BENCE MÉRŐ, ANDRÁS BORSOS, ZSUZSANNA HOSSZÚ, ZSOLT OLÁH,

NIKOLETT VÁGÓ

A HIGH RESOLUTION AGENT-BASED MODEL OF THE HUNGARIAN HOUSING MARKET

MNB WORKING PAPERS | 7

2022
OCTOBER

A HIGH RESOLUTION AGENT-BASED MODEL OF THE HUNGARIAN HOUSING MARKET

MNB WORKING PAPERS | 7

2022
OCTOBER

The views expressed are those of the authors' and do not necessarily reflect the official view of the central bank of Hungary (Magyar Nemzeti Bank).

MNB Working Papers 2022/7

A High Resolution Agent-based Model of the Hungarian Housing Market *

(A magyar lakáspiac komplex, ágensalapú elemzése)

Written by Bence Mérő**, András Borsos***, Zsuzsanna Hosszú****, Zsolt Oláh****, Nikolett Vágó****

Budapest, October 2022

Published by the Magyar Nemzeti Bank

Publisher in charge: Eszter Hergár

Krisztina körút 55., H-1013 Budapest

www.mnb.hu

ISSN 1585-5600 (online)

*We would like to thank Dirk Kohlweyer, Marco Pangallo, Doyne Farmer, Sándor Winkler, Bálint Dancsik, Alexandr Maxim Palicz, János Szakács, Ádám Banai, Tamás Nagy, Gergely Fábián and participants at the 2019 and 2022 WEHIA conferences, and ABM4policy seminars for their useful comments. Our working paper has been awarded the Third Prize in the 2022 "Rebuilding Macroeconomics: Complexity in Macroeconomics" Research Prize Competition funded by the UK's Economic and Social Research Council.

**Central Bank of Hungary, Krisztina körút 55, 1013 Budapest, Hungary and Corvinus University of Budapest, Fővám tér 8, 1093 Budapest, Hungary

***Central Bank of Hungary, Krisztina körút 55, 1013 Budapest, Hungary and Complexity Science Hub Vienna, Josefstädter Strasse 39, A-1080 Vienna, Austria

****Central Bank of Hungary, Krisztina körút 55, 1013 Budapest, Hungary

Contents

Abstract

This paper presents a complex, modular, 1:1 scale model of the Hungarian residential housing market. All the 4 million households and their relevant characteristics are represented based on empirical micro-level data coming from the Central Credit Information System, the Pension Payment database and transaction data of property sales collected by the National Tax and Customs Administration and the largest real estate agencies. The model features transactions in the housing and rental markets, a construction sector, buy-to-let investors, housing loans, house price dynamics and a procyclical banking sector regulated by a macroprudential authority. The flats in the model are characterized with detailed attributes regarding their size, state and neighbourhood quality. Households choose the flat with the highest consumer surplus according to standard utility maximization theory. Additionally, we have also implemented demographic trends, including childbearing, marriage and inheritance. This way the model is suitable for analysing various types of macroprudential, fiscal and monetary policies as well as for the assessment of exogenous shock scenarios. Initiating the model simulation from 2018, it managed to reproduce the number of transactions and the observed house price dynamics in most of the regions of Hungary for 2018-2019, while the volume of new housing loans and their distribution regarding income deciles and loan-to-value ratios were also in compliance with the empirical data.

JEL Classification: C63, D1, D31, E58, R21, R31

Keywords: agent-based modelling, macroprudential policy, housing market, housing loans

Összefoglaló

Jelen cikk a magyar lakáspiac komplex, moduláris, 1:1 arányú modelljét mutatja be. Mind a 4 millió háztartást és azok releváns jellemzőit megjelenítettük mikroszintű adatbázisok, a Központi Hitelinformációs Rendszer, az Országos Nyugdíjfolyósító, a Nemzeti Adó- és Vámhivatal által gyűjtött ingatlanértékesítési és illetékfizetési, valamint a legnagyobb ingatlanügynökségek tranzakcióit tartalmazó adatbázisok alapján. Modellünk szereplői a háztartások, az építőipar, a professzionális befektető és a makroprudenciális hatóság által szabályozott bank. Keretrendszerünkkel elemezhetővé válnak a lakás- és a bérleti piaci tranzakciók, a lakásár-dinamika alakulása és a lakáspiac szempontjából fontos hitelpiaci folyamatok. A lakások tulajdonságait is részletesen jelenítettük meg a modellben, minden lakásnak három jellemzője van: a mérete, az állapota, és a területi elhelyezkedése. A háztartások lakásvásárlás és -bérlés esetén a haszonmaximalizálási elméletnek megfelelően döntenek: a legnagyobb fogyasztói többlettel rendelkező lakást választják. Emellett demográfiai folyamatokat is tartalmaz a modellünk a gyermekvállalásra és a házasságkötésre vonatkozóan, amelyet többek között az öröklés során vettünk figyelembe. Így a modell alkalmas makroprudenciális, fiskális és monetáris politikai eszközök elemzésére, valamint különböző exogén sokkokat tartalmazó forgatókönyvek kiértékelésére. A modellszimulációt 2018-tól indítva sikerült reprodukálni a tranzakciók számát és a lakásárak dinamikáját Magyarország legtöbb régiójában 2018-2019 negyedéveiben, míg az újkihelyezésű lakáshitelek volumene és megoszlása jövedelmi decilisek, illetve hitelfedezeti-mutató kategóriák szerint szintén megfeleltek az empirikus adatoknak.

1 Introduction

The housing market is one of the most fundamental segments of the economy with salient influence on the financial position of households, investors, the construction sector and even on the stability of the banking system. Hence, disturbances on this market can have far-reaching economic and societal implications, which grants a special role to the housing market for policy makers.

However, there are also several characteristics of this market which make it difficult to represent all the relevant mechanisms and to obtain detailed policy assessments using the conventional toolkit of economics. The first type of these complications consists of relatively standard considerations in the literature, such as cyclical behaviour (Ceron and Suarez (2006), Cunningham and Kolet (2011)); interlacement with the financial sector due to the typically highly leveraged transactions (Acharya and Richardson (2009), Glaeser and Sinai (2019)); the illiquidity of real estates (especially during distressed periods) (Lin and Vandell (2007), Garriga and Hedlund (2020)); or the geographic fragmentation of the housing market (Goodman and Thibodeau (1998), Muth (2017)). However, there is also a second branch of difficulties not addressed sufficiently in standard economic models, which arises from the high level of heterogeneity and non-linearity due to the many types of economic agents interacting simultaneously with diverse motivations and irrational expectations. These features result in a complex system, which is impossible to capture in a sufficiently detailed manner without circumventing many of the typical assumptions of economic models.

Despite these intricacies, there are several papers in the mainstream literature attempting to examine the housing market either by using empirical econometric analysis (Kelly et al. (2018), Saunders and Tulip (2020), De Stefani (2020)) or DSGE models. In the latter case, the most typical strategy is to introduce the housing sector by considering housing as collateral for borrowing (Nobuhiro and Moore (1997), Iacoviello (2002), Kiyotaki et al. (2011)). As the size of the available loan depends on the changes in property prices, it can also be influenced indirectly in these models by monetary policy shocks. There are also several DSGE models which consider the interactions between monetary policy and – usually macroprudential – measures that are more directly linked to house prices. Funke and Paetz (2012), Lambertini et al. (2013) and Rubio and Carrasco-Gallego (2014) examine the role of loan-to-value (LTV) regulations, while Gelain and Bank (2011) and Gelain et al. (2012) expand the scope of macroprudential instrument under scrutiny also to loan-to-income and leverage ratios. Furthermore, Kannan et al. (2012) emphasizes the role of the underlying shock's type in the determination of the optimal monetary and macroprudential policies, while Mendicino and Punzi (2014) and Rubio and Comunale (2017) consider two-economies settings in determining the optimal policy mix. There are also a few papers with DSGE models in which the authors try to deviate somewhat from the usual assumptions in order to accommodate with some of the above described characteristics of the housing market. In the model of Kuang (2014) – contrary to the rational expectation assumption – households are not aware of each other's preferences and expectations, which feature results in endogenously generated credit-, and house prices cycles. Ortalo-Magne and Rady (2006) relaxed another typical assumption by developing an equilibrium life-cycle model in which households are heterogeneous regarding their income and preferences, and restricted by credit constraints. This setup has an impact on households' savings, on the timing of the house purchases, on the size of the purchased homes and ultimately on the volatility in house prices.

The introduction of bounded rationality and household heterogeneity are indeed important for the analysis of the housing market, however, these extensions are only partial remedies to the shortcomings of DSGE models. The easiest way to represent more realistically the considerable heterogeneity of the agents and the complex behaviour of owners, renters, investors, banks and regulators is the use of agent-based models (ABMs).

The literature contains a large number of agent-based models suitable for analysing the housing market. These are often macroeconomic ABMs with moderately elaborated housing markets. For example Erlingsson et al. (2014) and Ozel et al. (2019) are both based on one of the most complex agent-based models at present, EURACE (see Deissenberg et al. (2008)), which makes it possible to analyse the trade-off between macroeconomic performance and financial stability. In the same line of research, Lauretta (2018) extends the model of Erlingsson et al. (2014) with interactions between securitization and financial innovations, while Kaszowska-Mojsa and Pipień (2020) presents a EURACE-based ABM tailored to the Polish economy with macroprudential applications. Fatouh et al. (2019) also focuses on macroprudential regulation in an ABM by modelling the UK banking system and several Basel III policies. In connection with output and stability considerations, economic inequality can also be analyzed with general ABMs through the housing market; e.g. Cardaci (2018) demonstrates that growing inequality is conducive to excessive indebtedness in home equity based loans, which can lead to higher default rates, credit crunches and eventually decline in the output.

There are also several “pure” housing market ABMs without a detailed macroeconomic environment. Some of these consider fictive agents to demonstrate theoretical claims, while others attempt to mimic the entities and real estates of a real city or country. One of the first papers in the former category is the model of Devisch et al. (2009), which elaborates fictive agents’ housing decisions and the price negotiation process by using detailed utility functions. Magliocca et al. (2011) describes the housing market and the expansion of a fictive city through the urbanization process of farmland areas. Ge (2013) and Ge (2017) also work without empirical data on studying how the easing of certain lending constraints leads to endogenous bubbles with high volatility in property prices. Finally, Pangallo et al. (2019) also builds a simulation of a fictive city, and they examine the connections between income inequality, segregation and house prices.

However, many research projects have been conducted recently on housing market ABMs based on empirical data. With this paper we also attempt to contribute to this branch of the literature. The works most closely related to our research are Axtell et al. (2014) and Baptista et al. (2016). Axtell et al. (2014) model the housing market of Washington, D.C. and generate certain characteristics relying on the available data, most importantly all housing market transactions concluded between 1997 and 2009. This model could be scaled up to more than 2 million households of Washington D.C., and it also features endogenously evolving asset price bubbles. They conclude that the tighter interest rate policy reduces the size of the bubble only to a limited extent, while the tightening of leverage may have substantial effect. This model was further developed and applied to the British housing market by Baptista et al. (2016). In this study, houses differ from each other only in one characteristics, which can be regarded as a quality parameter. Households always buy the highest quality house on the market which is available within their reservation price. Cycles in this model are amplified by the decisions of buy-to-let investors. The paper examined the impact of a macroprudential instrument (the loan-to-income (LTI) limit) introduced by the Bank of England in 2014, and found that the application of the LTI limit can mitigate the housing market cycles. Cokayne (2019) has used this model as a starting point to build a housing ABM calibrated to the Danish housing market. Furthermore, Laliotis et al. (2020) has also built a simplified version of Baptista et al. (2016) to examine the effects of the LTV regulation at the country-level within Europe. There are a few further papers featuring housing market ABMs based on empirical data. Gilbert et al. (2009) performed simulations using a model of buyers, sellers and real estate agents calibrated for the British housing market. Kouwenberg and Zwinkels (2015) analysed the cyclical dynamics in the US housing market modelling different investment strategies. Carstensen (2015) examined the Danish housing market, while Glavatskiy et al. (2021) and Evans et al. (2021) modelled the housing market of Sydney at the neighbourhood level utilizing many granular datasets for calibration.

Our paper contributes to these endeavours by proposing a comprehensive model of the Hungarian housing market. Its central motive is to model transactions on the market in order to enable the detailed analysis of house price and housing loan dynamics. Since investors’ demand contributes to a large extent to the Hungarian housing market, not only purchases but also renting is an integral part of our framework. This element is also essential to assess the accumulation of households’ savings which is an important determinant of house purchases and housing lending. While these building blocks are common in many housing market ABMs, our model presents several new features and improvements compared to the existing practices in the literature. Among these the most important ones are the following.

Full-scale empirical mapping

As discussed above, there are several studies using empirical data to build and calibrate ABMs, however, our approach takes a step forward by actually modelling all the 4 million households and their flats in Hungary at a 1:1 scale representation. The agents and their relevant characteristics are reconstructed by a comprehensive mapping of the households and the housing market to micro-level data coming from the Central Credit Information System, the Pension Payment database and transaction data of property sales collected by the National Tax and Customs Administration and the largest real estate agencies in Hungary. Furthermore, the model also considers detailed demographic trends and marital dynamics following empirical tendencies. Besides the higher level of plausibility in applications, this data-intensive framework has an additional benefit of minimizing the burn-in effect in the simulations.

Characteristics of flats

Most of the papers consider only a single "quality" measure to quantify the desirability of flats. In contrast, we characterize flats by three attributes: *size, state* and *location*. While the *size* of flats is straightforward to define, the other two features are more arduous to specify. In the case of *state*, we utilized several empirically observed traits which we compressed into a single continuous variable using a hedonic regression model. Regarding *location*, we introduced two novel solutions: (i) we have divided the country into actual, interpretable neighbourhoods instead of the typical grid-, or graph-based spatial representation of the housing market (which can be found in e.g. Gilbert et al. (2009), Devisch et al. (2009), Ge (2017), Pangallo et al. (2019) or Evans et al. (2021)); (ii) then – similarly to the procedure applied in the case of *state* – we assigned a cardinal quality measure to each neighbourhood as well.

Changes in the housing stock

The model also takes into account the changes of the housing stock, which can happen in the form of *(i) amortization, (ii) renovation* and *(iii) new constructions*. The quality, i.e. the *state*, of flats deteriorates at a predefined rate, but in every month a random set of households can renovate their apartment. While – to our knowledge – renovations cannot be found in other housing market ABMs, constructions are sometimes modelled, but only for smaller urban areas (e.g. Magliocca et al. (2011), Ge (2017)), and not for a whole country. The construction sector in our model consists of one representative company, and it builds flats with the highest possible *state* value. The location, size and quantity of new flats are set by the construction sector based on simulated, fictive demand (see A.2.2). The price of newly built flats is determined by the price of the building site, the construction cost and the extra margin of new flats.

Households' decision making process

Agents in the model can decide on moving not only at a certain age or at a predefined time, but in any time period depending on their assessment on the potential utility gain. Households make offers on flats based on their consumer surplus calculated from the utility of the purchase expressed in monetary terms (in the spirit of Magliocca et al. (2011)). The higher this surplus is, the higher the probability and the size of an offer will be. Before making an offer, households take into account also the cost of renovation. In order to carry out these calculations we assigned each household a utility function which has been calibrated individually to their preferences based on empirical data about their actual homes. The feature of adding utility functions to the agents is only present in a handful of housing ABMs (e.g. Magliocca et al. (2011), Ge (2017), Pangallo et al. (2019)), and it is constructed usually in a Cobb-Douglas form. Our model, in turn, uses a composite utility function encompassing trade-offs not only between housing and consumption, but also among the attributes of the flats.

Credit market and the banking system

In order to prepare the model for conducting macroprudential analyses in a plausible manner, we incorporated several details of the credit market and the banking system. First of all, the bank increases and reduces its credit supply procyclically. Furthermore, there are not only housing loans, but also bridge loans and personal loans. Housing loans can have either fixed or variable rates. A household is eligible for a loan if (i) it meets the relevant LTV and DSTI (Debt Service to Income) rules; (ii) its expected income covers the credit payments and a minimal consumption level; (iii) and it did not have a defaulting loan in the past five years. If a household becomes non-performing, they will try to restructure the loan in order to meet their payment obligation. If this adjustment is not sufficient, foreclosure (and thus, liquidation at a discount) of the flats are also possible in the model.

Endogenous cycles

Although there are several papers featuring the analysis of cycles in housing ABMs (e.g. Kouwenberg and Zwinkels (2015), Baptista et al. (2016), Lauretta (2018), Cokayne (2019)), we offer a novel mechanism resulting in endogenous cycles. In the case of a tight housing market, typical households have to wait longer to find and purchase a suitable flat compared to normal market conditions. Hence, the longer they have been stranded on the market, the more tolerant they become regarding both the price and the characteristics of flats. While this mechanism drives the prices upwards, plateauing will be inevitable as credit constraints become more and more binding, which limits the possibilities of households for purchasing using loans. Finally, after the impatient households have managed to buy at higher prices the tightness of the market relaxes.

