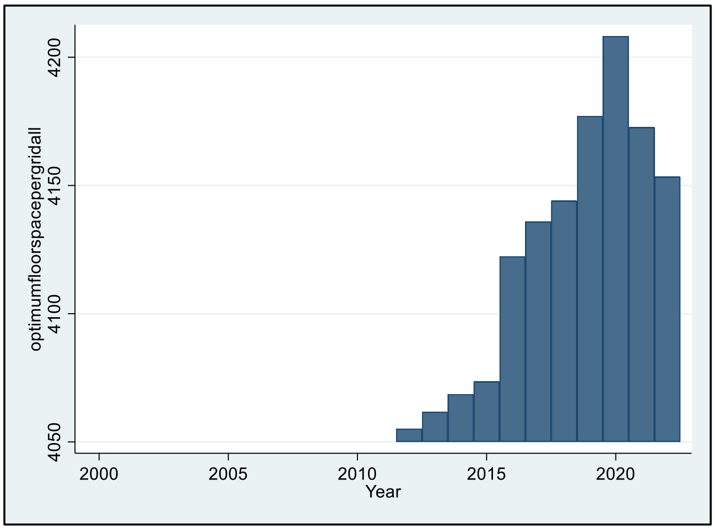


*Figure 9: Amount of floor space in square meters per grid cell. Building started five years before or in 2007*
*Source: Authors' computations*

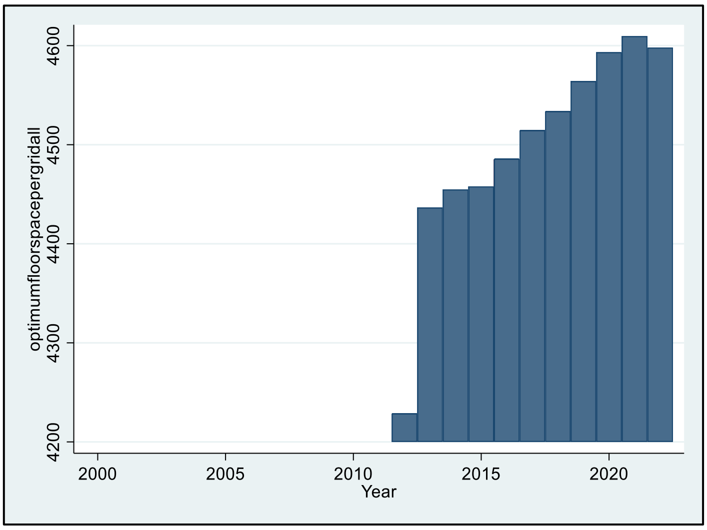


*Figure 10: Amount of floor space in square meters per grid cell. Building started five years before or in 2012*
*Source: Authors' computations*

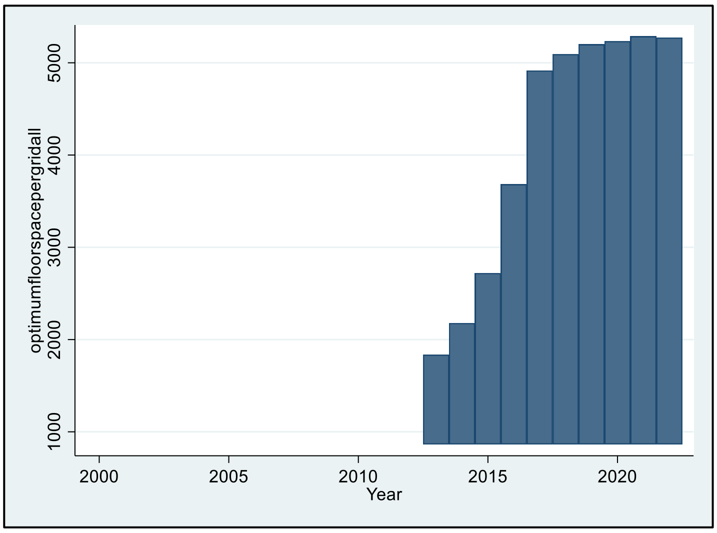


*Figure 11: Amount of floor space in square meters per grid cell. Building started five years before or in 2017*
*Source: Authors' computations*

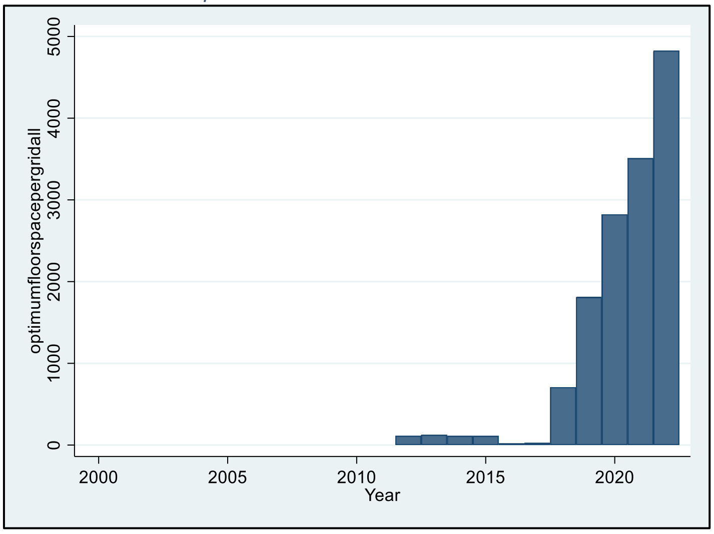


*Figure 12: Amount of floor space in square meters per grid cell. Building started five years before or in 2022*
*Source: Authors' computations*