While these features increase the complexity considerably, they also make it possible to surpass existing tools in granularity and precision, which enhances the models' usefulness in many applications. This way our proposed framework is suitable for analysing various types of macroprudential and fiscal policies, as well as for the assessment of exogenous shock scenarios. (However, in addition to this empirically designed model we also provide a version which uses only simulated data (see Appendix A.8) and therefore can be freely shared upon publication. We will refer to this version as *sample model* throughout the paper.)

The remainder of the paper is structured as follows. Section 2 gives a high-level overview of the model to facilitate the general understanding. Section 3 highlights and exhaustively describes the most important features (while the less emphasized details are delegated to the appendix). Section 4 concerns with the calibration, validation and sensitivity analysis of the model. Section 5 discusses the key contributions, proposes directions for future research and concludes the paper.

2 The model in a nutshell

We provide a general outline of the model by describing the agent types, reviewing the most important mechanisms and finally the timeline of the simulation steps. Readers interested in the finer specifications of the model can consult Section 3 and Appendix A, while more detailed information on the generation of agents and their attributes from empirical data can be found in Appendix B.

2.1 CONSTITUENTS OF THE MODEL

The model incorporates (i) the population of individuals, households and flats; (ii) a representative investor, construction company and bank; (iii) furthermore authorities responsible for fiscal policies and macroprudential regulation. (See Figure 1.) As the rules and parameters of the fiscal and macroprudential policies are given exogenously, we only provide a description of these in the Appendix A.9.

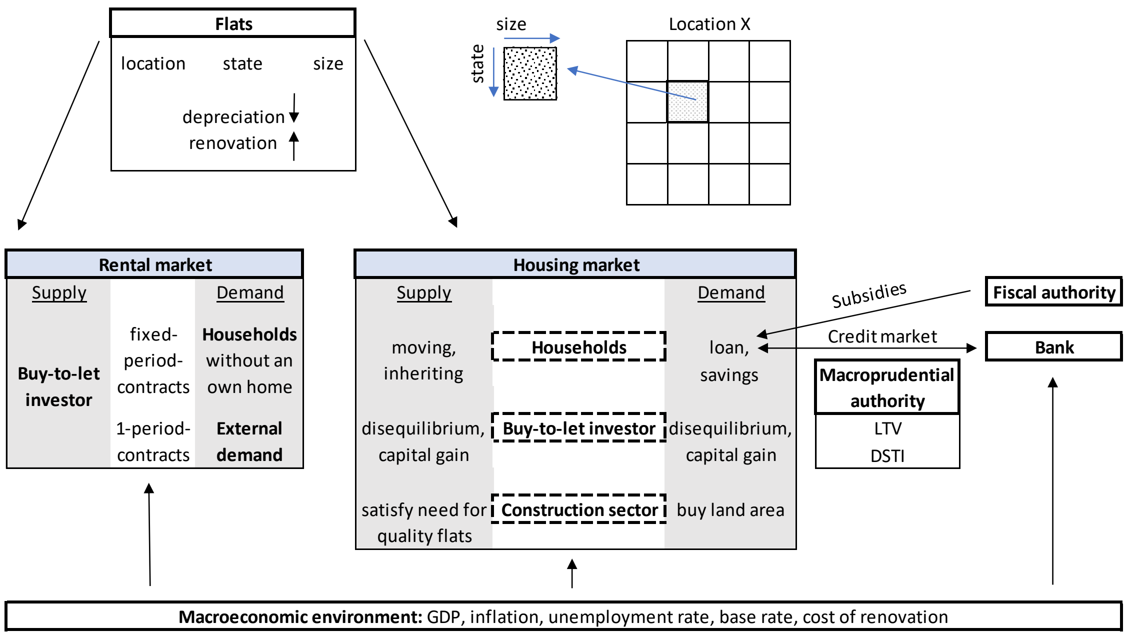
2.1.1 INDIVIDUALS AND HOUSEHOLDS

Every household in the model pursues having an own flat, however, as long as they cannot afford it, they can only have a dwelling via the rental market. If a household has a sufficiently favorable financial position, it cannot only own a home, but also additional flats as investments. Although these transactions are carried out at the level of households in the model, the granularity of our framework expands even deeper by representing also the internal structure of almost 4 million households. They consist of either an individual adult or pairs of adults possibly with children. The shifts between these configurations as well as the characteristics of the individuals within a household can fundamentally influence housing preferences. This way, the concept of individuals and households are closely related, thus, they are described jointly in this section.

We also represent several demographic phenomena of individuals and households in the model. Children after a certain age leave their parents' household and start a new one either alone or with a spouse. Newly formed married couples can also have children who will be part of the household. Although there is no divorce in the model, individuals age and pass away independently of their partners. Albeit the time horizon of the model simulations includes typically only a few years, demographic events still have an important role. Starting from 2016, several subsidies have been introduced in Hungary to support the purchases of flats depending on the number of children in the family. These policies can materially ease the credit constraints for the affected families because additionally to the very favorable credit conditions the subsidies can even be taken into account as part of the down payment. The two most important subsidy programs are the *Home Purchase Subsidy Scheme for Families* and the *Prenatal Baby Support Loan*. While the former affected 15-30% of the loan contracts (depending on the year), the latter can be connected to almost half of the non-housing loans issued since July, 2019. The implications of these policies can only be taken into account accurately if one adds all the relevant demographic processes in the model as well.

From the point of view of the labor market status, each individual's life has three stages: they are inactive at first, then they enter the labor market (and become either employed or unemployed), and finally they retire. Members of a household can become unemployed independently from each other, furthermore, external shocks can also affect their salaries. The age of entering the labour market, the probability and length of unemployment and also the wage dynamics of individuals are determined by their educational level which is classified into three categories. At the age of retirement, pensioners receive a fraction of their last gross salary, which will be corrected each year by the growth rate of the nominal GDP. The details of generating the above mentioned characteristics and life events of individuals and households from empirical data is described comprehensively in Appendix B.3.

Figure 1
Schematic outline of the agent-based model of the Hungarian housing market.



2.1.2 FLATS

Although we do not observe all the flats in Hungary, we could still reconstruct an approximation of the housing stock by using several partial datasets. To ensure representativity, we multiplied these empirical observations using weights such that the distribution of the generated flats would match the aggregate housing statistics in Hungary. The details of this procedure as well as the description of the empirical housing data can be found in Appendix B.1.

Each flat in the model can be characterized by three attributes: *size, state* and *location*, which are represented by continuous variables. While this is straightforward in the case of the *size* of flats, the other two variables had to be constructed by reducing the high-dimensional space of various characteristics of the flats. In the case of the *state*, we compressed several empirically observed traits into one categorical variable along which the flats have been divided into 21 categories. Then, using a hedonic regression model it became possible to calculate a continuous quality measure for all of these groups. (More details on this procedure can be found in Appendix B.1). Regarding the *location* of flats, firstly we have divided the country into actual, interpretable neighbourhoods, among which 40 can be found in the capital, and 84 cover the rest of the country. Then – similarly to the method applied in the case of *state* – we estimated a cardinal quality value to each neighbourhood as well. The main principle in this process was to divide the country as homogeneously as possible from the point of view of pricing while keeping the number of observations sufficiently high everywhere to be able to calculate neighbourhood-level price indices. (More details on the creation of neighbourhoods can be found in B.1). Finally, flats have been grouped into *buckets* which are defined within each neighbourhood along *size* and *state* intervals.

The model also incorporates three different sources of change in the housing stock: *(i) amortization, (ii) renovation* and *(iii) new constructions*. The *state* of flats deteriorates every month at a predefined rate, which can be compensated by the households through renovating their apartment. The construction of new flats takes place by a representative company which builds flats with the highest possible *state* value.

2.2 DESCRIPTION OF THE MARKETS

2.2.1 HOUSING MARKET

Demand

The demand side of the housing market features not only households but also a representative investor and a construction company. These agents can have four distinct motivations for house purchases. (i) Young (first-time buyer) households want to buy flats to provide a home for themselves. (ii) Secondly, households appear on the demand side again if they decide to move. This can happen if the oldest member of the household reaches a certain age, or the household finds another flat which is sufficiently closer to their optimal choice than their current home. First-time buyer households can choose between flats in more than one region, but all the other households can consider flats only in their own regions[1]. (For further details of these mechanisms see Appendix A.4.4 and A.5.3.) (iii) Thirdly, the construction sector is also present on the demand side of the market. In order to build new houses, a building site proportional to the size of the new flats is required. For this, the construction sector purchases the cheapest flats (based on the square meter unit price) in the neighbourhood. (See Appendix A.5.1.) (iv) Lastly, households and the representative professional investor might want to buy flats to realize profit on these investments. This type of demand is influenced by the potential increase of the flats' market value and the obtainable profit from rents. Purchases for investment purposes generally happen in *buckets* with excess demand. The introduction of the central investor was important to represent all the entities whose decisions are not constrained by the Hungarian macroprudential rules, e.g. corporate investors (especially from abroad). Even if these actors are credit constrained, they are not influenced by the same limits Hungarian households are. Due to this asymmetry, LTV and DSTI regulation might be less effective if the market is at least partly driven by investment activities. Apart from the susceptibility to Hungarian macroprudential policies there is another aspect from which we differentiate between the two types of investors. Buy-to-let households can only buy one flat for investment purposes in a given period, thus, the outcome of their behavioral rule is binary, while the central investor determines the sum it wants to invest in each period. (More details on the investment-motivated flat purchases can be found in Appendix A.5.2.)

Supply

The same agent types appear also on the supply side of the housing market. First of all, moving households can sell their previous homes. However, this step does not need to necessarily precede the purchase of a new flat owing to the opportunity of taking out a bridge loan. Additionally, flats inherited after the death of relatives are also taken to the market. (Further details of inheritance can be found in Appendix A.3.1.) Flats purchased for investment purposes can also be taken to the market if the expected yield reduces. This can happen due to the lowering of house prices, or in the case of excess supply in a given *bucket*. (For further details see Appendix A.7.) Finally, the construction sector also sells new flats already during the construction period.

Pricing

Households make offers on one or more flats after evaluating them based on the available consumer surplus of the purchase. This is calculated using the utility functions of households which are calibrated individually to their preferences from empirical data about their actual homes. (See details on the calibration of the utility functions in Appendix B.5.) The utility function takes into account the three above described characteristics, *size, state* and *location*, furthermore, the financial situation of the household, and assigns a value – expressed in monetary terms – to each flat. (The exact formulation can be found in Appendix A.4.2.)

In the process of bidding, potential buyers choose a „perfect" fictive flat and then they can place bids on multiple flats. The closer a flat's consumer surplus is to the surplus of this ideal fictive flat, the higher the probability and the size of the offer will be[2]. The exact size of a household's offer is the weighted average of the reservation price and the market price. In the case of a tight housing market, households have to wait longer to find and purchase a suitable flat compared to normal market conditions. The longer they have been stranded on the market, the more tolerant they become regarding both the price and the characteristics of flats. (See details in Appendix A.4.2.) On the other side of the transactions sellers determine the market unit price of a flat by taking the two closest, recently sold flats in the neighbourhood. Every month the sellers set the ask price as a decreasing fraction of the market price. (See details in Appendix A.2.1.) At the end of the bidding process, the highest

[1] *Regions* in the model correspond to the 7 (+Budapest) NUTS 2 regions of Hungary.

[2] Before making an offer, households take into account also the cost of renovation.

bidder can buy the flat at the price of the second highest bid (i.e. by Vickrey auction)[3], or at the market price if there is only one bidder. If a household buys a flat, all the other offers of this household will be withdrawn.

2.2.2 CREDIT MARKET

The model features fixed and variable rate housing loans which can be used by the households to buy either their own homes or flats for investment purposes, provided that they have sufficient savings and stable income. This way, the size of the approved loans influences greatly the offers households can make on the market. In case of moving, households can also take out bridge loans, moreover, there are also personal loans to be used for renovation.

Apart from bridge loans, households' eligibility depends on three types of requirements: (i) macroprudential policy rules (i.e. LTV and DSTI regulations); (ii) the bank's own conditions regarding a minimum consumption level (i.e. the household's expected income must cover the credit payments and a minimal consumption level); (iii) and finally, a household is only eligible if it did not have a defaulting loan in the past five years[4].

The interest rate of a newly issued loan is calculated as the sum of the base rate and a spread. The bank determines the credit spreads with a regression model where the coefficients are estimated on actual empirical data (Appendix B.4). Furthermore, the credit supply of the bank also depends on macroeconomic factors, which leads to procyclical behavior on the credit market.

Since the members of a household can experience unemployment, or their income dynamics can be diverted by exogenous shocks, loans of the distressed agents might become temporarily or even permanently non-performing. In these predicaments, the bank first tries to restructure the loan in order to meet their payment obligation. If the payments can still not cover the interest calculated after the outstanding principal, we account a loss for the affected bank. If the household is continuously unable to deliver the restructured monthly installment, the collateral will be liquidated on a discounted price.

The credit market is discussed in more detail in Appendix A.6.

2.2.3 RENTAL MARKET

Flats can be taken to the rental market by the representative investor or buy-to-let households. We distinguish in the model between short-term (one period) and long-term renting. The short-term market segment is meant to represent the influence of online marketplace platforms (e.g. Airbnb), which mostly serve the demand coming from tourism. We represent this phenomenon by introducing time-dependent, exogenously given external demand for every bucket based on empirical data. (That is, this factor is completely independent from the preferences of the households.) This type of short-term renting is very concentrated geographically, as mostly only the inner districts of Budapest are affected by it directly. The increased amount of purchases for investment purposes in this area led to an upsurge in the real estate prices in 2018 and 2019; furthermore, as these flats have been mostly utilized for short-term renting, the rental prices also skyrocketed due to the massive outflow of flats from the long-term rental segment (Boros et al. (2018)).

Long-term renting corresponds to the traditional interpretation of the rental housing market. If a household does not have an own home, but needs a separate flat to live in, it can go to the long-term rental market[5]. Firstly, households look through the vacant flats in the rental market in their preferred region. Choosing between the flats happens similarly to the way described for purchases: households select the flat with the highest consumer surplus calculated as the difference between the rental reservation price (coming from households' utility functions) and the rental fee. After the expiration of the rental contracts

[3] While it is obvious that potential buyers in this case do not have full knowledge of all the bids – which would be a necessary condition for a plausible Vickrey auction design, this solution still have the very useful feature that under tight market conditions both the highest and the second highest bids will be very close to the ask price, but in the case of a loose market the gap between the bids will be larger representing the potential for bargaining (which mechanism is not present in the model otherwise).

[4] The third type of the requirements is based on the fact that banks can obtain the credit history of potential clients via the Central Credit Registry.

[5] Households in the model always try to purchase an own home. They only turn to the rental market if they do not succeed to do so. The main justification for this assumption is that Hungarian households have a very strong preference for owning their homes. More than 90% of the Hungarian households live in their own flat, and young households also pursue this strategy.

households either try to buy an own flat, or they can look at the rental market again. In the rental market there is no bidding process. Households come sequentially in a random order and they rent out the chosen flat for the ask rental price.

For further details see appendix A.7.

2.3 THE TIMELINE OF EVENTS

Our model simulations run through several periods, each of which represents a one-month time interval. One period consists of the following steps: (i) updating individuals and their characteristics, (ii) updating the market environment, (iii) transactions in the housing market, (iv) updating the housing stock, (v) updating the rental market and (vi) accounting of the money flows. The following subsections give a schematic overview of the events happening in each of these blocks of the simulations.

2.3.1 UPDATING INDIVIDUALS AND THEIR CHARACTERISTICS

- Life events of individuals
  - Individuals age and die (with a probability depending on their characteristics).
  - Young adults leave their parents' household and start their own households.
  - If there are no more adults in a household, the flat will be inherited.
  - Individuals marry and create new households. The selection of a spouse has been determined based on data coming from the Central Credit Registry.
  - Children are born in selected households. Children inherit the characteristics of their parents.
- The labour market status of individuals
  - The wages get updated. Wages depend on individuals' educational level, work experience and on the nominal GDP index.
  - Individuals become (un)employed and retired.
  - The life-time income of individuals gets updated.

2.3.2 UPDATING THE MARKET ENVIRONMENT

- Flats' market price gets updated.
- Households decide on moving. (If they have been trying to move for a certain amount of time, but they were not able to sell their own flat, they withdraw from buying a new flat.)
- The state of the flats deteriorates.
- Flats with expiring rental contracts become empty.
- The expected return is calculated on housing market investments in each neighbourhood.
- Rental market mark-up is calculated in each neighbourhood based on the respective vacancy rates.
- The construction sector's mark-up is calculated in each neighbourhood.
- The central investor determines the investment amount for each neighbourhood, furthermore, the investment probabilities for each bucket within the neighbourhoods.
- The probability of selling is calculated in each bucket in the case of flats owned for rental purposes (based on the expected return and the vacancy rate).
- Collecting the flats to be sold and calculating their ask prices:
  - The construction sector takes to the market (or keeps on the market) its flats to be sold.