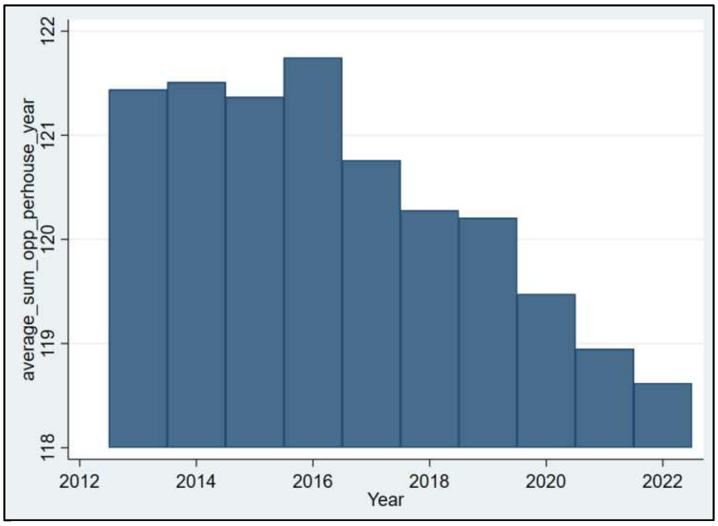


*Figure 13: Average floor space per house newly developed sites aggregated*
*Source: Authors' computations*

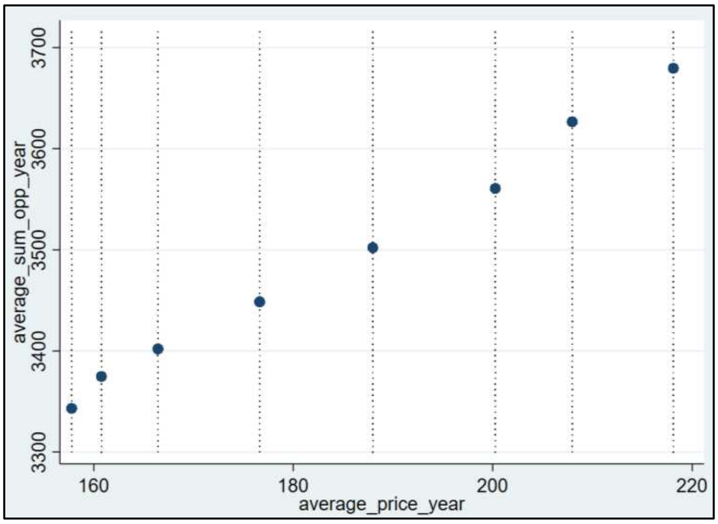


*Figure 14: Average floor space per grid cell Utrecht and price index (2013-2020, left to right)*
*Source: Authors' computations*

## 3.2 Empirical method

Rouwendal (2022) hypothesizes by using the outcomes of the Muth model of housing construction “that in each period a part of the gap between actual and optimal floor space is filled”. The optimal amount of floor space per housing unit in a grid cell is considered to be an increasing function of the local housing price index $PI_i$ and other variables $X_{i,t}$. The other variables are considered to reflect restrictions from spatial planning.

$$FA_{i,t}^{opt} = g\left(PI_{i,t},\ X_{i,t}\right) \tag{1}$$

The data available is used to estimate a regression equation derived by following equations. The gap in optimal floor area (FA) is assumed to be filled over each period.

$$\Delta \log FA_{i,t} = \gamma\left(\log FA_{i,t}^{opt} - \log FA_{i,t}\right) \tag{2}$$

$\gamma$ is the part of the gap being filled, the equation is equivalent to:

$$\frac{FA_{i,t+1}}{FA_{i,t}} = \left(\frac{FA_{i,t}^{opt}}{FA_{i,t}}\right)^{\gamma} \tag{2'}$$

Next, the optimal floor space is suggested to be a linear function of the house price and other variables:

$$\log FA_{i,t}^{opt} = \alpha' PI_{i,t} + \beta' X_{i,t}. \tag{3}$$

Substituting equation 3 into 2 and adding an error term $\varepsilon'_{i,t}$ then gives the estimating equation

$$\Delta \log FA_{i,t} = -\gamma \log FA_{i,t} + \alpha PI_{i,t} + \beta X_{i,t} + \varepsilon'_{i,t} \tag{4}$$

with $\alpha = \gamma\,\alpha'$ and $\beta = \gamma\,\beta'$.

To account for dependence between the errors terms of grid-cells over time we adopt a fixed effect approach, to get rid of the unobserved effect $\varphi_i$. $\varepsilon_{i,t}$ is assumed to be i.i.d. Rewriting the regression equation in order to deal with endogeneity leads to the next equation.

$$\left(\log FA_{i,t+1} - \log FA_{i,t}\right) = \beta_0 - \gamma \log FA_{i,t} + \alpha PI_{i,t} + \beta X_{i,t} + \varepsilon'_{i,t}$$

$$\log FA_{i,t} = \beta_0 + (-\gamma + 1) \log FA_{i,t-1} + \alpha PI_{i,t-1} + \beta X_{i,t-1} + \varphi_i + \varepsilon_{i,t} \tag{5}$$