  - Empty flats owned for investment purposes are taken to the market (or kept on the market) by a given probability.
  - Households take their inherited flats to the market (or keep on the market).
  - Movers take their previous flats to the market (or keep on the market).
- Assessing households' housing preferences by fictive flats
  - Fictive flats are generated (with realistic market prices).
  - Every potential buyer chooses a fictive flat which she would buy if it was an existing flat on the market. (This information is used as a reference for households' actual purchases, furthermore, the construction sector estimates the demand for newly built flats in each neighbourhood based on the fictive demand.)

2.3.3 TRANSACTIONS IN THE HOUSING MARKET

- The construction sector determines the number of flats it wishes to build in each bucket, and it buys flats (on their market price) in the corresponding neighbourhoods to ensure the necessary building sites.
- The central investor buys flats in each neighbourhood (if there are available flats) until it reaches the planned investment amount.
- Each household attempts to invest with a given probability: We randomly assign them a neighbourhood and then a bucket, in which we examine if there is a flat for sale which is possible to buy for the given household.
- Purchases of occupant owners take place in multiple rounds (where some of the households might not succeed on the market even after the last round). One round consists of the following steps:
  - Potential buyers evaluate the list of flats for sale and place unique bids (offers) on some of them.
  - If there is at least one valid offer for a flat, the household with the highest bid purchases the flat (at the price of the second highest bid, or at the market price if there is only one offer).
  - Sold flats are deleted from the list of flats for sale and buyers are removed from the list of potential buyers.

2.3.4 UPDATING THE HOUSING STOCK

- The construction sector continues the building of ongoing projects.
- Renovation
  - A random fraction of the households (and those who recently purchased a flat) decide on the extent they wish to renovate their flats.
  - Households pay the costs of the renovation (they can take out a loan for this if it is necessary), and the state of the flats will improve instantaneously.

2.3.5 UPDATING THE RENTAL MARKET

- Assessing households' renting preferences by fictive flats:
  - Fictive flats are generated (with realistic rental prices).
  - Every potential renter chooses a fictive flat which she would rent if it was an existing flat on the market. (This information is used to estimate the rental demand in each bucket in the next period, which determines the activity of investors.)
- The available (empty) flats for rent get collected (and their rental price is also generated).
- Short term renting: According to the exogenous demand, some of the available flats become rented out for one month in each bucket.

- Long term renting: Every potential renter choose a flat which is still available and moves there for a given number of periods.

2.3.6 ACCOUNTING OF THE MONEY FLOWS

- Individuals receive their wage.
- Renters pay the rental fee.
- Households determine and spend the amount allocated for consumption.
- Repayment of loans:
  - Households pay monthly instalments.
  - If they cannot pay, the given loan becomes non-performing.
  - After a certain time, the bank tries to restructure non-performing loans.
  - If a loan is still non-performing even after the restructuring, the collateral will be liquidated. (The liquidation starts in the next period).
  - Non-performing loans can become re-performing if the debtor household can resume the payment of the full monthly instalment again.
  - After paying the last instalment, the loan ceases to exist.
- The interest rate of adjustable-rate loans might change. The new amount of the instalments is applied from the next period.
- Housing market price indices are getting calculated.
- The average price of building sites are getting calculated.

3 Highlighted features of the model

This chapter presents several simulation results which provide justification for the unique features and the most important assumptions of the model.

3.1 CHARACTERISTICS OF FLATS

In the model, every flat has three attributes: *size, state* and *location*. While the *size* is constant over time (as it does not change even due to renovation), the *state* is time-variant, i.e. it changes with depreciation and renovation. If it remained unchanged, it would be impossible to model decisions about renovations. Furthermore, we assume that the neighbourhood quality characterizing each *location* is constant over time, however, the decision about time-invariance is not that obvious in this case. For example Ge (2017) uses the residents' income position to model this characteristic dynamically. In this paper, the ranking of neighbourhoods could vary over time if some neighbourhoods become more popular. However, these processes usually take several years, much more than the typical time horizon simulated in our model.

According to the results of our hedonic price regressions described in Appendix B.1, all of these attributes are highly significant, and they increase the goodness-of-fit considerably. To illustrate this, Table 1 shows the R^2 values if we used just two of the three flat features:

Table 1
R^2 of the hedonic price regression with different sets of flat attributes

Flat characteristics	R^2
Original estimation	0.84
Without "state"	0.75
Without "size"	0.54
Without "neighbourhood"	0.27

These results also suggest the outstanding importance of our neighbourhood quality variable. As it is described in details in Appendix B.1, our strategy regarding the formation of the neighbourhoods aims at creating a partition of the country which represents an optimal trade-off between minimizing the fragmentation of the data and maximizing the homogeneity within the categories[6]. Our data-driven solution for this is unique in the literature, and it has the advantage that house prices indirectly contain every relevant information about the location (public transportation, health and education institutions, residents' income situation, etc).

Table 2
R^2 of the hedonic price regression with different location variables

Location categories	R^2
None	0.27
Region	0.66
Counties	0.67
Neighbourhoods	0.84

[6] We formed 124 neighbourhoods with on average 80,000 inhabitants each.

Table 2 shows the performance of our neighbourhood level location variable compared to more coarse-grained options in the price regression estimation. According to these results, the estimation's R^2 decreases considerably if we use administrative location categories (like regions or counties) instead of our approach.

3.2 HOUSEHOLDS' DECISION MAKING PROCESS

During their decision making process about moving, households choose the flat with the highest consumer surplus (which is not necessarily the flat with the best quality) according to standard utility maximization theory. In order to do this, each household in the model has a reservation price function which assigns a utility to flats expressed in monetary terms, hence, for the sake of simplicity we call it a utility function. (The details of the exact form of this function and the decision making process can be found in Appendix A.4.2 and Appendix A.4.4.)

Since standard functional forms in consumer theory are used in a different spirit, we did not constrain ourselves to a CES-design, instead, we investigated alternative forms which meet the following requirements: (i) the reservation price function is continuous; (ii) the reservation price is a monotonically increasing function of a household's lifetime income, and of the size ($size_f$), state ($state_f$) and neighbourhood quality (Q_{N_f}) of a given flat f; (iii) the second derivatives with respect to $size_f$ and $state_{t,f}$ are negative.

We selected six different options, which differ either regarding their fundamental type (Cobb-Douglas, CES, exponential or sigmoid function), or whether they handle flat characteristics and the households' life-time income (I_h^{lt}) in an additive or a multiplicative way.

- Cobb-Douglas:

$$U_h^l = \delta_h \times size_f^{\alpha_h} \times state_f^{\beta_h} \times Q_{N_f}^{1-\alpha_h-\beta_h} \times I_h^{lt} \tag{1}$$

- Cobb-Douglas–Sigmoid:

$$U_h^l = \delta_h \times \left(size_f^{\alpha_h} \times state_f^{1-\alpha_h} + \frac{K_h}{(1+e^{-Q_{N_f}})^{1/\gamma_h}} \right) \times I_h^{lt} \tag{2}$$

- CES:

$$U_h^l = \delta_h \times \left[\alpha_h \times size_f^{-\beta_h} + \gamma_h \times state_f^{-\beta_h} + (1-\alpha_h-\gamma_h) \times Q_{N_f}^{-\beta_h} \right]^{-1/\beta_h} \times I_h^{lt} \tag{3}$$

- CES–Sigmoid:

$$U_h^l = \delta_h \times \left[\left(\alpha_h \times size_f^{-\beta_h} + (1-\alpha_h) \times state_f^{-\beta_h} \right)^{-1/\beta_h} + \frac{K_h}{(1+e^{-Q_{N_f}})^{1/\gamma_h}} \right] \times I_h^{lt} \tag{4}$$

- Exponential with additive size and state variables:

$$U_h^l = \left(c\_m_h \times size_f^{\beta_h^m} + c\_a_h \times state_f^{\beta_h^a} + \frac{K_h}{(1+e^{-Q_{N_f}})^{1/\gamma_h}} \right) \times I_h^{lt} \tag{5}$$

- Exponential with multiplicative size and state variables:

$$U_h^l = \left(c\_m_h \times size_f^{\beta_h^m} \times (1+c\_a_h) \times state_f^{\beta_h^a} + \frac{K_h}{(1+e^{-Q_{N_f}})^{1/\gamma_h}} \right) \times I_h^{lt} \tag{6}$$

To determine the best candidate, we tested the six different functions in the following way. We calibrated the six reservation price functions for 1,000 randomly sampled households by finding the parameter set for each household and for each function, with which we can get the highest consumer surplus for exactly the flat that the given household owns in reality (or for one which is very similar to it). (The details of this calibration procedure are described in Appendix B.5.) We assessed the goodness of the six functions based on their success rate in this exercise[7].

[7] The success rate is defined in the following way: We calculate d, which is the maximum of the percentage differences between two flats' price, size, state value and neighbourhood quality score, and we minimize the value of this function during the parameter search. We consider the calibration successful if the value of the objective function is less than d%. (The zero value implies that the household picked the same flat after the optimization as it did in reality.)

Table 3
Calibration performance of different reservation price functions

	Perfect match (d=0)	Successful calibrations (d<=0.1)	Successful calibrations (d<=0.2)	Successful calibrations (d<=0.3)	Number of parameters
A*X^B+sigmoid (multiplicative)	80	159	562	845	6
A*X^B+sigmoid (additive)	45	103	400	700	6
CES+sigmoid	19	36	104	249	5
CES	26	55	286	587	4
CD+sigmoid	85	161	498	713	4
CD	85	150	510	778	3

While Cobb-Douglas functions are the most widely used in the economic literature, and the more general CES functions are very popular in consumption theory, we can see in Table 3 that actually the exponential functional forms perform better, especially with multiplicative interaction between size and state. Additionally, we can also observe in the case of the CD and CES results, that in this non-linear, non-convex optimization problem more parameters do not necessarily yield better performance.

3.3 CONSTRUCTION SECTOR

In the existing housing market ABMs the construction sector is either not modeled at all (Laliotis et al. (2020)), or it is included only to a limited extent. For instance, Erlingsson et al. (2014) represented the construction sector merely as one of the industries producing capital goods, and Baptista et al. (2016) used this sector only during the burn-in phase to accumulate the sufficient housing stock which remained constant afterwards during the model simulation. We decided to model the construction sector in a more detailed way due to the following considerations:

- Prices and therefore the building up of bubbles on the housing market can be substantially influenced by the construction sector. On the one hand, exogenous or endogenous labor and material cost shocks fundamentally determine the price of newly built houses, which trickles down to the market of used flats as well. On the other hand, the increase of prices driven by excessive demand can be offset by the construction sector if they build houses in the right segments of the market. However, the effectiveness of this mechanism is highly dependent on the time requirement of the constructions and the availability of the necessary capacities. While the above mentioned papers acknowledge and somewhat utilize the construction sector as a means to adjust the housing stock, the constraints on building new housing stock after demand shocks on the market has not been taken into account so far.

- Our model features demographic phenomena such as the reducing birth numbers or the migration within the country. These mechanisms have heterogeneous implications across regions which results in different dynamics on the local housing markets. Thus, the assumption of time-invariant market conditions would lead to more and more distorted results as the time span of the simulation expands. To ensure the flexibility which is required to handle these processes from an economic point of view, one needs to elaborate also the mechanisms responsible for the adaptability of the market.

- As a more practical consideration, our model wishes to be applicable also for the analysis of the most recent tendencies in economic policy making. Since 2016, vast amount of resources have been allocated on the subsidization of Hungarian families' home purchases. At first, these programs focused on newly built flats, but later also the transactions of used flats became targeted. The most important elements of these subsidies are direct money transfers and very favourable loan constructions where some subsidized loans could even be accounted as part of the down payment of the housing loans. These policies affected immensely not only the housing loan markets, but also the housing market through several channels. To capture the implications in a plausible way, the detailed representation of the construction sector has become inevitable in the model.

The importance of the construction sector is also shown by the fact that in the period between 2018 and 2019 13% of the transactions on the Hungarian market concerned newly built flats, which accounts for the renewal of 0.4-0.5% (depending on the year) of the total housing stock in the country. Furthermore, the amount of subsidies in the same period were equal to 7% of the newly issued loans.

3.4 THE SIZE OF THE MODEL

According to our knowledge, there is no other paper in the relevant literature, which analyses the housing market of a whole country by representing all the households and also the housing stock in its entirety. For our purposes however, it was necessary to create the model in this fashion as this design contributed greatly to the accuracy and granularity of the simulation results in a threefold way.

(i) As discussed above, in order to obtain sufficiently homogeneous divisions based on empirical data, we had to divide the country into 124 neighbourhoods. Since the price formation of the offered flats on the market is based on the transactions of the previous periods in the given area, the number of households making deals has to be high enough in each neighbourhood.

(ii) Secondly, by representing the whole market it becomes possible to analyse the effects of various economic policies at more disaggregated levels. This way, we can examine relatively small groups of households, such as low income households, families with many children, first-time home buyers, etc.

(iii) Additionally, there is also a third justification for the high resolution strategy, namely, that the model is not scale invariant. If we use fewer agents, the long term averages of the main time series generated by the model will change considerably. To demonstrate this effect, we ran the *sample model* using only simulated data with 50,000 and 500,000 agents in a fictive economy. In the results, we show the averages of three realizations for both parameterizations. One realization consists of 3,000 periods, out of which we ignored the first 500 to exclude burn-in effects.

Table 4
Model simulation results for different model sizes

Number of individuals	500,000	50,000
Avg. house prices (million HUF)	21.1	15.2
House price autocorrelation	0.94	0.04
Avg. # of transactions per month	955	33
Transaction number autocorrelation	0.77	0.33
Ratio of transactions / individuals	0.19%	0.07%
Ratio of transactions / flats for sale	27.2%	1.1%

According to the results in Table 4, the long term average house prices decreased by 28% and the number of transactions on the housing market dropped by almost 97% in the case of 50,000 agents in the model compared to the scenario using 500,000 agents. While the first order autocorrelation in the prices is 0.94 for the larger model, it is only 0.04 for the smaller version. We can observe similar pattern – although to a lesser extent – in the autocorrelation of the number of transactions.

The main driver of this level of scale dependence is the agents' behaviour regarding the house selection mechanism. Buyers would place a bid with a high probability only on flats with which they can obtain similar consumer surplus as that of the best fictive flat. If there are fewer agents in the economy, there will be fewer flats on the market, which leads to lower probabilities for the households to find a suitable home, i.e. frictions on the market become more severe. This phenomenon is well illustrated by the fact that with 500,000 agents in the model 27% of the flats for sale can actually be sold, while this number drops astonishingly to 1% for the model with 50,000 individuals. The difference in the autocorrelation is driven by the same mechanism: with fewer households in the model there is a larger variance regarding the type of flats which can be sold on the market.

As it has been clearly shown by this example, the scale dependence can lead to serious distortions in policy analyses if the number of agents differs too much from the actual size of the system. In these situations, the effect of the policy shocks can be completely dominated by the amplified frictions, which would render the model unserviceable.

3.5 CREDIT MARKET

Housing market models usually do not include detailed credit markets, only standard housing loans typically with long maturity. In contrast, we propose two further loan contract types which are available for the households in the model: (i) personal loans which can be used for renovation purposes and (ii) bridge loans for households wanting to buy a new flat and sell the old one simultaneously. Personal loans make it possible to carry out renovations even for those who would not be able to afford it without a loan[8], and bridge loans can greatly mitigate the frictions on the market as it is no longer necessary to sell the old flat before buying the new one in case of moving. Even though this problem can have a large influence on the timing of transactions, it is usually not handled realistically in the existing models.

The inclusion of these additional loan types has an important contribution also to the model's suitability to analyse macroprudential regulation. Since in reality households can have multiple loans, one can represent their indebtedness, and this way also systemic risks more accurately, which enhances the precision when measuring the effects of macroprudential policy changes. This line of reasoning leads to another novel feature of our model: households have to comply with the actual, complex LTV and DSTI regulatory rules in order to take out a loan. In accordance with the Hungarian regulation, the DSTI limit can have six different values depending not only on the income of the given household, but also on the length of the interest rate fixation[9]. The higher the income is and the longer the interest rate is guaranteed to remain unchanged, the higher the DSTI limit will be. Since the heterogeneity of the households in the model allows for these types of differentiation, we can analyse the regulatory tools with all the details present in reality.

There is another important macroprudential consideration which we incorporated in the behaviour of the credit market in the model. After the 2008 crisis, the procyclicality of the banking system played a central role in the formation of the new regulatory framework. We introduced this aspect into the model by implementing a rule which makes the bank tighten its lending standards if the unemployment rate increases in the economy. To examine the cycle-amplifying effect of this attitude of the bank we calculated the average difference between 10-10 model simulations with and without this rule. (See Table 5.)