A linear dynamic panel data approach is needed in order to deal with endogeneity. The Arellano-Bond method is a difference or system GMM-estimation approach that deals with this. These two methods are carried out in Stata. First, it is essential to get rid of the (individual) unobserved effect $\varphi_i$ by applying a method of first differencing. Secondly, to do a proper regression we have to deal with the problem of endogeneity. The lagged dependent variable $\log FA_{i,t-1}$ in the model is a cause for autoregressive paths by the error term. That means, $\log FA_{i,t-1}$ might be correlated to $\log FA_{i,t}$ causing endogeneity. After first differencing, the error term is uncorrelated with the explanatory variable $\log FA_{i,t-2}$. This lagged dependent is used as an instrument. The Arellano-Bond approach assumes the error term to be uncorrelated to the lagged dependent that is used as an instrument. That is why this research assumes todays floor space not to be correlated with future error terms.

$$\log FA_{i,t} - \log FA_{i,t-1} = \beta_0 + (-\gamma + 1)\left(\log FA_{i,t-1} - \log FA_{i,t-2}\right) + \alpha\left(PI_{i,t-1} - PI_{i,t-2}\right) + \beta\left(X_{i,t-1} - X_{i,t-2}\right) + \left(\varepsilon_{i,t} - \varepsilon_{i,t-1}\right) \tag{6}$$

On the right hand side the terms $\log FA_{i,t-1}$ and $\varepsilon_{i,t-1}$ are correlated, first differencing corrects for that. The Arellano-Bond approach uses earlier lags as instrumental variable. Hereby, assuming that past values are not correlated with future error terms. It is possible to do a difference or a system GMM estimation (Roodman, 2009). Applying system GMM requires an extra level equation to be estimated the original level equation (7). Earlier differences are used as instruments for the endogeneity stemming from $\varphi_i$. That is why the system GMM assumes that first differences of instrument variables are not correlated with the individual fixed effects, in other words it is assumed to be a random walk model.

$$\log FA_{i,t} = \beta_0 + (-\gamma + 1) \log FA_{i,t-1} + \alpha PI_{i,t-1} + \beta\, X_{i,t-1} + \varphi_i + \varepsilon_{i,t} \quad (7)$$

This offers the possibility to use more instruments which improves the efficiency. Blundell & Bond (1998) indicate that the use of difference GMM estimators yields inefficient and biased estimates of $(-\gamma + 1)$. This is for finite samples and when T is short.

Carrying out the Arellano-Bond approach the rule of thumb according to Bond et al. (2001, p. 7) is taken. This means both a pooled OLS and a fixed effects panel regression are done. This rule of thumb makes use of an estimation of the pooled OLS is used as an upper-bound estimate for the difference GMM estimate and the fixed effects estimate as a lower-bound estimate. When the difference GMM estimate obtained is close to or below the fixed effects estimate it suggests that the estimate is downward biased, because of weak instrumentation and a system GMM should be preferred instead. Next to that, tests for autoregressive patterns are carried out and the Hansen test of overidentifying restrictions is carried out. Following assumptions are made when carrying out the system GMM method, which is the used regression method in this research. First assumption: there is sequential exogeneity. This means that the dependent variable in the equation is unrelated to future errors. However, the process is allowed to be dynamic. That means: current dependent variables may be influenced by past ones. Second assumption: there is no autocorrelated error. This means the $\varepsilon_{i,t} - \varepsilon_{i,t-1}$ terms are not autocorrelated. For this assumption an Arellano-Bond test for autocorrelation exists. This tests looks after first and second order autocorrelation in the error terms. First order correlation is expected and should be no problem. Second order autocorrelation is a problem. A Hansen-Sargan test checks the exclusion of the instrumental variables. In the model some regressors may be endogenous. In this research this is the lagged dependent for floor space. The Arellano-Bond tests are for autocorrelation of the error term. An additional assumption for system GMM is that the instrumental variables are uncorrelated with the individual fixed effects. So for every observation the predetermined (independent) variables have constant correlation over time with $\varphi_i$ (Blundell & Bond, 1998; Bond et al., 2001; Kripfganz, 2019; Roodman, 2009).

Several options exist for carrying out a GMM regression. A one-step or a two-step regression variant is possible, Windmeijer-corrected standard errors, instruments small-sample adjustment and orthogonal deviations. The two-step regressions have modest efficiency gains, this research only uses two-step when results appear to be very different and efficiency is needed. Using Windmeijer correction for standard errors reduces the gap between two-step and one-step estimations. The robust option uses quietly calculated two-step GMM errors (in xtabond2). After a one-step system GMM estimation Hansen test is not asymptotically valid, so not valid for sizes going to infinity. The two-step outcomes are used within this research. Small-sample adjustment shows t-test instead of z-test statistics. The orthogonal deviations subtract the average of all future available observations of a variable, instead of subtracting the previous observation from the current one. This is used in this research to correct for gaps in observations, which are present due to sites built upon after some 'empty' years (Blundell & Bond, 1998; Bond et al., 2001; Kripfganz, 2019; Roodman, 2009).

# 4. Results and discussion

## 4.1 Result and discussion

Estimating by the Arellano and Bond approach is done by using different subsets of variables (reflecting local characteristics), taking into account the aforementioned tests for second order autocorrelation and Hansen-Sargan to look for exogeneity of instruments (Roodman, 2009). Thereby, it is expected by theory that $\gamma$ will appear to be between 0 and 1. Otherwise the theory does not hold in practice, or due to other reasons there is no adjustment factor.