Table 5
Average differences between model variants with and without procyclical behaviour rules in the banking system

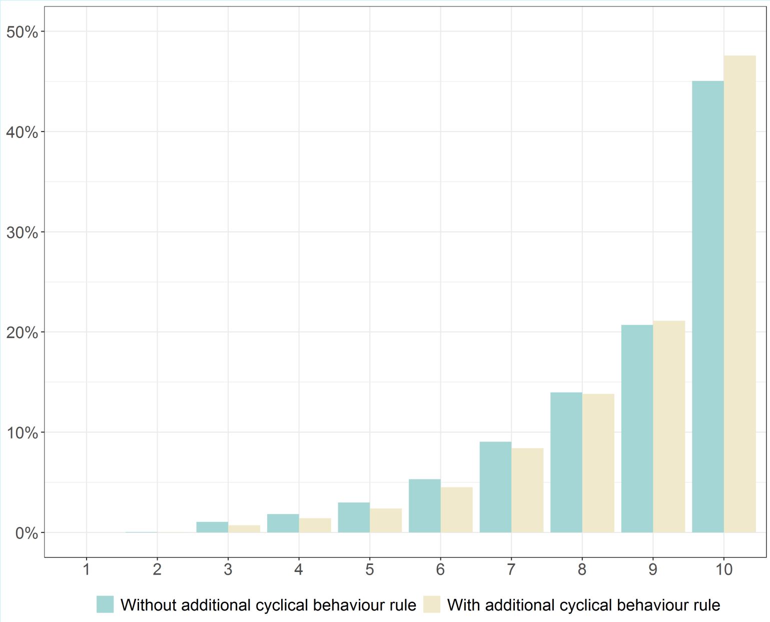
	Difference in the # of transactions	Difference in the new housing loan volume	Difference in the default rate (pp)
2018	0%	0%	0.0
2019	0%	0%	0.0
2020	-5%	-7%	0.0
2021	-2%	-5%	-0.1
2022	1%	2%	-0.1
2023	1%	1%	-0.1
2024	0%	0%	0.0

After the economic growth in the period between 2018-2019, the COVID-19 shock caused a major decline in 2020. Regarding the period after 2021, we used the forecast of the Central Bank of Hungary, according to which the economy will start growing again from 2022. As it is shown in Table 5, the number of transactions and the volume of newly issued loans in the depression period between 2020-2021 are considerably lower in the model simulations with stricter lending standards. Despite the correction in the following few years, we can clearly observe reduced lending activity during the time span of the simulation due to the procyclical behaviour of the banking system. Additionally, the lower amount of newly issued loans also entailed lower default rates until the end of the simulation's time horizon.

[8] In Hungary the majority of the loans used for renovation purposes are unsecured loans.

[9] The DSTI regulation in Hungary is also differentiated based on the denomination of the loans, however, as there are no loans in the model denominated in foreign currencies, this aspect of the regulation is not relevant.

Figure 2
Distribution of newly issued loans along households' income deciles in the model variant with procyclical banking system.



Furthermore, we could analyse not only the volume of the new loans, but also their distribution along the income deciles of the households. According to Figure 2, the tightening of the credit supply affects more heavily the households with lower income level.

3.6 CYCLICAL BEHAVIOUR

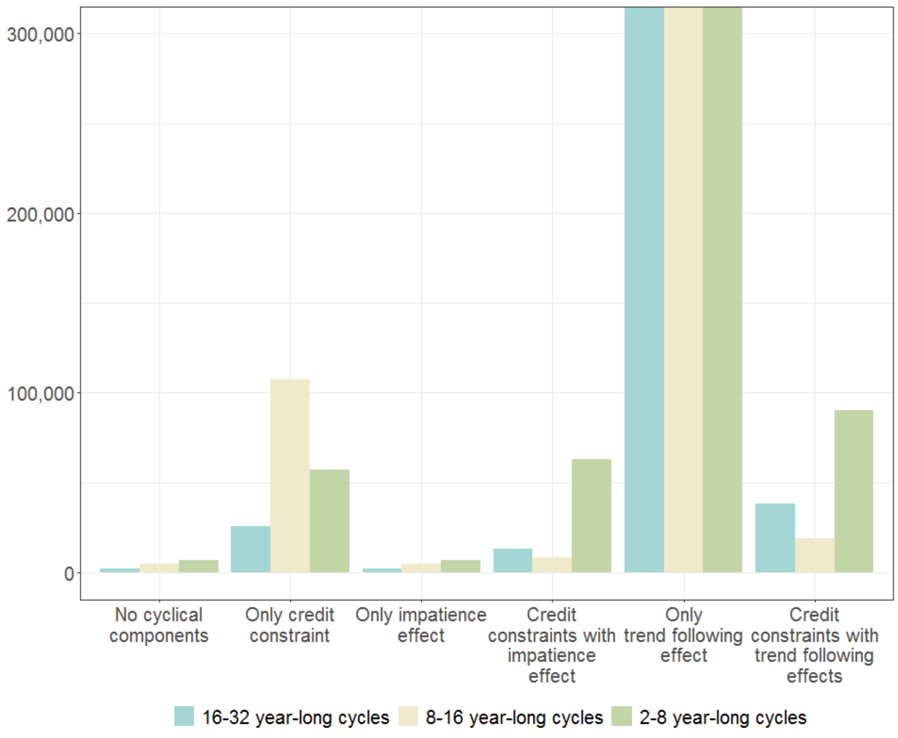
The possibility of cycles' emergence in the housing market is ensured by two mechanisms in our model: (i) credit constraints and (ii) the impatience of the buyers. In the case of a tight housing market, some households might have to wait longer to buy a suitable flat. The more they have to wait, the more tolerant they become regarding the price and the characteristics of flats. This behavior exerts upwards pressure on the prices, however, after a while this effect will be overpowered by the more and more often binding credit constraints.

According to our knowledge, the introduction of *impatience* is a novel approach to induce cyclicality on the housing market. The underlying intuition for this mechanism is based on the concept of the *cost of waiting* and the uncertainty of this cost in the future. If a household is not able to buy an adequate flat, it has to spend on renting, moreover, it also has to face with the uncertainty that house prices might increase faster than the income of the household. Since households have only limited knowledge about the tightness of the housing market and the price dynamics, they cannot predict accurately the expected *waiting costs*. However, if they turn out to be unsuccessful during multiple rounds of bidding on the market, we assume in the model that they experience and anticipate higher waiting costs, therefore they will be willing to buy a flat even with relatively low consumer surplus. A more detailed description of similar behaviour patterns can be found in Mell et al. (2021).

Due to computational constraints, we can demonstrate the implications of these rules by running the *sample model*. In this environment, we also have the opportunity to compare our strategy for generating cyclical behaviour to the assumption of *trend follower* behaviour, which is an often applied strategy in similar models. (See e.g. Baptista et al. (2016).) This is built on the assumption that households are willing to spend more (less) on housing if the prices increase (decrease) on the market, which strengthens the trends in the price dynamics.

We examined the strength of the cyclical behaviour of the model in six cases:

Figure 3
Results of the periodogram analysis of cyclical patterns in the time series of house prices coming from our model simulations with different cycle generating mechanisms



1. We assumed that households have very high savings, which makes them immune to any constraints on the credit market, and additionally we turned off the *impatience* trait of the agents. These modifications should eliminate all the cyclicality from the simulation results.

2. We left the *credit constraints* in place, but eliminated the *impatience* of the agents.

3. We eliminated the *credit constraints*, but left the *impatience* factor in the model.

4. Both the *credit constraints* and the *impatience* are present.

5. There are no *credit constraints*, but we introduced the *trend following*, i.e. herding mechanism.

6. Combined *credit constraints* and *trend following* attitude.

We ran the model for 3,000 periods, out of which the first 1,500 were ignored to exclude burn-in effects. We examined the cyclicality in the time series data of house prices coming from the remaining 1,500 periods with periodogram analysis. Figure 3 shows the presence of cyclical patterns in the above listed six cases. The higher the values are, the stronger the presence of cyclicality is with the given cycle length. One can see that there is only very low cyclicality in the time series without cyclical components or with only impatience effect in the simulation, however, in all the other cases the values are considerably higher. If there are no credit constraints, households are able to buy their optimal flat, they do not have to wait for it, thus impatience effect can not influence them. In contrast, if we consider the trend follower attitude without credit constraints, idiosyncratic shocks generate long cycles with high amplitudes, because households could pay also very high prices for flats. Enabling only the credit constraints increased the presence of cycles with longer examined lengths (8-16 and 16-32 years). This finding is in line with the results in the literature focusing on financial cycles, where several paper estimates the typical cycle length to be in this interval (Drehmann et al. (2012), Borio (2014), Aikman et al. (2015), Dias (2017), Hiebert et al. (2018)). If we combine the *credit constraints* with the *impatience* mechanism, the results clearly indicate stronger cyclicality compared to the scenario with only impatience. This is due to the interaction between the two effects: if credit constraints are present, households have to wait more often and for longer time to find an optimal flat, which enhances the effects of the impatient attitude of the buyers. In contrast, if we consider the *trend follower* attitude, credit constraints mitigate its cyclical effect because households cannot afford to spend that much on house purchases. Interestingly, when the credit constraints are present, the cyclicality is very similar regardless of whether it is paired with the *impatience* or the *trend follower* attitude. The case with *impatience* has slightly lower values for all the three cycle lengths. and both behaviour rules decrease the role of long cycles, while business cycles become more important.

4 Model diagnostics

4.1 CALIBRATION RESULTS

The model has been calibrated to the following four empirical time series: (i) regional house prices, (ii) the quantity of newly issued housing loans, (iii) the number of transactions of the housing market and (iv) the number of transactions of newly built flats and houses. We could observe these variables between 2018Q1 and 2019Q4 except from the number of transactions which was only available until 2019Q2. We compared these data to the average results of the model calculated using ten simulation runs[10]. It is important to highlight that by using the 1:1 representation of many parts of the economy our model starts from an initialization close to the steady-state of the considered markets. This way, it was possible to reduce the burn-in period of the model to merely one or two quarters. However, because of this, our results for 2018Q1 are still somewhat distorted due to the remaining burn-in effect.

Table 6
Quarterly average house prices in 2018 and 2019 in the model and in the empirical data (in million HUF)

Period	Budapest		Pest county		N. Hungary		N. Great Plain	
	Data	Model	Data	Model	Data	Model	Data	Model
2018Q1	28.5	25.8	21.2	17.1	7.1	6.5	9.1	8.6
2018Q2	30.7	29.6	21.4	20.4	7.2	6.4	8.6	8.9
2018Q3	31.3	30.5	22.2	21.1	7.2	6.5	9.3	9.1
2018Q4	34.4	31.6	23.2	22.9	7.7	7.1	9.7	8.4
2019Q1	35.3	33.3	25.9	24.3	8.3	7.3	10.9	7.8
2019Q2	36.9	35.1	25.6	25.3	8.4	7.9	10.9	7.6

Period	S. Great Plain		C. Transdanubia		S. Transdanubia		W. Transdanubia	
	Data	Model	Data	Model	Data	Model	Data	Model
2018Q1	9.5	9.0	12.4	12.2	10.6	11.2	14.8	12.0
2018Q2	9.5	9.6	13.1	14.9	10.5	11.1	14.9	13.6
2018Q3	9.8	9.7	13.6	17.3	10.9	11.1	15.1	14.2
2018Q4	10.0	9.9	13.9	19.0	11.6	11.0	15.7	14.8
2019Q1	11.2	10.5	15.6	20.0	12.4	10.9	17.0	14.8
2019Q2	11.3	11.2	16.6	20.7	12.9	11.1	16.9	14.7

Source: MNB.

In the case of the regional house prices in 2018Q1, we could calibrate the model to produce results which are less than 10% different compared to the empirical data for six out of the eight regions, while the largest difference was 24%. The ranking of the regions is also in line with the actual data, furthermore, the model was able to match the dynamics in the increasing prices in most of the regions (see Table 6). These results are outstandingly accurate for the two most important regions, Budapest and Pest county, in which more than one third of all the transactions happen. In contrast, the models' output is somewhat further from the empirically observed data in the case of the Northern Great Plain and Southern Transdanubian regions where also the number of the transactions is the lowest in the country.

[10] We decided for using the average of ten runs because it proved to be an optimal choice regarding the trade-off between running time and the accuracy of the results. For only five runs we experienced much greater instability, however, if we increased the number of runs above ten, the results remained practically unchanged.

Table 7
Gross flow of housing loans in 2018 and 2019 in the model and in the empirical data (in billion HUF)

Period	Data	Model
2018Q1	172	99
2018Q2	232	202
2018Q3	243	199
2018Q4	206	194
2019Q1	203	197
2019Q2	253	201
2019Q3	222	202
2019Q4	233	208

Source: MNB.

Table 8
Number of transactions in the housing market (total and newly built) in 2018 and 2019 in the model and in the empirical data

Period	Data (Total)	Model (Total)	Data (Newly built)	Model (Newly built)
2018Q1	44,015	42,562	6,496	3,366
2018Q2	49,890	44,626	6,627	6,723
2018Q3	49,023	44,775	6,432	6,173
2018Q4	41,722	45,326	6,768	5,515
2019Q1	44,396	48,325	4,245	5,319
2019Q2	45,746	49,035	4,827	4,511

Source: MNB.

Considering the quantity of newly issued housing loans in 2018-19 (Table 7), we can see that the model predicts altogether 1501 billion HUF, which is close to the actual 1,763 billion HUF. The quarterly averages in the model show around 13% difference compared to the empirical numbers (except for the first quarter which is dominated by the burn-in effect).

Regarding the number of transactions of the housing market we can observe the missing quarterly fluctuation in the predictions due to the lack of seasonality in the model (Table 8). The number of transactions is lower in the model for the first three quarters of the examined period, but the situation is reversed afterwards. However, the distance from the actual values is less than 15% in all quarters, while it is only 4% if we consider the whole 1.5 years period.

The number of transactions on the market of newly built flats is also very close to the data. After the first quarter the mean quarterly deviation is only 2.6%. Additionally to the seasonal effects, the empirical numbers have been influenced by the change in the VAT rate and by the introduction of several family support policies, the impacts of which were not modeled in the simulations.

4.2 VALIDATION RESULTS

After the calibration, we also performed several validation steps by comparing the model's outputs with time series which were not used during the calibration process. Since generating the households and their demographic and financial characteristics have been based on empirical data – coming from the Hungarian Central Statistical Office and the Central Administration of National Pension Insurance – socio-economic traits cannot be the basis of the validation. Furthermore, we could not use several out of the variables describing the housing market as the total number of transactions and the regional prices as these pieces of information have been already involved in the calibration process.

Table 9
Number of transactions in 2018 and 2019 in the model and in the empirical data

Period	Budapest		Pest county		N. Hungary		N. Great Plain	
	Data	Model	Data	Model	Data	Model	Data	Model
2018Q1	10,937	9,592	4,845	6,372	4,346	3,445	5,805	4,800
2018Q2	11,106	11,141	5,822	6,617	5,562	3,466	6,705	4,623
2018Q3	10,658	10,910	5,586	6,120	5,547	3,674	6,824	4,578
2018Q4	9,749	10,572	4,594	6,230	4,418	3,720	5,578	4,722
2019Q1	10,671	10,612	5,097	6,571	4,688	4,056	5,771	5,146
2019Q2	10,222	10,708	5,321	6,517	5,040	4,173	5,946	5,156

Period	S. Great Plain		C. Transdanubia		S. Transdanubia		W. Transdanubia	
	Data	Model	Data	Model	Data	Model	Data	Model
2018Q1	5,476	5,131	4,662	4,470	3,764	3,555	4,180	3,984
2018Q2	6,295	5,495	5,288	4,349	4,618	3,551	4,494	3,994
2018Q3	6,204	5,645	5,196	4,326	4,497	3,754	4,511	4,067
2018Q4	5,087	5,862	4,364	4,383	3,774	3,969	4,158	4,227
2019Q1	5,441	6,406	4,699	4,726	3,929	4,164	4,100	4,560
2019Q2	5,766	6,430	4,916	4,935	4,304	4,167	4,231	4,584

Source: MNB.

However, the regional transaction numbers are still usable as a means of validation (see Table 9). Based on this, we can see that the model was able to capture the outstanding role of Budapest, and it matches the actual numbers in each quarter with an error smaller than 10%. The results are also promising for the rest of the regions, where the difference between the model's output and the actual numbers is usually only a couple hundred transactions. We could detect only one minor systematic bias in the results. The model tends to underestimate the transaction numbers for the less developed areas (Northern Hungary and Northern Great Plain), and overestimates them for one of the most developed region (Pest county).

Unfortunately, there is not many empirical data at our disposal about the transactions on the housing market, however, we could perform one additional comparison which concerns with the average neighbourhood quality of the flats involved in the transactions in the different regions (see Table 10). In this regard, the model is able to match the average neighbourhood quality with a difference less than 0.1 in four cases out of the eight regions, while for two other regions this difference is around 0.2. The ranking of the regions from this aspect is also correctly reconstructed by the model with the exception of one region (Northern Great Plain), which has slightly lower average neighbourhood quality value in reality.

As the model produces several outputs related to the credit market, we could also use these results for validation. For instance, we can observe both in the model and in reality the ratio of transactions for which the buyers used bank financing (see Table 11). Apart from the first two quarters – during which the burn-in effect could distort the result to some extent – we see only 1-3 percentage point differences. Together with the calibration of the newly-issued loans, this suggests that the model can capture accurately the aggregate credit flows for housing purposes, which is a very important result from a macroprudential point of view.

The agent-based approach makes it also possible to compare the distribution of the indebted households along several characteristics to the empirical data. Firstly, we considered the distribution based on the income deciles of the households (see Figure 4). The model is in line with the empirical observation that households in the higher income deciles have much larger share on the credit market (especially the highest decile) than the lower deciles. While the difference between the model and the data is very small (1.4 percentage points on average), there is only one category – the 9th percentile group – for which we can see a slightly more considerable, 4.5 percentage points deviation.

Table 10
Average neighbourhood quality of flats in transactions occurred between 2018 and 2019 in the model and in the empirical data

Region	Data	Model
Budapest	7.0	6.9
Pest county	3.8	3.4
Northern Hungary	1.9	1.9
Northern Great Plain	2.3	2.7
Southern Great Plain	2.2	2.4
Central Transdanubia	2.8	2.8
Southern Transdanubia	2.4	2.4
Western Transdanubia	3.2	3.0

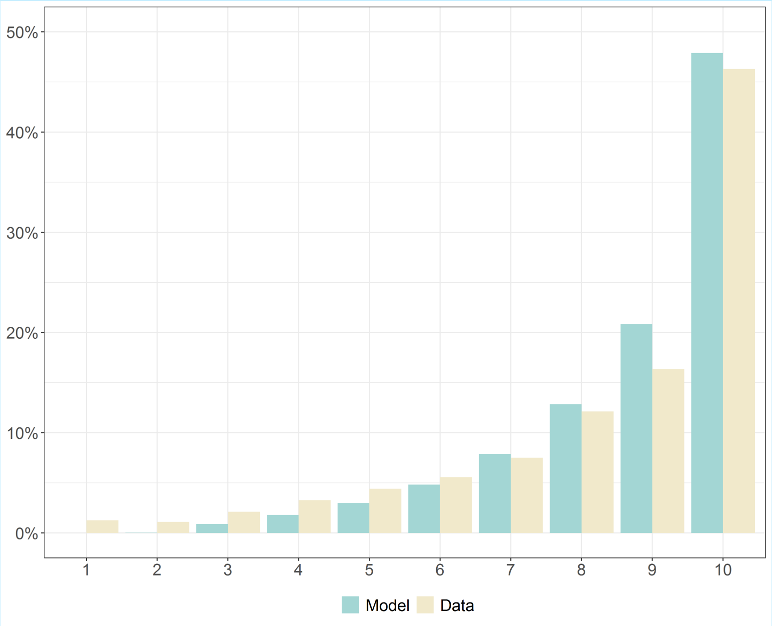
Source: MNB.

Table 11
Number of newly issued housing loans / Number of transactions in 2018 and 2019 in the model and in the empirical data

Period	Data	Model
2018Q1	41%	31%
2018Q2	43%	47%
2018Q3	45%	46%
2018Q4	46%	44%
2019Q1	46%	44%
2019Q2	46%	43%

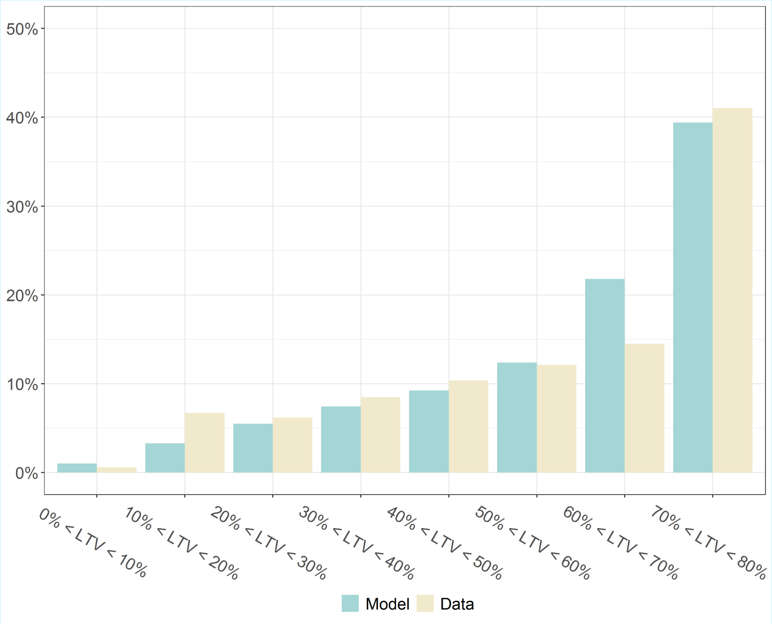
Source: MNB.

Figure 4
The distribution of the volume of newly issued housing loans in 2018 and 2019 based on the income deciles of the households.



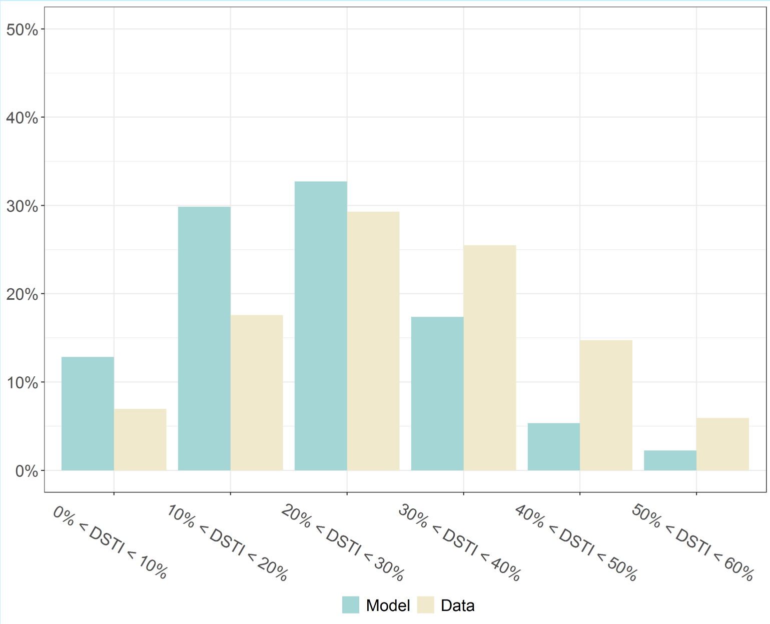
Source: MNB.

Figure 5
The distribution of the volume of newly issued housing loans in 2018 and 2019 based on the loan-to-value (LTV) ratio of the loans.



Source: MNB.

Figure 6
The distribution of the volume of newly issued housing loans in 2018 and 2019 based on the debt service-to-income (DSTI) ratio of the loans.



Source: MNB.

The model predicts similarly well the distribution of newly issued housing loans based on the loan-to-value (LTV) ratio of the households having loan contracts (see Figure 5). We can see that the model could capture the fact that around 60 % of the loans belong to the two highest LTV deciles. The highest decile has far the largest weight, while the lower a decile is, the less its weight is. This implies that the model is capable of accurately measuring the risks stemming from not being completely covered by collateral, and also the varying extent to which the LTV regulation is binding for the households.

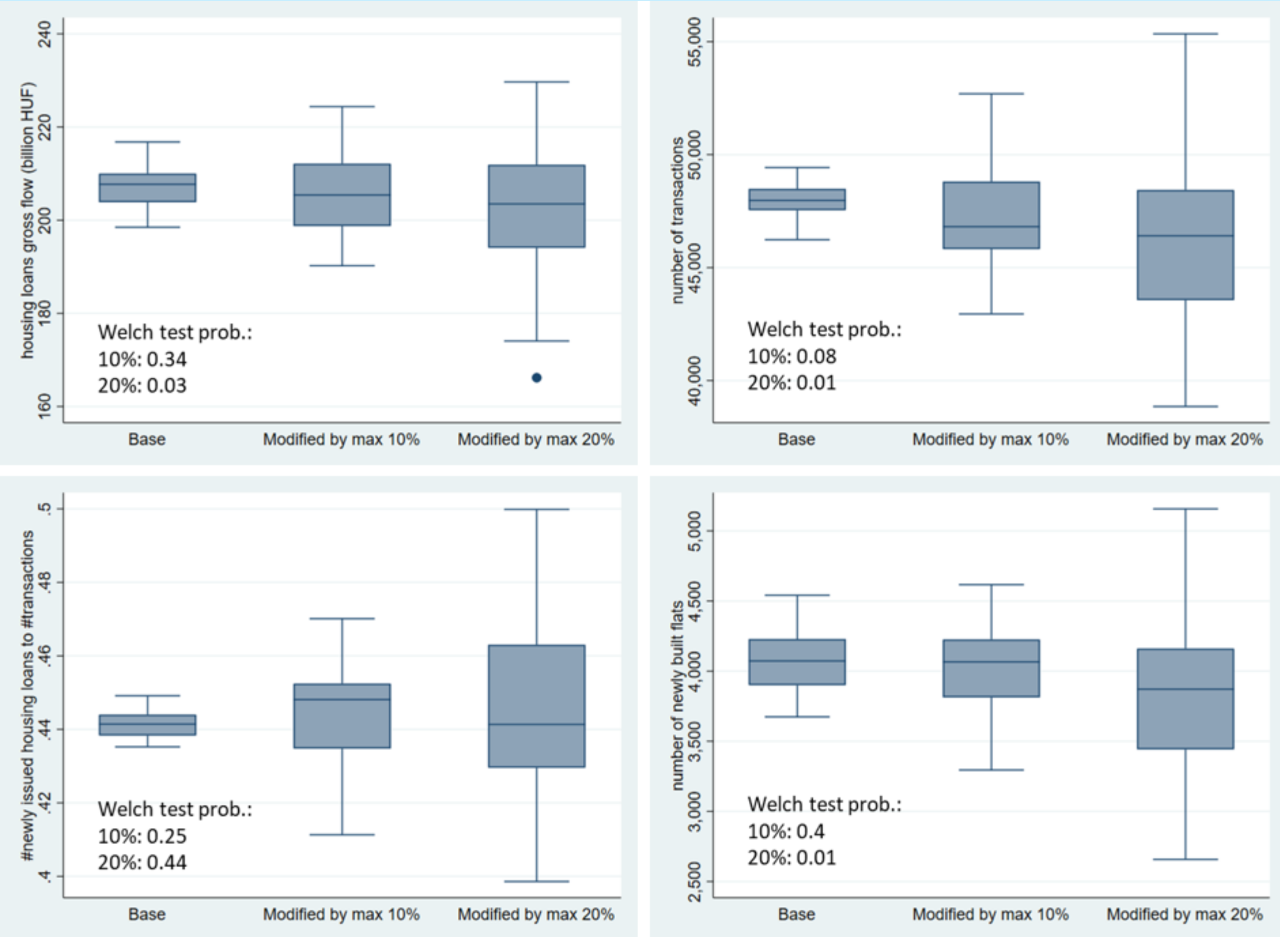
The fit of the model is less convincing in the case of the debt service-to-income (DSTI) distribution of the loan contracts (see Figure 6). The main reason for this is that the empirical data contains the payments of all the loan types while we only modeled housing loans and personal loans which can be used for renovation purposes. This leads to smaller DSTI values in the model which also results in the underestimation of default probabilities. This caveat could only be corrected by modelling the consumption behaviour of the households in a more detailed way, however, this extension would go far beyond the scope of a housing market model[11].

4.3 SENSITIVITY ANALYSIS

Although the reliability of calibrating a large number of parameters is a focal point in the criticisms against the ABM paradigm in economics, in the case of most large-scale ABMs in this literature there is no emphasis put on sensitivity analysis (Borgonovo et al., 2022). An elaborated exercise in this pertinent matter could not only enhance the credibility of the results, but it also makes it possible to assess the contributions of the different mechanisms to the outputs of the model. However, computational constraints – especially for large, data-driven models – pose a significant barrier in the present. In this section, we attempt to overcome this obstacle and provide a detailed sensitivity analysis for our model.

[11] In an extended version of our model, we also take into consideration a special consumption loan type (prenatal baby support loans), which takes nearly the half of the new outstanding consumption loans from June 2019. The distribution generated by this version of the model was much closer to the empirical distribution.

Figure 7
Box plots of 40 realizations of four highlighted output variables of the model with respect to three parameter sets: original values, 10% perturbation, 20% perturbation.



4.3.1 ROBUSTNESS ANALYSIS

Since our model uses 60 calibrated parameters, it was not possible to examine their robustness one by one. As an alternative strategy, we tested the sensitivity of the model's output to these parameters simultaneously in a perturbation exercise. For all of these parameters, we modified their values using random numbers drew firstly from the $U(-10\%, +10\%)$, then from the $U(-20\%, +20\%)$ uniform distributions. We compared the results of 40 runs generated with the original parameters to the results of 40-40 realisations with these random perturbations. To evaluate the differences among the results of these three different specifications, we compared the median and also the mean values of the variables used in the calibration and validation phases. Our measurement strategy consists of three techniques: (i) visualization of the distribution of the model outputs using box plots for the three different cases, (ii) hypothesis testing for the equality of the median values of the output variables across the three specifications and (iii) hypothesis testing using Welch's test for the equality of the mean values of the output variables across the three specifications. While the median test considers the equality of the three cases simultaneously, the Welch's test is only suitable for pairwise comparisons. The main advantage of Welch's test over the t-test is that it is robust to the different variances of the distributions out of which the three samples have been generated. According to our knowledge, this is the first attempt in the literature of large-scale economic ABMs to perform a sensitivity analysis in this fashion.

Figure 7 shows the results for four of the most important output variables of the model, while the rest of the outcomes of the robustness analysis can be found in Appendix C.1. The results of the robustness analysis are in accordance with the lessons of the calibration and validation processes, i.e. that usually the variables for which the calibration was less successful also proved to be less robust to perturbations and showed larger variance across realisations. A typical example for this situation is the average prices and the number of transactions in the least developed regions of Hungary. Nevertheless, according to the box plots there are only very few such problematic variables. In the case of the realisations generated with the original parameters, the difference between the minimum and the maximum values of the box plots are less than 10% for most of the cases, and it almost never exceeds 20%. For the perturbed scenarios this difference is higher, but it increases only by a smaller factor than the maximum level of the perturbation.

In the case of the hypothesis tests, we found that the first two quarters (2018Q1-Q2) are – similarly to the calibration – heavily influenced by the burn-in effect. In these periods we can almost always reject the hypothesis that the means and also the medians are identical across the three parameter scenarios. Similarly, in the last two quarters of the time horizon (2019Q3-Q4) the modifications to the parameters cause larger and larger deviations across the model outcomes in the three cases. However, in the remaining part of the time horizon (and especially in the case of the variables describing the loan distributions based on LTV, DSTI and income deciles) the results indicate on all of the usual significant levels that the three cases produce identical outcomes. The highest match between the three cases happens in the period between 2018Q4 and 2019Q1: in the case of 20 out of the 29 variables in this exercise we found that the means (and usually also the medians) are statistically identical between the runs with the original parameter set and the runs with 10% perturbation of the parameter values. In the case of 20% perturbation we got basically the same results for the 2018Q4 period. We can overall observe that the increase in the perturbation from 10% to 20% did not lead to larger divergence in the means of the examined outcome variables, and the robustness of the most important model outputs proved to be relatively high in the case of the simultaneous perturbation of the model parameters.

4.3.2 PARAMETER IMPORTANCE ASSESSMENT

As the next phase of the sensitivity analysis, we investigated the individual importance of the parameters on the model outputs. Our strategy for this consisted of two steps: (i) In order to reduce the complexity of the task, we filtered and grouped together the parameters into 4 blocks. (ii) In the second step we conducted a variance-decomposition analysis based on the Sobol method for the four groups of parameters obtained in step (i).