Different results appear when running the Arellano-Bond approach with different subsets of variables. The first regression below is taken from the areas Overvecht, Noordoost, Leidsche Rijn and Vleuten-De Meern. These somewhat similar outskirts were the neighborhoods with largest (greenfield) housing developments. Therefore, they were expected to be more likely to adjust to an optimum. Next to that, it is expected that these parts of Utrecht will show the quickest adjustment. A second regression is taken from only newly built sites in Overvecht, Noordoost, Leidsche Rijn and Vleuten-De Meern. This regression is taken as a check, since the first regression did not appear to have proper results (according to the theory). $\gamma$ was higher than 1. Therefore the theory is not used to interpret the results from table 1 and 2. The second regression of table 1 and 2 is not suitable for obtaining very firm conclusions about the whole city, because the theory of the monocentric city model expects existing housing to adapt in size. Existing housing (before 2013) is left out in this second regression. The plots in the second regression had zero houses in 2013, but end with at least one house in 2022. The following regression is made by a one-step system GMM approach using Windmeijer robust standard errors and collapsing instruments to prevent proliferation of it (Roodman, 2009). Too many instruments will cause inefficient results. A first regression with only Price_index as explanatory variable was taken, however this came with bad outcomes for the Hansen-Sargan overidentification test. Too little instruments were used. The Price_index variable was not significantly different from zero. However, it is significant when estimated with other control variables. It appeared to be negative, which was against expectations. The statistics needed to check assumptions are added below the table.

According to Roodman (2009) the Sargan-Hansen statistic is one the be very cautious with when values are just above the 0.05 or 0.10 p-values. When they are near to 1, they are considered 'too good to be true'. The Sargan-Hansen statistic values shown in the tables below are trustworthy. Also the second order autocorrelation tests are good. Several explanations can be drawn from these outcomes. For the first regression, an increase of 1 in price index (base year 2000 is 100) leads to a decrease of 0.0593% in floor area per hectare. Reason for this might be that less dense but commodious areas have bigger gardens and less houses, leading to higher prices. Both columns show the impact of several variables upon the log of floor space. Across all model specifications, the past floor space is a significant predictor of its current level. This denotes that floor space in grid cells tends to be path-dependent or is sticky, which means that the floor space level in the present year has a strong influence in determining her size the following year. In the first column the factor for filling the optimal floor space gap appears to be above 1, which is not expected. The $\gamma$ of minus 0.448 indicates the floor space in the neighborhoods did not move towards an optimum. It can be concluded that the outcomes are not as expected by theory. However, in the newly built grid cells the theoretical 'optimum gap' was filled by a factor of 0.499. These results might indicate that newly built sites move more towards an optimal floor space, compared to their previous year. This is logical, for newly built sites grow to a certain floor area in a short time period, adjustment growth rate declines after being developed. However, it should be noted that the 0.499 adjustment is estimated by taking newly built dwellings at one point in time only. Existing houses (from before 2013), setting a certain initial floor space, are not included in that estimation. That means theoretical predictions cannot be verified. The price component (Price_index) in the second regression appears to be slightly negative (at a p-value of 0.1). The increase of 1 in price index leads to a decrease of approximately 0.671% in floor area used. This factor might be explained by the fact that one batch (at one point in time) of houses is used in this regression. More expensive areas will have bigger houses and therefore use less floor area per hectare. The different factors for L.lnsum_woon_opp indicate that there is an accelerating growth rate of floor area within already built areas, while there is a declining growth in the recent years of the newly built areas of 2013. For the first regression, one percentage increase in floor area in the prior period leads to 1.448 percentage increase in the current period. This accelerating growth rate does not implicate the theoretical gap gets filled (theory is not applicable in this case). It might be that the optimum grows at a higher pace. Spatial restrictions or stickiness in housing adjustment may be the reason that floor space stemming from prior periods does not adjust to an optimum. In other words, the floor area used by existing housing probably grows slower than the optimum does. Taking into account the descriptive part, this makes sense. Newly developed sites eventually result in floor area of over 5000 square meter per hectare (figure 11 and 12), while existing areas remain around 3700 square meter per hectare (figure 14).

***Table 1: Regression outskirts***

| | All grids of Overvecht (WK034403), Noordoost (WK034404), Leidsche Rijn (WK034409) and Vleuten-De Meern (WK034410) | Empty in 2013, but built upon in later periods |
|---|---|---|
| VARIABLES | Dependent: lnsum_woon_opp | |
| L.lnsum_woon_opp | 1.448*** | 0.501** |
| | (0.198) | (0.223) |
| Price_index | -0.000593*** | -0.00671** |
| | (0.000225) | (0.00307) |
| count_woningen | -0.00848** | 0.00596* |
| | (0.00388) | (0.00335) |
| sum_footprint | -4.75e-05* | 1.92e-05 |
| | (2.57e-05) | (1.41e-05) |
| sd_bouwjaar | 0.00208** | 0.186** |
| | (0.00104) | (0.0747) |
| aant_dominantwoon | -0.0110** | 0.0170** |
| | (0.00495) | (0.00690) |
| Constant | -2.752** | 4.765*** |
| | (1.234) | (0.951) |
| Year dummies included | No (not significant) | No (not significant) |
| Groups/instruments | 2178/12 | 101/10 |
| AR (2) | 0.707 | 0.929 |
| F Statistic | 252237.40 | 16880.24 |
| Hansen Statistic | 0.519 | 0.255 |
| Observations | 14,603 | 301 |
| Number of grid_id | 2,178 | 101 |