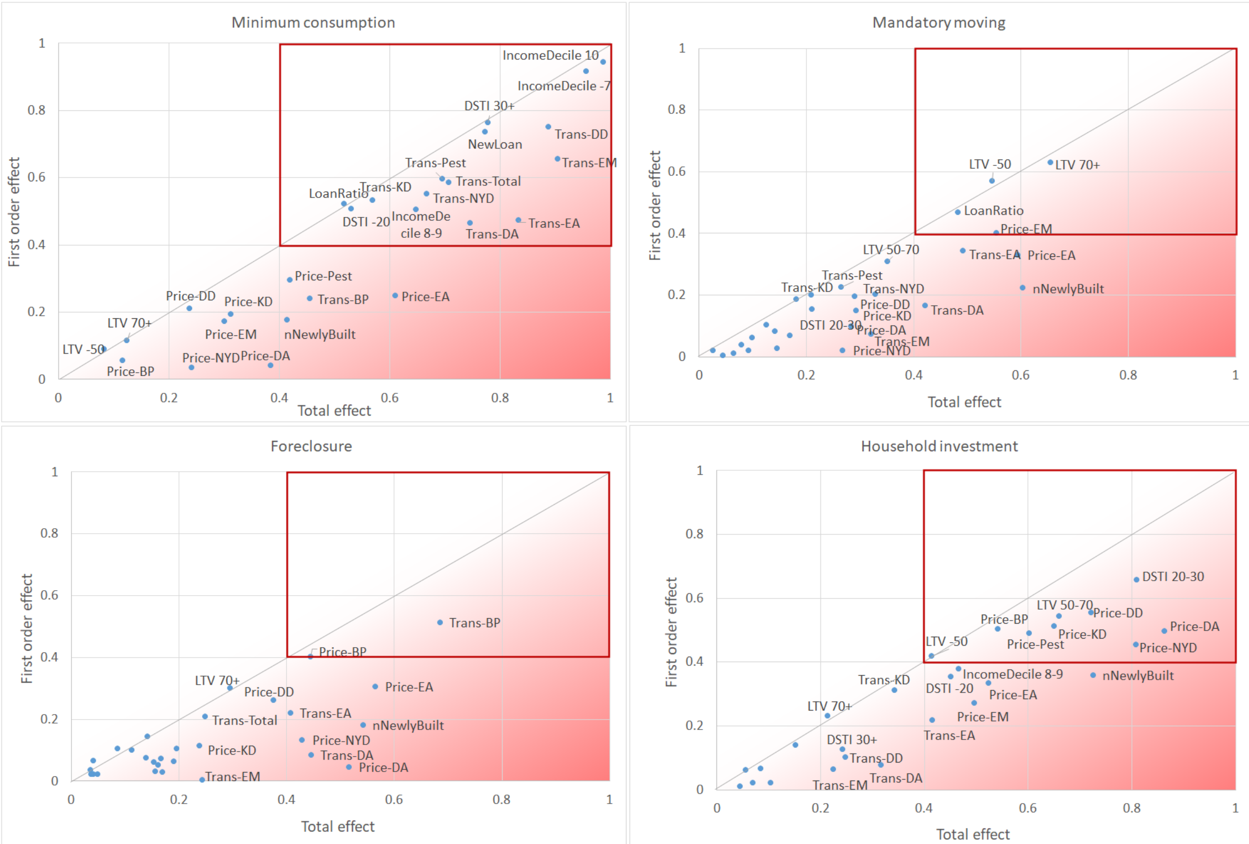
For the first step we could utilize the data generated by the robustness checks of the previous section to carry out a regression analysis. The 40-40 runs with perturbed parameters provided us with a set of dependent variables (the 29 output variables of the simulations) and a set of explanatory variables (the values of the 60 parameters)[12]. To determine the importance of the parameters, we counted the number of regressions in which a given parameter had significant coefficient, and we narrowed down the list of parameters to 12 based on this ranking. (The significance level did not influence materially the results of this step. Appendix C.2 contains the results for 5% significance level.) The first parameter on this narrowed list was significant in 21 regressions, while the last one was significant in 8. There were 34 parameters in the original list which had a significant coefficient in less than five regressions. These 12 parameters could be divided into four categories: parameters governing (i) the minimum consumption level of households, (ii) the mandatory moving behaviour of households, (iii) foreclosure in case of defaulting loans and (iv) households' purchasing for investment purposes. For the detailed parameter description see Appendix C.2.

In the second step, we conducted a Sobol-type variance decomposition in which we could separately measure the first-order effects (FOE) and the total-order effects (TOE), which takes into account indirect channels and interactions as well. For both of these calculations we used the estimator method introduced by Jansen (1999). Similarly to the robustness analysis, we ran our model with maximum 20% perturbed parameter values, however, in this case we modified the parameters belonging to the same group in a synchronized way to keep them consistent with each other. Furthermore, we did not modify the parameters which were filtered out in step (i).

The required number of simulations for this decomposition technique depends of course on the model and also on the number of parameters, but it is considered to be sufficient if the results converge. To increase the accuracy of the results one has to

[12] In the cases where we observed more than one time periods, we ran panel regressions. In these regressions we also included the one period lagged value of the dependent variable on the right-hand side. For the nine variables representing the LTV, DSTI and income categories we used pooled regressions.

Figure 8
Scatter plots of the first-order and total effects of the four highlighted parameter groups.



double the number of runs each time. Given our computational capacities the maximum number of simulations to complete the FOE and TOE calculations was 1,530, which resulted in sufficiently reliable outcomes for most cases[13].

According to our results (Figure 8), the parameters influencing the minimum consumption level of households have both the largest first-order and total effects for most of the model's output variables. Their role is especially strong for the distribution of new lending across income deciles. The underlying reason behind this is the mechanism that if households in low income deciles consume less, they will be able to accumulate the down payment for buying a flat more easily, and they can also tolerate higher installments after their loans. Similarly, the same reasons can be responsible for the high impact on the transaction numbers in almost all regions in the country. These parameters even have a relatively large effect on the distribution of new loans along DSTI categories.

The parameters determining the moving behaviour of households mostly affect (at the first-order) the ratio of purchasing flats using loan and the distribution of new loans along LTV categories. A typical situation explaining this phenomenon can be the following. If households decide to move earlier, then they will be more likely to be in need for a (larger) bank loan, which would lead to higher LTV values. Regarding the total effect, we can observe the strongest impact of these parameters in the case of Northern Great Plain and Northern Hungary, and mostly on the average house prices and the number of transactions in these regions. Additionally, the total effect is also relatively strong for this category on the number of transactions of newly built flats.

[13] The variables for which the reliability of the results seems to be unsatisfactory are the following: house prices and number of transactions in the region of Northern Hungary, house prices in the region of Southern Transdanubia, and the credit market variables for the lowest and highest LTV categories.

The parameters influencing the foreclosure in case of defaulting loans have the strongest first-order effect on the house prices and on the number of transactions in Budapest. Furthermore, the total effect is quite high in the case of the number of transactions of newly built flats, and on the average house prices and the number of transactions in the Northern and the Southern Great Plain regions. The reason why these variables are affected the most strongly by this parameter group is that in the case of foreclosure events flats are sold faster and with a higher discount, which leads to greater supply of flats and lower prices on the market.

The first-order impacts of the parameters governing households' purchasing behaviour for investment purposes is very high on the house price variables in almost every region, but they are also important in the case of the loan distribution along DSTI and LTV categories. The total effect is outstandingly high here on the number of transactions of newly built flats. As these parameters largely determine the investment demand, their strong influence on the prices and on lending is intuitively understandable.

Finally, it is important to highlight the differences between the first-order and total effects. Where they deviate from each other to a great extent, it implies a strong role of interactions. This is the case in particular for the number of transactions of newly built flats, and also for the average house prices and for the number of transactions in certain regions. On the other hand, this difference is relatively small in the case of the loan distributions across DSTI and LTV categories, and it is insignificant for the amount of newly issued loans and the ratio of flat purchases with loans.

5 Conclusion

In this paper we introduced a complex, modular, 1:1 scale model of the Hungarian residential housing market, which is based on several empirical, micro-level datasets. This modelling framework enables policy makers to analyze detailed economic scenarios with macroprudential, fiscal and monetary policies in a way which surpasses the analytic capabilities offered by traditional tools – especially regarding the granularity and the precision of the results. This became possible by tackling at the same time several of the fundamental challenges present in the literature, such as the cyclical behavior, the interlacements with the financial sector, the geographic fragmentation and the large extent of heterogeneities among the agents on this market. The most important features and improvements compared to the existing practices in these regards are the following.

To make it possible to analyze the effects of various economic policies at disaggregated levels (e.g. along geographic regions, income deciles of households, etc.) we divided the country into 124 relatively homogeneous neighbourhoods. To ensure the reliability of the model, we had to have sufficient amount of data at this level of granularity, which was only possible by simulating the housing market with all the households and the housing stock in its entirety. To justify this modelling strategy, we showed that the model is not scale-invariant, i.e. the long term averages of the generated main time series change considerably depending on the number of agents. The main driver of the scale dependence is that buyers would place a bid with a high probability only on flats with which they can obtain a consumer surplus close to that of their ideal dwelling. Fewer agents in the economy would lead to lower transaction numbers, in which case the effects of policy shocks can be masked by the amplified frictions.

The second significant contribution considers the characteristics of flats. Most of the papers consider only a single "quality" measure to quantify the desirability of flats. However, our choice of characterizing flats by three attributes (*size, state* and *location*) have several benefits: (i) while the *size* is constant over time (as it does not change even due to renovation), the *state* needs to be time-variant to make it possible to model decisions about renovations. (ii) According to the results of our hedonic price regressions all of these attributes are highly significant, and they increase the goodness-of-fit considerably. (iii) The regression results also suggest the outstanding importance of our neighbourhood quality variable (especially compared to more coarse-grained administrative location variables), which contains every relevant information (public transportation, health and education institutions, residents' income situation, etc) about the location of the flats.

To account also for the changes in the housing stock, it is pivotal to represent the construction sector in a detailed way, which criterion is not met sufficiently in the existing housing market literature. In this paper we emphasized the importance of this modelling block using a threefold argument. Firstly, the building up of bubbles on the housing market can be substantially influenced by the construction sector given its potential to offset the effect of excessive demand, but also by its sensitivity to increased labor and material cost shocks. Secondly, we showed that the construction sector has important role on the housing market also in the case of long-term demographic phenomena which can have heterogeneous implications across regions resulting in different dynamics on the local housing markets. Lastly, the construction sector often has central role in the outcome of various policies on the housing market. To capture the implications in a plausible way, the detailed representation of the construction sector has become inevitable in the model.

To further increase the reliability of the model, we attempted to represent household's decision making process in an innovative way. When choosing a flat, households consider the consumer surplus according to standard utility maximization theory. Since traditional functional forms in consumer theory are used in a different spirit, we did not constrain ourselves to a CES-design, instead, we investigated alternative forms and eventually opted for an exponential function with multiplicative size and state variables. In order to take into account the heterogeneity of agents' preferences, each household in the model has been assigned a reservation price function which was calibrated uniquely using a stochastic optimization procedure.

Regarding the financial connections of the housing market, models usually only include standard housing loans typically with long maturity. In contrast, we propose two further loan contract types which are available for the households in the model: (i) personal loans which can be used for renovation purposes and (ii) bridge loans for households wanting to buy a new flat and sell the old one simultaneously. The inclusion of these additional loan types has an important contribution to the model's suitability to analyse several aspects of the macroprudential regulation. Firstly, households have to comply with the real-life LTV and DSTI rules in order to take out a loan, which enhances the plausibility of the model's assumptions. Secondly, with this detailed representation of the banking block of the model we could examine not only the volume of the new loans, but also their distribution along the income deciles of the households, which analysis showed that the tightening of the credit supply affects more heavily the households with lower income level.

In our model we also ensured the possibility of cycles' emergence in the housing market by two mechanisms: (i) credit constraints and (ii) the impatience of the buyers. In the case of a tight housing market, some households might have to wait longer to buy a suitable flat. The more they have to wait, the more tolerant they become regarding the price and the characteristics of flats. This behavior exerts upwards pressure on the prices, however, after a while this effect will be overpowered by the more and more often binding credit constraints. We compared and combined our strategy for generating cyclical behavior with another mechanism, which is often applied in similar models. It is built on the assumption that households are willing to spend more (less) on housing if the prices increase (decrease) on the market, which strengthens the trends in the price dynamics.

The model has been used for several further applications at the Central Bank of Hungary, e.g. assessing the effects of monetary policy shocks, simulating fiscal subsidies for households to buy flats, analysing macroprudential policy changes, etc. However, there are still ambitious plans for further development, most importantly to embed this agent-based housing market model into a macroeconomic environment to be able to generate endogenous cycles and feedback mechanisms with the rest of the economy. While this step would undoubtedly be a great challenge, the experiences regarding the versatile applicability of this model suggest that it is worth further extending the frontiers of computational economics.

References

Acharya, V. V., & Richardson, M. (2009). Causes of the financial crisis. *Critical review*, *21*(2-3), 195–210.

Aikman, D., Haldane, A. G., & Nelson, B. D. (2015). Curbing the credit cycle. *The Economic Journal*, *125*(585), 1072–1109.

Axtell, R., Farmer, D., Geanakoplos, J., Howitt, P., Carrella, E., Conlee, B., … others (2014). An agent-based model of the housing market bubble in metropolitan Washington, DC. In *Deutsche Bundesbank's Spring Conference on "Housing markets and the macroeconomy: Challenges for monetary policy and financial stability* (pp. 5–6).

Baptista, R., Farmer, J. D., Hinterschweiger, M., Low, K., Tang, D., & Uluc, A. (2016). Macroprudential policy in an agent-based model of the UK housing market. *SSRB*.

Borgonovo, E., Pangallo, M., Rivkin, J., Rizzo, L., & Siggelkow, N. (2022). Sensitivity analysis of agent-based models: a new protocol. *Computational and Mathematical Organization Theory*, *28*(1), 52–94.

Borio, C. (2014). The financial cycle and macroeconomics: What have we learnt? *Journal of Banking & Finance*, *45*, 182–198.

Boros, L., Dudás, G., Kovalcsik, T., Papp, S., & Vida, G. (2018). Airbnb in Budapest: analysing spatial patterns and room of hotels and peer-to-peer accomodations. *GeoJournal of Tourism and Geosites*, *10*(1), 26–38.

Cardaci, A. (2018). Inequality, household debt and financial instability: An agent-based perspective. *Journal of Economic Behavior & Organization*, *149*, 434–458.

Carstensen, C. L. (2015). An agent-based model of the housing market. *Steps toward a computational tool for policy analysis. University of Copenhagen, MSc-szakdolgozat*.

Ceron, J. A., & Suarez, J. (2006). *Hot and cold housing markets: International evidence*. Centre for Economic Policy Research.

Cokayne, G. (2019). *The effects of macroprudential policies on house price cycles in an agent-based model of the Danish housing market* (Tech. Rep. No. 138). Danmarks Nationalbank Working Papers.

Cunningham, R., & Kolet, I. (2011). Housing market cycles and duration dependence in the united states and canada. *Applied Economics*, *43*(5), 569–586.

Deissenberg, C., Van Der Hoog, S., & Dawid, H. (2008). Eurace: A massively parallel agent-based model of the european economy. *Applied Mathematics and Computation*, *204*(2), 541–552.

De Stefani, A. (2020). House price history, biased expectations, and credit cycles: The role of housing investors. *Real Estate Economics*.

Devisch, O. T., Timmermans, H. J., Arentze, T. A., & Borger, A. W. (2009). An agent-based model of residential choice dynamics in nonstationary housing markets. *Environment and Planning A*, *41*(8), 1997–2013.

Dias, G. F. (2017). The time-varying garch-in-mean model. *Economics Letters*, *157*, 129–132.

Drehmann, M., Borio, C. E., & Tsatsaronis, K. (2012). Characterising the financial cycle: don't lose sight of the medium term! *BIS Working Papers* (380).

Dudás, G., Kovalcsik, T., Boros, L., Kovalcsik, B., & Vida, G. (2018). Az airbnb szálláskínálatának jellemzői a megyei jogú városokban. *Területi Statisztika*, *58*(5), 462–488.

Erlingsson, E. J., Teglio, A., Cincotti, S., Stefansson, H., Sturlusson, J. T., & Raberto, M. (2014). Housing market bubbles and business cycles in an agent-based credit economy. *Economics: The Open-Access, Open-Assessment E-Journal*, *8*(2014-8), 1–42.

Evans, B. P., Glavatskiy, K., Harré, M. S., & Prokopenko, M. (2021). The impact of social influence in Australian real estate: market forecasting with a spatial agent-based model. *Journal of Economic Interaction and Coordination*, 1–53.

Fatouh, M., Markose, S., & Giansante, S. (2019). The impact of quantitative easing on UK bank lending: Why banks do not lend to businesses? *Journal of Economic Behavior & Organization*.

Funke, M., & Paetz, M. (2012). A DSGE-based assessment of nonlinear loan-to-value policies: Evidence from Hong Kong. *SSRN*.

Garriga, C., & Hedlund, A. (2020). Mortgage debt, consumption, and illiquid housing markets in the great recession. *American Economic Review*, *110*(6), 1603–34.

Ge, J. (2013). Who creates housing bubbles? An agent-based study. In *International workshop on multi-agent systems and agent-based simulation* (pp. 143–150).

Ge, J. (2017). Endogenous rise and collapse of housing price: An agent-based model of the housing market. *Computers, Environment and Urban Systems*, *62*, 182–198.

Gelain, P., & Bank, N. (2011). Macro-prudential policies in a DSGE model with financial frictions. In *7th Dynare Conference, Atlanta, USA.*

Gelain, P., Lansing, K. J., & Mendicino, C. (2012). House prices, credit growth, and excess volatility: Implications for monetary and macroprudential policy. *Forthcoming, International Journal of Central Banking, Norges Bank Working Paper*, *8*.

Gilbert, N., Hawksworth, J. C., & Swinney, P. A. (2009). An agent-based model of the English housing market. In *AAAI spring symposium: technosocial predictive analytics* (pp. 30–35).

Glaeser, E. L., & Sinai, T. (2019). *Housing and the financial crisis*. University of Chicago Press.

Glavatskiy, K. S., Prokopenko, M., Carro, A., Ormerod, P., & Harre, M. (2021). Explaining herding and volatility in the cyclical price dynamics of urban housing markets using a large-scale agent-based model. *SN Business & Economics*, *1*(6), 1–21.

Goodman, A. C., & Thibodeau, T. G. (1998). Housing market segmentation. *Journal of housing economics*, *7*(2), 121–143.

Hiebert, P., Jaccard, I., & Schüler, Y. (2018). Contrasting financial and business cycles: Stylized facts and candidate explanations. *Journal of Financial Stability*, *38*, 72–80.

Iacoviello, M. (2002). House prices and business cycles in europe: A VAR analysis. *Working Papers in Economics*, 81.

Jancsik, A., Michalkó, G., & Csernyik, M. (2018). Megosztás megosztottság nélkül–az airbnb és a budapesti szálláshelypiac átalakulása. *Közgazdasági szemle*, *65*(3), 259–286.

Jansen, M. J. (1999). Analysis of variance designs for model output. *Computer Physics Communications*, *117*(1-2), 35–43.

Kannan, P., Rabanal, P., & Scott, A. M. (2012). Monetary and macroprudential policy rules in a model with house price booms. *The BE Journal of Macroeconomics*, *12*(1).

Kaszowska-Mojsa, J., & Pipień, M. (2020). Macroprudential policy in a heterogeneous environment – An application of agent-based approach in systemic risk modelling. *Entropy*, *22*(2), 129.

Kelly, R., McCann, F., & O'Toole, C. (2018). Credit conditions, macroprudential policy and house prices. *Journal of Housing Economics*, *41*, 153–167.

Kiyotaki, N., Michaelides, A., & Nikolov, K. (2011). Winners and losers in housing markets. *Journal of Money, Credit and Banking*, *43*(2-3), 255–296.

Kouwenberg, R., & Zwinkels, R. C. (2015). Endogenous price bubbles in a multi-agent system of the housing market. *PloS one*, *10*(6), e0129070.

Kuang, P. (2014). A model of housing and credit cycles with imperfect market knowledge. *European Economic Review*, *70*, 419–437.

Laliotis, D., Buesa, A., Leber, M., & Población, J. (2020). An agent-based model for the assessment of LTV caps. *Quantitative Finance*, *20*(10), 1721–1748.

Lambertini, L., Mendicino, C., & Punzi, M. T. (2013). Leaning against boom–bust cycles in credit and housing prices. *Journal of Economic Dynamics and Control*, *37*(8), 1500–1522.

Lauretta, E. (2018). The hidden soul of financial innovation: An agent-based modelling of home mortgage securitization and the finance-growth nexus. *Economic Modelling*, *68*, 51–73.

Levine, A. S., Frank, R. H., & Dijk, O. (2010). Expenditure cascades. *Available at SSRN 1690612*.

Lin, Z., & Vandell, K. D. (2007). Illiquidity and pricing biases in the real estate market. *Real Estate Economics*, *35*(3), 291–330.

Magliocca, N., Safirova, E., McConnell, V., & Walls, M. (2011). An economic agent-based model of coupled housing and land markets (chalms). *Computers, Environment and Urban Systems*, *35*(3), 183–191.

Mell, H., Baumard, N., & André, J.-B. (2021). Time is money. Waiting costs explain why selection favors steeper time discounting in deprived environments. *Evolution and Human Behavior*, *42*(4), 379–387.

Mendicino, C., & Punzi, M. T. (2014). House prices, capital inflows and macroprudential policy. *Journal of Banking & Finance*, *49*, 337–355.

Muth, R. F. (2017). *The spatial structure of the housing market*. Routledge.

Nobuhiro, K., & Moore, J. (1997). Credit cycles. *Journal of Political Economy*, *105*(2), 211–248.

Ortalo-Magne, F., & Rady, S. (2006). Housing market dynamics: On the contribution of income shocks and credit constraints. *The Review of Economic Studies*, *73*(2), 459–485.

Ozel, B., Nathanael, R. C., Raberto, M., Teglio, A., & Cincotti, S. (2019). Macroeconomic implications of mortgage loan requirements: an agent-based approach. *Journal of Economic Interaction and Coordination*, *14*(1), 7–46.

Pangallo, M., Nadal, J.-P., & Vignes, A. (2019). Residential income segregation: A behavioral model of the housing market. *Journal of Economic Behavior & Organization*, *159*, 15–35.

Rubio, M., & Carrasco-Gallego, J. A. (2014). Macroprudential and monetary policies: Implications for financial stability and welfare. *Journal of Banking & Finance*, *49*, 326–336.

Rubio, M., & Comunale, M. (2017). Lithuania in the euro area: monetary transmission and macroprudential policies. *Eastern European Economics*, *55*(1), 29–49.

Saunders, T., & Tulip, P. (2020). A model of the Australian housing market. *Economic Record*, *96*, 1–25.

Appendix A Details of the model

A.1 EXOGENOUS MACROECONOMIC ENVIRONMENT

The macroeconomic environment is exogenous and can be described by the following time series:

- $GDP_t = GDP_t^b \times GDP_t^{sh}$. GDP_t^b is the base real GDP index and GDP_t^{sh} is a time dependent parameter altering the GDP according to the shock scenario (in normal times it is set to 1). In normal times individuals form expectations for the GDP using GDP_t^b, but under a shock even for the expectations they consider $GDP_t^b \times GDP_t^{sh}$
- P_t - price index
- IR_t^b - base interest rate
- r_t^{TAX} - income tax rate
- The probability of becoming unemployed ($U_{t,c}$) is differentiated based on three possible educational levels of individuals (c). The length of unemployment is defined with two variables (which are also differentiated along the same educational categories): in each category individuals stay unemployed for at least p_c^{um} periods, after which they get hired with the probability of $U_{t,c}^f$.

A.2 FLATS

A.2.1 PRICING OF FLATS

In this section we only discuss flat prices without going into details about other characteristics of flats. Details on these can be found in Appendix B.1.

We distinguish between the *market price* ($PR_{t,f}^m$), *ask price* ($PR_{t,f}^a$), *transaction price* ($PR_{t,f}^t$) and – in the case of newly built flats – the *cost-based price* ($PR_{t,f}^c$) of a flat f in period t.

In the case of used flats, the *market price* of flat f is a preliminary estimation of the equilibrium price based on former transactions. This will also be the first ask price of a flat after it is taken to the market. It is calculated from the transactions of the previous period provided that there were at least two transactions ($j \in TR_{t-1}$, where TR_{t-1} is the set of transaction records of period $t-1$) in the same neighbourhood in the previous period, for which the absolute size deviation ratio ($sdr_{t,f,j} = |size_j - size_f|/size_f$) and the absolute state deviation ratio ($qdr_{t,f,j} = |state_{t,j} - state_{t,f}|/state_{t,f}$) do not exceed the parameters th^{cns} and th^{cnq} respectively. To get the two closest records, we use the following distance measure: $DIST_{t,f,j} = w^{ds} sdr_{t,f,j} + w^{dq} qdr_{t,f,j}$, where w^{ds} and w^{dq} are exogenous parameters. If we cannot take the two closest neighbours in a given period, we adjust the previous general market price by the increase in the nominal GDP ($PR_{t,i}^m = PR_{t-1,i}^m \times P_t GDP_t/(P_{t-1} GDP_{t-1})$). We calculate the market price of each flat in each period, but it does not necessarily equal the ask price. For a used flat f, $PR_{t,f}^a = PR_{t,i}^m \times (1 - c^{mpd})^{p_f^{on}}$, where p_f^{on} is the number of periods in the market, and c^{mpd} is the monthly price decrease.

Regarding the demand side, agents place bids on flats. If at the time of purchase, there is only one valid bid, the transaction price equals the ask price, otherwise it equals the second highest bid price[14], i.e. we follow the logic of Vickrey auctions.

In case of the forced liquidation of flats serving as collateral we use a somewhat altered formula. The ask price is similar to the price of a used flat, but there is an additional discount term (DC), which can even be further augmented in the case of a negative macroeconomic shock event, i.e. :

[14] All bids must be greater than or equal to the ask price.

$$PR^{a}_{t,f} = PR^{m}_{t,i} \times (1 - c^{mpd})^{p^{on}_{f}} \times \begin{cases} DC, \text{ if } GDP^{sh}_{t} > thr^{dc} \\ DC \times DC^{a} \text{ otherwise} \end{cases} \tag{A.1}$$

where DC^a represents the augmentation term to the discount factor when the shocked GDP is less than thr^{dc} fraction of the base GDP.

In case of newly built flats, the ask price depends on the land and construction costs and the mark-up:

$$PR^{c}_{t,f} = size_f \times (LP_{t,N_f} + CUC_{t,R_f}) \times (1 + \mu^{C}_{t,R_f}) \tag{A.2}$$

where LP_{t,N_f} is the land price in the neighbourhood of a given flat f in period t. To calculate LP_{t,N_f}, we take the transactions of period $t-1$ in a given neighbourhood and calculate the unit sale price of flats[15].

CUC_{t,R_f} is the construction unit cost in the region of the flat. For region i, it is calculated by taking a base cost (CUC^b) which can be altered by the regional average wage ($\overline{W}_{t,i}$), other unobserved region specific factors (θ^{CUC}_{i}) and the nominal GDP index (P_tGDP_t):

$$CUC_{t,i} = CUC^{b}_{t,i} + c^{CUC}\overline{W}_{t,i} \tag{A.3}$$

$$CUC^{b}_{t,i} = P_tGDP_t \times CUC^{b} \times (1 + \theta^{CUC}_{i}) \tag{A.4}$$

where c^{CUC} is an exogenously given parameter.

The construction mark-up ($\mu^{C}_{t,i}$) depends on the tightness of the market which we capture by calculating as the ratio of newly built flats sold under construction compared to all newly built flats sold in a given region ($SR^{NB}_{t,i}$):

$$\mu^{C}_{t,i} = \mu_{C,0} + \mu_{C,1} \times SR^{NB}_{t,i} \tag{A.5}$$

where the constant $\mu_{C,0}$ and the coefficient $\mu_{C,1}$ are exogenously given parameters.

Newly built flats can be sold even under construction using the cost-based price (no matter how many periods are left until completion). After p^{nba} periods after completion, the construction sector starts adjusting the ask price to converge to the market price of the given flat according to the following formula:

$$PR^{a}_{t,f} = PR^{c}_{t,f} - (PR^{c}_{t,f} - PR^{m}_{t,f}) \times (1 - (1-\gamma)^{p^{on}_{f} - p^{nba}}), \tag{A.6}$$

where p^{on}_{f} is the number of periods passed since completion and γ is the monthly price convergence rate.

A.2.2 FICTIVE FLATS

There are three occasions in the simulation where we use fictive flats to support the decision making process of agents: (i) One is to determine the optimal level of renovation (see A.4.2 and A.2.3). (ii) The second is to determine an ideal choice in

[15] At this point we filter for outliers regarding the size of flats and we only consider flats with relatively low state value (i.e. below $state^{max}_{LP}$).

the housing market which can serve as a reference for households and help the construction sector estimating demand for different types of newly built flats. Since flats for sale in the market can vary period by period (or even within a single period), this ideal choice represents the flat which is worth waiting for at current market prices for a given household. (iii) Finally, we use fictive flats also to determine the ideal choice in the rental market to help investors estimate rental demand. Intuitively we can interpret this – and also the situation of the construction sector in (ii) – as conducting a market survey supporting investment decisions.

To determine the ideal choice in the housing market, in each period, we generate the set of fictive flats F^{fh}, which contains n^f flats in all the buckets in every neighbourhood, randomly selecting a size and a state within the intervals of the bucket. This list includes newly built flats as well. For all these flats, we determine an ask price, which equals the market price in case of non newly built flats and which equals the cost-based price in case of newly built flats. In each period t every household h selects its ideal flat $f^*_{t,h}$ out of the fictive sample based on the consumer surplus (see Appendix A.4.2), considering its preferred regions (see Appendix A.4.3).

A.2.3 RENOVATION

The *state* of flats deteriorates every month at a predefined rate δ, which can be compensated by the households through renovating their apartment. Each period, households are selected with probability pr^r to renovate their homes. In this case, they calculate the optimal level of renovation using the strategy described in Appendix A.4.2. Flats in the rental market are automatically renovated in every period to offset the deterioration of flats.

The cost of renovation of flat f to state $state_r$ is denoted by RC_{t,f^r} and calculated the following way:

$$RC_{t,f^r} = (state_r - state_{t,f}) \times size_{t,f} \times RUC_{t,R_f} \tag{A.7}$$

where RUC_{t,R_f} is the renovation unit cost in the region of flat f. In region i

$$RUC_{t,i} = RUC^b_{t,i} + c^{RUC}_1 \overline{W}_{t,i} + c^{RUC}_2 \overline{W}_{t,i} \times \omega_{t,i} \tag{A.8}$$

$$RUC^b_{t,i} = P_t GDP_t \times RUC^b \times (1 + \theta^{RUC}_i), \tag{A.9}$$

where RUC^b is the base renovation unit cost which can be altered by regional mark-ups (θ^{RUC}_i) and is always adjusted by the nominal GDP index. c^{RUC}_1 and c^{RUC}_2 are exogenous parameters. $RUC^b_{t,i}$ can be altered as a function of the regional average wage ($\overline{W}_{t,i}$). We also take into account the extent to which the renovation volume in period t deviates from the long term average within the region ($\omega_{t,i}$). This captures the sudden imbalances between supply and demand factors.

A.3 INDIVIDUALS

Individuals may be referred to as adults or children. Children live with their adult parents until they form an own household. Individual i is characterized by the following attributes:

- $a_{t,i}$: Age of individual i in period t
- e_i: Educational level of individual i
- w^0_i: Real starting wage of individual i
- $we_{i,t}$: Work experience of individual i in period t
- sex_i: Sex of individual i

Each individual may live up to a_{max} periods, but in each period it may die with probability pr^d which is dependent on sex_i, e_i and $a_{t,i}$. Each individual starts working at age $a^w_{e_i}$ and retires at age $a^p_{e_i}$. Individuals may get wage income or pension income.

In addition to these, households can get rent income and transfers. The net potential income of an adult individual i is

$$W^p_{t,i} = \begin{cases} P_t GDP_t \times (1 - r^{TAX}_t) \times w^0_i \times WR_{e_i, we_{i,t}}, & \text{if ind. } i \text{ is active} \\ prr_{e_i} \times P_t GDP_t \times (1 - r^{TAX}_t) \times w^0_i \times WR_{e_i, we_{i,t}}, & \text{if ind. } i \text{ is retired,} \end{cases} \tag{A.10}$$

where $WR_{e_i, we_{i,t}}$ is the wage ratio of a given educational level and work experience. WR is exogenously given and based on empirical data. Its elements can be interpreted as multipliers. The real starting wage and the multipliers of WR define the wage dynamics of an individual. This is then adjusted with the nominal GDP index. Pension income is a prr_{e_i} fraction of the last real wage earned (pension replacement rate), again, adjusted with the nominal GDP index. The net income of an individual ($W_{t,i}$) equals $W^p_{t,i}$ if an individual is employed or retired, and 0 in case of unemployment.