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1
Source: Authors' computations

The results of Utrecht as a whole are shown in table 2. The AR(2) value of 0.071 is a point of concern, the test for autocorrelated error is beyond a five percent level, but not that far. Therefore, these results should be taken with some caution. The outcomes of the Hansen-Sargan statistic show lower p-values, but the exogeneity of the instruments can be trusted. The outcomes look similar to the newer outskirts in the first regression. The coefficient for the log of floor space still has a value larger than 1, as it was in the first regression of the previous part, though it is somewhat smaller. Largest difference is the 0.11 difference in the lagged dependent factor of the first column. Apparently, the newer parts of the city had more fluctuations in floor space compared to Utrecht as a whole. Reason for this might be that dense inner city parts adjust at a slower pace due to more stringent spatial restrictiveness. This results in higher values for the lagged dependent beta in the first regression, causing subsequent years to grow more in floor space. The price variable is still negative as it was in the previous regression. The outcomes suggest that floor area of older city parts grows slower compared to the newer outskirts. It can be argued that the coefficient for floor space indicates that the floor space grows at decelerating pace, while newer parts show declining growth rates. Though this explanation seems reasonable, it does not fit in the context of the theoretical part.

***Table 2: Utrecht regression***

| VARIABLES | All Utrecht grids<br>Dependent: lnsum_woon_opp | Empty in 2013, but built upon in later periods |
|---|---|---|
| L.lnsum_woon_opp | 1.334*** | 0.522** |
| | (0.129) | (0.199) |
| Price_index | -0.000310*** | -0.00694** |
| | (0.000115) | (0.00278) |
| count_woningen | -0.00542** | 0.00652* |
| | (0.00218) | (0.00353) |
| sum_footprint | -3.36e-05** | 1.61e-05 |
| | (1.48e-05) | (1.18e-05) |
| sd_bouwjaar | 0.000545** | 0.178** |
| | (0.000259) | (0.0728) |
| aant_dominantwoon | -0.00622*** | 0.0152*** |
| | (0.00234) | (0.00579) |
| Constant | -2.125** | 4.690*** |
| | (0.829) | (0.858) |
| Year dummies included | No (not significant) | No (not significant) |
| Groups/instruments | 3891/12 | 103/10 |
| AR (2) | 0.071 | 0.842 |
| F Statistic | 848065.31 | 18692.76 |
| Hansen Statistic | 0.223 | 0.169 |
| Observations | 26,480 | 310 |
| Number of grid_id | 3,891 | 103 |

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1
Source: Authors' computations

Both regression show an accelerating growth rate for floor area in existing housing and decelerating rates for newly built grids. Looking at table 2 the first regression for existing housing, one percentage increase in floor area in the prior period leads to 1.334 percentage increase in the current period. Whereas newly built sites gain one 0.522 percent increase if their prior period floor area was one percentage larger. It is logical to assume that growth rates in floor area are accelerating if the gap between optimal and actual floor area grows bigger, taking into account the low number of approximately 3650 square meter per hectare in 2020 (figure 14) and the 5100 square meter per hectare (figure 11 and 12). Apparently, newly built sites are built close to the current optimum and grow bit by bit onwards. The interpretation of above results without theoretical perspective are complementary to the following theoretical interpretation.

Previous results suggested several concluding thoughts. The regressions for Utrecht (table 2) and the subset regression (table 1) were not confirming the theory. Price had some impact upon floor area and the regression for newly built area's showed the floor area gap on a site filled over time, development was (almost) fulfilled after a certain period. The results suggest that the existing housing in Utrecht does not adjust to the optimal floor area, as theoretical predictions were not met. A negative adjustment factor was not expected. Spatial planning restrictions are likely to be severe. A reasonable explanation is that the optimal floor area grew more than the actual growth in floor area for existing housing. This is also suggested by the descriptive data from figure 9 to 14. The floor area for newly built sites grew from 4200 square meter per hectare to 5100 square meter, whereas existing housing grew from 3350 square meter to 3650 square meter.

A new data subset is taken to investigate the floor area used per hectare over time. The time span batches from figure 9 to 12 are taken. This regression with newly built grids of different time spans offers the opportunity to verify theory without using newly built grids of one point in time. Table 3 (next page) shows the regression of the grid cells described in figure 9 up to figure 12. That means this regression contains only grids with newly built housing from the batches of 2001 up to 2022. This data does not take into account housing built before 2001, therefore it does not measure the exact adjustment of all existing housing. Using only newly built grids leads to less accurate testing of the theoretical hypothesis. This regression shows a large adjustment factor, over 70% of the gap to an optimum is filled every year. All beta's appear to be significant, although very small. The coefficient for price is again small. If the price index goes up by 1 around the optimal floorspace rises with about 0,0455%. Note that the variables have to be multiplied by $\gamma$ , since the formula was rewritten by: $\alpha = \gamma \alpha'$ and $\beta = \gamma \beta'$. It is arguable that the floor space is adjusting towards an optimum. The 0.293 coefficient (significant at 0.01 level) for the lagged dependent fits the theory. According to theory this would mean the gap between actual and optimal floor space would be filled by 70.7% every year. This effect seems reasonable, taking into account the graphs from the descriptive statistics about newly built housing. Average floor space used per hectare of newly built dwellings increased over a decade by approximately 800 square meters per 10,000 square meters. This suggests that 80 square meters per year on average must be equal to the 70.7% adjustment. This means the gap would consist of approximately 114 square meters (100%) per year per grid cell. That is almost equal to the size of one house. To compare: existing housing grew approximately 40 square meters per hectare. Checking these numbers using descriptive data about average floor space per house confirms these approximations. The average floor space in 2012 was 121.5 square meter using about 4600 square meter per hectare leads to an average of 37.9 houses per hectare in 2012. The same numbers for 2017 (5100 square meter and 118 square meter) lead to 43.22 houses. That leads to approximately one house per year.

These findings suggest that the scarcity of land forces developers to use larger amounts of areas to be built. The impact of price may be argued differently for this regression with newly built dwellings, given the fact that municipalities and regulations have influence upon the price setting process when new housing in lower segments is developed. However, it is evident that sizes of newly built dwellings are smaller and more densely developed. This also clarifies the declining profitability of low and middle segment type of dwellings. These findings also indicate that the tension existing in the land market must have a certain impact upon (intergenerational) wealth (living space) distribution.

*Table 3: Newly built grids (in batches) 2001-2022*

| VARIABLES | Batch grids<br>Dependent:<br>lnsum_woon_opp |
|---|---|
| L.lnsum_woon_opp | 0.293*** |
| | (0.0856) |
| Price_index | 0.000644*** |
| | (0.000157) |
| count_woningen | 0.00761*** |
| | (0.00158) |
| sum_footprint | 4.74e-05*** |
| | (1.70e-05) |
| aant_dominantwoon | 0.0110*** |
| | (0.00164) |
| Constant | 5.047*** |
| | (0.592) |
| Year dummies included | No (not significant) |
| Groups/instruments | 694/40 |
| AR (2) | 0.138 |
| F Statistic | 503229.63 |
| Hansen Statistic | 0.339 |
| Observations | 4,968 |
| Number of grid_id | 694 |

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1
Source: Authors' computations

## 4.2 Recommendations

It is important to know what factors drive the floor space of both existing housing and newly built housing. Further research might investigate the impact of reconstructing or adjustment upon floor spaces. This research did not have outcomes predicted by theory for existing housing. Therefore, the impact of (or change in) spatial regulations and governmental interventions on the amount of floor space used by existing housing might be researched. Secondly, it might be interesting to investigate the impact of declining prices in real estate and land markets. This dataset uses floor space measured from 2013, so this research only covers periods of price hikes. Thirdly, it would be useful to conduct the same research in other countries to compare their growth rates and adjustment factors. Fourthly, the densification by (real) prices results real estate to be developed differently (more dense) compared to other periods. It is useful to know what the 'option value' of a certain densification is. Research upon price input differences and value outcomes might clarify the value of these options.

# 5. Conclusion

Given the sparse empirical evidence about the impact of prices for the size of urban dwellings and newly built apartments, this research investigates the impact of real estate price hikes upon the size of urban dwellings in ten neighborhoods in Utrecht from 2013-2022.

Results of this study are quite convincing, at least for newly developed sites. However, the sys-GMM results should be interpreted with certain caution. The small influence of price upon floor space used, indicates that prices play a minor but yet significant role in the density of both newly built and existing housing. The growth coefficient for floor area of Utrecht, suggests that existing housing is very sticky to its size. The optimal floor space shifts away from existing floor space. The descriptive statistics of this study show that newly built dwellings are built by using more floor space per grid cell and below the average floor space per house of existing housing. These two facts are supported by the estimated effects in the empirical part. Existing housing sticks to its original size, whereas newly built housing is built close to more or less optimized floor spaces. This stickiness is likely to be the consequence of strict spatial land use policy. The findings are in line with the signs of tension seen on the land and housing market. New housing is developed more and more by using lower amounts of land per house. In addition, land is used more intensively, meaning that lower amounts of land can be used for other purposes. The findings are also relevant when considering large scale sustainability measurements for future housing projects. These measurements should focus upon densely built areas.

# 6. References

Ahlfeldt, G. M., & McMillen, D. P. (2018). Tall buildings and land values: Height and construction cost elasticities in Chicago, 1870–2010. *Review of Economics and Statistics*, *100*(5), 861–875.

Blundell, R., & Bond, S. (1998). Initial conditions and moment restrictions in dynamic panel data models. *Journal of econometrics*, *87*(1), 115–143.

Boelhouwer, P. (2020). The housing market in The Netherlands as a driver for social inequalities: Proposals for reform. *International Journal of Housing Policy*, *20*(3), 447–456.

Boelhouwer, P., & Hoekstra, J. (2009). Towards a better balance on the Dutch housing market? Analysis and policy propositions. *European Journal of Housing Policy*, *9*(4), 457–475.

Bond, S., Hoeffler, A., & Temple, J. (2001). *GMM Estimation of Empirical Growth Models* (Economics Paper Nr. 2001-W21; p. 7). Economics Group, Nuffield College, University of Oxford. https://econpapers.repec.org/paper/nufeconwp/0121.htm

Brueckner, J. (1987). *The Structure of Urban Equilibria: A Unified Treatment of the Muth-Mills Model*. http://www.socsci.uci.edu/~jkbrueck/course%20readings/handbook%20chapter.pdf

Buitelaar, E. (2021). *De werking van de grondmarkt en de rol van de overheid: Verkenning en reflectie*.

Colwell, P. F., & Munneke, H. J. (1999). Land prices and land assembly in the CBD. *The Journal of Real Estate Finance and Economics*, *18*(2), 163–180.

DiPasquale, D. (1999). Why don't we know more about housing supply? *The Journal of Real Estate Finance and Economics*, *18*(1), 9–23.

DiPasquale, D., & Wheaton, W. C. (1994). Housing market dynamics and the future of housing prices. *Journal of urban economics*, *35*(1), 1–27.

Doodeman, M. (2021, april 15). *Dure nieuwbouwwoningen vinden gretig aftrek*. Cobouw. https://www.cobouw.nl/294852/dure-nieuwbouwwoningen-vinden-gretig-aftrek

Ebrahimigharehbaghi, S., Qian, Q. K., Meijer, F. M., & Visscher, H. J. (2019). Unravelling Dutch homeowners' behaviour towards energy efficiency renovations: What drives and hinders their decision-making? *Energy Policy*, *129*, 546–561.

Fuller, G. W., Johnston, A., & Regan, A. (2020). Housing prices and wealth inequality in Western Europe. *West European Politics*, *43*(2), 297–320.

Groenemeijer, L. (2021). *Terugblik ontstaan woningtekort*. ABF Research.

Kripfganz, S. (2019). Generalized method of moments estimation of linear dynamic panel data models. *London Stata Conference*, *17*.

McDonald, J. F. (1981). Capital-land substitution in urban housing: A survey of empirical estimates. *Journal of Urban Economics*, *9*(2), 190–211. https://doi.org/10.1016/0094-1190(81)90040-1

Muellbauer, J., & Murphy, A. (2008). Housing markets and the economy: The assessment. *Oxford review of economic policy*, *24*(1), 1–33.

Needham, D. B., Kruijt, B., & Koenders, P. (2018). *Urban land and property markets in the Netherlands*. Routledge.

Nieuwbouwwijzer. (2020, oktober). *Nederland blijft verbouwen*. https://www.nieuwbouwwijzer.nl/hypotheek/nederland-blijft-verbouwen/

PBL. (2019, september 5). *PBL/CBS Regionale bevolkings- en huishoudensprognose 2019-2050: Belangrijkste uitkomsten* [Text]. PBL Planbureau voor de Leefomgeving. https://www.pbl.nl/publicaties/pbl-cbs-regionale-bevolkings-en-huishoudensprognose-2019-2050-belangrijkste-uitkomsten

Roodman, D. (2009). How to do xtabond2: An introduction to difference and system GMM in Stata. *The stata journal*, *9*(1), 86–136.

Rouwendal, J. (2022). *High prices, old buildings and urban housing dynamics: Evidence from Amsterdam*.

Van der Krabben, E. (2021, november 12). *De werking van de grondmarkt: Gevolgen voor woningbouw en woningmarkt | ACM.nl* [Onderzoek]. https://www.acm.nl/nl/publicaties/de-werking-van-de-grondmarkt-gevolgen-voor-woningbouw-en-woningmarkt

van Haaster-de Winter, M. A., Dijkshoorn-Dekker, M. W. C., Polman, N. B. P., Mattijssen, T. J. M., & Kortstee, H. (2020). *Natuurinclusief bouwen: Wat beweegt de Nederlandse vastgoedsector?: Onderzoek naar motieven, barrières en gedrag in de transitie naar meer groen*. Wageningen Economic Research.

Vermeulen, W., & Rouwendal, J. (2007). *Housing supply and land use regulation in the Netherlands*.

Voigtländer, M. (2014). The stability of the German housing market. *Journal of Housing and the Built Environment*, *29*(4), 583–594.

# 7. Appendix

## 7.1 Descriptive statistics

Number of observations, mean, standard deviation, minimum and maximum by neighborhood (WK, neighborhoods are shown graphically in figure 4 by the last two numbers). The counts for 'woningen' (houses) are total observations, other variables contain cells left empty. For example, WK034401 has 20608 observations and only 9856 cells have an observation for sum footprint.

**WK031200**

| | N | mean | sd | min | max |
|---|---|---|---|---|---|
| count panden | 11 | 1 | 0.000 | 1 | 1 |
| sum opp woon | 11 | 0 | 0.000 | 0 | 0 |
| sum footprint | 11 | 37.725 | 0.000 | 37.725 | 37.725 |
| count woningen | 23 | 0 | 0.000 | 0 | 0 |
| avg bouwjaar | 11 | 1956.818 | 24.722 | 1939 | 1988 |
| sd bouwjaar | 11 | 0 | 0.000 | 0 | 0 |
| **WK034401** | | | | | |
| count panden | 9856 | 13.406 | 25.198 | 0 | 142 |
| sum opp woon | 9856 | 1300.403 | 2855.434 | 0 | 98570 |
| sum footprint | 9856 | 2225.962 | 3804.973 | 0 | 54315.699 |
| count woningen | 20608 | 13.21 | 28.009 | 0 | 192 |
| avg bouwjaar | 7318 | 1966.723 | 28.388 | 1891 | 2021 |
| sd bouwjaar | 7318 | 8.36 | 11.434 | 0 | 55.509 |
| **WK034402** | | | | | |
| count panden | 4961 | 45.441 | 37.640 | 0 | 161 |
| sum opp woon | 4961 | 3751.384 | 2655.678 | 0 | 13894 |
| sum footprint | 4961 | 2449.698 | 1528.392 | 0 | 14908.5 |
| count woningen | 10373 | 42.624 | 32.446 | 0 | 238 |
| avg bouwjaar | 4546 | 1951.904 | 30.581 | 1650 | 2019 |
| sd bouwjaar | 4546 | 16.231 | 14.003 | 0 | 118.24 |
| **WK034403** | | | | | |
| count panden | 9328 | 10.358 | 19.557 | 0 | 122 |
| sum opp woon | 9328 | 1647.307 | 2757.336 | 0 | 34968 |
| sum footprint | 9328 | 1191.408 | 1786.999 | 0 | 13298 |
| count woningen | 19504 | 17.578 | 31.115 | 0 | 410 |
| avg bouwjaar | 6450 | 1974.114 | 23.653 | 1800 | 2021 |
| sd bouwjaar | 6450 | 6.958 | 12.249 | 0 | 103.784 |
| **WK034404** | | | | | |
| count panden | 5588 | 30.815 | 30.624 | 0 | 117 |
| sum opp woon | 5588 | 3324.445 | 3179.895 | 0 | 18706 |
| sum footprint | 5588 | 1852.323 | 1666.878 | 0 | 18632.301 |
| count woningen | 11684 | 30.912 | 35.270 | 0 | 497 |
| avg bouwjaar | 4538 | 1953.764 | 35.925 | 1864 | 2021 |
| sd bouwjaar | 4538 | 13.188 | 14.569 | 0 | 93.921 |
| **WK034405** | | | | | |
| count panden | 12166 | 9.999 | 19.981 | 0 | 109 |
| sum opp woon | 12166 | 1509.911 | 9151.915 | 0 | 480487 |
| sum footprint | 12166 | 1249.488 | 2854.523 | 0 | 62701.301 |
| count woningen | 25438 | 11.539 | 27.666 | 0 | 506 |
| avg bouwjaar | 6441 | 1957.814 | 38.093 | 1670 | 2020 |
| sd bouwjaar | 6441 | 12.888 | 16.395 | 0 | 157.882 |

| | | | | | |
|---|---|---|---|---|---|
| **WK034406** | | | | | |
| count panden | 2992 | 19.799 | 22.178 | 0 | 113 |
| | N | mean | sd | min | max |
| sum opp woon | 2992 | 3756.331 | 3957.716 | 0 | 21157 |
| sum footprint | 2992 | 3907.479 | 7608.255 | 0 | 113968 |
| count woningen | 6256 | 31.918 | 37.870 | 0 | 266 |
| avg bouwjaar | 2557 | 1885.491 | 90.937 | 1646 | 2021 |
| sd bouwjaar | 2557 | 77.677 | 66.893 | 0 | 186.293 |
| **WK034407** | | | | | |
| count panden | 5225 | 22.289 | 29.809 | 0 | 122 |
| sum opp woon | 5225 | 2330.642 | 2659.549 | 0 | 14062 |
| sum footprint | 5225 | 1386.539 | 1550.381 | 0 | 13098.1 |
| count woningen | 10925 | 25.842 | 32.346 | 0 | 282 |
| avg bouwjaar | 3692 | 1973.142 | 21.078 | 1900 | 2021 |
| sd bouwjaar | 3692 | 7.827 | 10.891 | 0 | 55.154 |
| **WK034408** | | | | | |
| count panden | 6083 | 18.277 | 27.304 | 0 | 126 |
| sum opp woon | 6083 | 2733.045 | 3293.785 | 0 | 21578 |
| sum footprint | 6083 | 2037.509 | 2588.200 | 0 | 33984.5 |
| count woningen | 12719 | 28.201 | 39.553 | 0 | 640 |
| avg bouwjaar | 4820 | 1968.712 | 24.071 | 1891 | 2021 |
| sd bouwjaar | 4820 | 8.965 | 11.327 | 0 | 63.422 |
| **WK034409** | | | | | |
| count panden | 12375 | 13.923 | 23.152 | 0 | 118 |
| sum opp woon | 12375 | 1399.492 | 2466.632 | 0 | 80899 |
| sum footprint | 12375 | 1190.491 | 2482.779 | 0 | 61062.398 |
| count woningen | 25875 | 8.077 | 18.246 | 0 | 399 |
| avg bouwjaar | 7422 | 1997.816 | 24.526 | 1750 | 2021 |
| sd bouwjaar | 7422 | 5.352 | 12.896 | 0 | 124.451 |
| **WK034410** | | | | | |
| count panden | 40546 | 7.62 | 18.214 | 0 | 126 |
| sum opp woon | 40546 | 623.147 | 1411.530 | 0 | 19557 |
| sum footprint | 40546 | 554.742 | 1149.228 | 0 | 18538.301 |
| count woningen | 84778 | 3.962 | 10.758 | 0 | 167 |
| avg bouwjaar | 14310 | 1975.386 | 45.520 | 1647 | 2021 |
| sd bouwjaar | 14310 | 9.388 | 19.533 | 0 | 179.5 |
| **WK035300** | | | | | |
| count panden | 11 | 0 | 0.000 | 0 | 0 |
| sum opp woon | 11 | 0 | 0.000 | 0 | 0 |
| sum footprint | 11 | 0 | 0.000 | 0 | 0 |
| count woningen | 23 | 0 | 0.000 | 0 | 0 |
| avg bouwjaar | 0 | . | . | . | . |
| sd bouwjaar | 0 | . | . | . | . |