We also introduce a lifetime income of individuals ($LTI_{t,i}$) which is taken into account in the decision making process of house purchases. It is updated every month and it consists of two parts: earned (already accumulated) wage income ($LTI^a_{t,i}$) and expected future income ($LTI^e_{t,i}$).

$$LTI_{t,i} = LTI^a_{t,i} + LTI^e_{t,i}, \tag{A.11}$$

where

$$LTI^a_{t,i} = \begin{cases} LTI^a_{t-1,i} + GDP_t \times (1 - r^{TAX}_t) \times w^0_i \times WR_{e_i, we_{i,t}}, & \text{for employed individuals} \\ LTI^a_{t-1,i}, & \text{otherwise} \end{cases} \tag{A.12}$$

$$LTI^e_{t,i} = \sum_{s=t+1}^{t+\min\{p^{LTI}, p^{ur}_{t,i}\}} GDP^b_s \times SA_{s,i} \times (1 - r^{TAX}_s) \times w^0_i \times WR_{e_i, we_{i,s}} \tag{A.13}$$

$$SA_{s,i} = \begin{cases} 1, & \text{if } GDP^{sh}_t = 1 \\ GDP^{sh}_s \times (1 - u_{s,e_i}), & \text{if } GDP^{sh}_t < 1 \end{cases} \tag{A.14}$$

When calculating expected future income, the individual considers p^{LTI} periods, or the number of periods left until retirement age ($p^{ur}_{t,i}$), whichever number is lower. In normal times, the individual do not consider the possibility of being unemployed, as opposed to crisis periods, in which they adjust downward their expected income, governed by $SA_{s,i}$. According to these formulas, the lifetime income keeps growing, and hence, households will be willing to pay more for housing. The available lifetime income reaches its maximum p^{LTI} periods before retirement, at which point they can access all their lifetime income using the credit market.

A.3.1 DEMOGRAPHIC EVENTS

Marriage

An unmarried woman i tries to marry an unmarried man with probability $pr^m_{e_i, a_{t,i}}$, which indicates that it depends on the age and the educational level of the woman. Before making couples we group the unmarried men into groups according to region, age, educational level and starting wage. Then, for each marrying woman we randomly select a group of the same region and then we assign a husband from this group. The probability of selecting a specific group has been determined by looking at the characteristics of couples identified in the Central Credit Registry. For the sake of simplicity, the parameters of the newly formed household will be the same as those of the wife's previous household.

Birthgiving

In each period, every married woman gives birth to a child with probability $pr^b_{e_i, a_{t,i}, n^c_{t,i}}$, where $n^c_{t,i}$ is the number of previous births of woman i. These probabilities were determined using the identified childbearings in the Pension Payment database. The child inherits the educational level of the more highly qualified parent as well as the higher starting wage of the two parents.

Death

In each period, individuals may die with probability $pr^d_{a_{t,i},sex_i,e_i}$. If they die and they were the last adult of a household, they bequeath their properties and deposits to an inheritor. If they had children, the oldest child gets the inheritance, otherwise we randomly select a household to inherit.

A.4 HOUSEHOLDS

Households may consist of one or two adults ($A_{i,t}$ denotes the set of adults in household i in period t) and children ($C_{i,t}$ denotes the set of children in household i in period t).

An individual i forms a new household when marrying or when reaching age $a^w_{e_i}$.[16] When a child leaves its parents' household and forms a new one, it inherits the parameters of the parents' household (e.g. for the utility function). If the parents own a home, then the child inherits a fraction r^{di} (deposit inheritance ratio) of the parent household's deposits. Additionally, individuals may also inherit a deposit which is worth the market value of the parents' home. The probability of such an inheritance increases as the average starting wage of the parents increases. These rules are set in a way to compensate for the missing capital income in the model and to match the empirical down payment data of first-time buyers.

A.4.1 CONSUMPTION

The level of consumption depends on several factors (described below), but most importantly on the wage and size of the household; furthermore it might be constrained by the available deposits. Hence, it is important to establish the order of cash flows for the households to determine their consumption. Firstly, we account for the transactions in the housing market. Secondly, households receive their wage (or pension) and they pay/get rents. Only after these they can make their decision on the level of consumption by taking into account their ideal saving rate as well as their monthly installments in the following way:

In each period t, a household i calculates a target savings rate ($tsr_{t,i}$) and a minimum consumption ($c^m_{t,i}$). It aims to save a fraction $tsr_{t,i}$ of its actual household income ($I^h_{t,i}$) upon paying monthly installments ($MI^h_{t,i}$) and the rent ($R^h_{t,i}$) and consumes the rest of $I^h_{t,i}$, but before making any savings it tries to achieve the minimum consumption. So the consumption ($C^h_{t,i}$) can be calculated as follows:

$$C^h_{t,i} = \begin{cases} I^h_{t,i} \times (1 - tsr_{t,i}) - MI^h_{t,i} - R^h_{t,i}, & \text{if } I^h_{t,i} > c^m_{t,i} + tsr_{t,i} I^h_{t,i} + MI^h_{t,i} + R^h_{t,i} \\ \text{min}(c^m_{t,i}, D^c_{t,i}) \end{cases} \tag{A.15}$$

where $I^h_{t,i}$ contains the sum of the net income of the adults ($W^h_{t,i} = \sum_j W_{t_j}, j \in A_{i,t}$) and also family benefits $FB^h_{i,t}$, which is set by the government depending on the number of children under 18 in the household. We define $D^c_{t,i}$ as the deposit at the time of the consumption decision, because the amount of deposits can change within one period in the model.

$MI^h_{t,i}$ and $R^h_{t,i}$ are calculated in a straightforward way:

$$MI^h_{t,i} = \sum_l MI_{t,l}, l \in L_{i,t} \tag{A.16}$$

$$R^h_{t,i} = \begin{cases} R^f_{t,j}, \text{ if household } i \text{ rents a flat } j \\ 0, \text{otherwise} \end{cases} \tag{A.17}$$

[16] When children leave the parents' household they do not necessarily buy or rent an own flat, but they start accumulating deposits. This represents the period of young adults' life when they still live with their parents, but they already have started to make savings. The lower their starting wage is, the longer this period lasts.

where $L_{i,t}$ is the set of active loan contracts of household i in period t and $MI_{t,l}$ is the monthly installment on loan l.

The target saving rate ($tsr_{t,i}$) depends on the actual net income of the household without family benefits:

$$tsr_{t,i} = c_0^{tsr} + c^{tsr} \times \log(W^h_{t,i} \times 12/P_t) \tag{A.18}$$

where c_0^{tsr} and c^{tsr} are parameters.

To calculate $c^m_{t,i}$, we apply the following strategy. Firstly, we define the subsistence level of a household ($c^s_{t,i}$) based on its size (which is independent of the household's income). Then we define a second consumption level ($c^i_{t,i}$) which is in contrast based on the income of the household. This is calculated as the r^{mc} fraction of the net potential income of the household without family benefits ($W^{ph}_{t,i}$). The household tries to achieve a minimum consumption of whichever of the two consumption levels is higher, but if it does not have enough deposits to do so, in practice it starts reducing the effective income-based consumption target $c^i_{t,i}$, approaching the subsistence level $c^s_{t,i}$ in p^{mc} periods. These rules can be specified according to the following formulas:

$$c^s_{t,i} = c^{spc}_{t,i} \times n^{ue}_{t,i} \tag{A.19}$$

$$c^i_{t,i} = r^{mc} W^{ph}_{t,i} \tag{A.20}$$

$$c^m_{t,i} = \begin{cases} \max\{c^s_{t,i}, c^i_{t,i}\}, & \text{if } D^c_{t,i} > \max\{c^s_{t,i}, c^i_{t,i}\} \\ c^i_{t,i} - \min\{p^{mc}_{t,i}, p^{mc}\}/p^{mc} \times (c^i_{t,i} - c^s_{t,i}), & \text{otherwise.} \end{cases} \tag{A.21}$$

where $p^{mc}_{t,i}$ is the number of periods since the household has to spend all its deposits trying to achieve the minimum consumption, $c^{spc}_{t,i}$ is the per capita minimum subsistence level of the household and $n^{ue}_{t,i}$ is the number of unit equivalents of the household. The first adult's equivalent is 1, and for any other member above age a^{aspc}, it is w^{aspc}, below that it is w^{cspc}.[17]

A.4.2 UTILITY FUNCTION

Each household in the model has a reservation price function which assigns a utility to flats expressed in monetary terms. We calibrated several different functional forms[18], and chose the one fitting best to the empirical observations. Our final choice is shown by the following equation:

$$RP_{t,f,i} = (FC^{rp}_{t,f,i} + NEIGH^{rp}_{f,i}) * LTI^h_{t,i} * P_t + NB^f_{t,i} \tag{A.23}$$

where $RP_{t,f,i}$ is the reservation price of household i in period t for flat f, $FC^{rp}_{t,f,i}$ encompasses flat characteristics (size and state), $NEIGH^{rp}_{f,i}$ captures neighbourhood quality, $LTI^h_{t,i}$ is the available real lifetime income ($LTI^h_{t,i} = \sum_j LTI_{t,j}, j \in A_{t,i}$) and $NB^f_{t,i}$ is the additional utility in case of newly built flats.

[17] Since there are many households in Hungary living under the legal subsistence level, we could not take simply the legal subsistence level as $c^{spc}_{t,i}$. Hence, besides the official (upper) per capita subsistence level we also introduced an effective lower subsistence level (c^{spcu} and c^{spcl}), and $c^{spc}_{t,i}$ can vary within this range, depending on the net potential per capita income of the household ($W^{ppc}_{t,i} = \sum_j W^p_{t,j}/(n(A_{i,t}) + n(C_{i,t})), j \in A_{i,t}$):

$$c^{spc}_{t,i} = P_t \times GDP_t \times \begin{cases} c^{spcl}, \text{if } W^{ppc}_{t,i} < th^{spcl} \\ c^{spcu}, \text{if } W^{ppc}_{t,i} > th^{spcu} \\ c^{spcl} + \frac{W^{ppc}_{t,i} - th^{spcl}}{th^{spcu} - th^{spcl}} \times (c^{spcu} - c^{spcl}), \text{otherwise} \end{cases}, \tag{A.22}$$

where th^{spcl} and th^{spcu} are threshold parameters regarding $W^{ppc}_{t,i}$.

[18] The detailed description of these calibration results can be found in 3.2.

As it can be seen in Equation (A.23), flat characteristics and neighbourhood quality are handled separately. The intuition behind this is that the utility gain of living in a given neighbourhood captures aspects such as commuting time, air quality, public security, noise and night pollution, etc. which are the same regardless of the size and state of the flat itself[19].

The expansions of the parts regarding flat and neighbourhood characteristics are the following:

$$FC^{rp}_{t,f,i} = c^{rps}_i \times size_f^{\alpha^{rps}_i} \times (1 + c^{rpq}_i \times state_{t,f}^{\alpha^{rpq}_i}) \tag{A.24}$$

$$NEIGH^{rp}_{f,i} = \frac{\sigma_{1,i}}{(1+\exp(-Q_{N_f}))^{1/\sigma_{2,i}}} \tag{A.25}$$

where $size_f$, $state_{t,f}$ and Q_{N_f} are respectively the size, state and neighbourhood quality of flat f, while c^{rps}, α^{rps}, c^{rpq}, α^{rpq}, σ_1, and σ_2 are parameters determined by the calibration process.

The interpretation of the sigmoid function form in the case of neighbourhood quality is that each household might have a preferred quality level from which it would be very costly to deviate downwards in terms of utility, but at the same time moving upwards would only mean less and less additional gain.

Moreover, the reservation price function also contains an additional utility gain in case of newly built flats. Two factors are contributing to this component: (i) in Hungary, households can get a transfer from the state in case of purchasing a newly built flat, and so this transfer $TR^f_{t,i}$ directly increases the reservation price. (ii) Secondly, there can be an inherent motivation of households to live in a newly built environment. We assume that this gain interacts with the size of the flat similarly to the base case, and is more pronounced in higher quality neighbourhoods:

$$NB^f_{t,i} = \begin{cases} (c^{nba_0} + c^{nba_1} \times Q_{N_f}) \times c^{rps}_i \times size_f^{\alpha^{rps}_i} \times LTI^h_{t,i} + TR^f_{t,i}, \text{if } f \text{ is newly built} \\ 0, \text{ otherwise} \end{cases} \tag{A.26}$$

Adjusting the reservation price considering renovation possibilities

When purchasing a flat, we let r^r fraction of the households to consider renovation, hence, we also define $RP^{f'}_{t,i}$ as the reservation price for a flat with optimal renovation. To determine the optimal level of renovation for flat f, we apply the following strategy. Firstly, we generate a set of fictive flats (F^r_f) which are identical to the flat in consideration except for their state attribute[20]. As a next step, we calculate for each of these fictive flats ($f^r \in F^r_f$) the consumer surplus ($SP^{f^r}_{t,i}$) considering the effective cost of renovation (RC^e_{t,f^r}):

$$SP_{t,f^r,i} = RP_{t,f^r,i} - PR^a_{t,f} - RC^e_{t,f^r} \tag{A.27}$$

where $PR^a_{t,f}$ is the ask price of the original flat. The effective renovation cost has two components: (i) market renovation cost (RC_{t,f^r}, see Appendix A.2.3) which is to be paid and (ii) a nonmaterial burden (including psychological costs and time demand) captured as a fraction ν of the market renovation cost.

Households choose the renovation level ensuring a state with the highest surplus (and the underlying fictive flat f^*). Finally, the adjusted reservation price for flat f is the difference between f^*'s reservation price ($RP_{t,f^*,i}$) and the corresponding renovation cost:

$$RP^a_{t,f,i} = RP^{f^*}_{t,f^*,i} - RC^e_{t,f^r} \tag{A.28}$$

[19] There can be arguments also for separating size and state, but the results of the calibration suggest some interaction between the two.

[20] The state of these fictive flats is $state_{t,f}$ plus an integer multiple of int^r (renovation interval), but state cannot exceed $state^{max}_{used}$.

A.4.3 COMMITMENT TO REGIONS

Every household has a preferred region to live in, and they only consider the flats of that region in their housing decisions. However, to take into account the most typical tendencies of Hungarian within-country migration we introduce an exception: first-time buyers have the option to move outside of the region of their origin. Depending on region dependent probabilities p^{fb}, some first-time buyers can evaluate the flats of a second preferred region as well: for households from Budapest the second preferred region is the agglomeration of the capital (Pest county), for all the other households it is Budapest. This affects all the housing decisions of households (choice of ideal flat, the decision on moving, placing bids and rental decisions).

A.4.4 DECISION ON MOVING

Household i in period t may be selected to consider moving with probability $pr^m_{t,i}$ based on the age of the oldest adult of the household ($a^o_{t,i}$):

$$pr^m_{t,i} = \begin{cases} pr^m_f, & \text{if } a^o_{t,i} = a^m_f \\ pr^m_0, & \text{if } a^o_{t,i} <= a^m_0 \text{ and } a^o_{t,i} \neq a^m_f \\ pr^m_0 \times MPD^{a^o_{t,i} - a^m_0} & \text{otherwise} \end{cases} \tag{A.29}$$

where pr^m_0 is the probability of moving under age a^m_0, and MPD is the moving probablity decay parameter. This formula (with the current parameter set) implies that most households would consider moving when reaching age a^m_f with probability pr^m_f, but there is a basic probability of moving as well, which is a decreasing function of age.

If household i is selected to consider moving, it first compares the consumer surplus of its ideal fictive flat ($f^*_{t,i}$) and its own home ($h_{t,i}$), calculating the surplus deviation rate $spdr^*_{t,i}$:

$$spdr^*_{t,i} = (SP_{t,f^*_{t,i}} - SP_{t,h_{t,i}})/SP_{t,h_{t,i}}, \tag{A.30}$$

If $spdr^*_{t,i}$ exceeds a threshold thr^m, then the household compares the market prices, the sizes, the neighbourhood qualities of the ideal fictive flat and its own home, calculating the price deviation rate $pdr^*_{t,i}$, the size deviation rate $sdr^*_{t,i}$ and the neighbourhood quality deviation rate $qdr^*_{t,i}$ in absolute terms:

$$pdr^*_{t,i} = |PR^l_{t,f^*_{t,i}} - PR^l_{t,h_{t,i}}|/PR^l_{t,h_{t,i}} \tag{A.31}$$

$$sdr^*_{t,i} = |size_{f^*_{t,i}} - size_{h_{t,i}}|/size_{h_{t,i}} \tag{A.32}$$

$$qdr^*_{t,i} = |Q_{N_{f^*{t,i}}} - Q_{N_{h_{t,i}}}|/Q_{N_{h_{t,i}}} \tag{A.33}$$

If any of these rates exceed thr^m, and the household can finance the ideal fictive flat, then it decides to move and takes its home to the market.

A.5 TRANSACTIONS IN THE HOUSING MARKET

There are three distinct types of purchases in the housing market: the purchase of the (i) construction sector, (ii) buy-to-let investors and (iii) home buyers. The sequence of the purchases in the simulation also follows this order.

A.5.1 PURCHASES OF THE CONSTRUCTION SECTOR

The construction sector determines the additional quantity of flats to build in each bucket i ($n^{NBP}_{t,i}$). This decision is based on the fictive choices of households in the housing market, i.e demand for newly built houses, and the stock of newly built flats of the construction sector ($n^{NB}_{t,i}$, including flats under construction). $n^{NBP}_{t,i}$ is set according to the following formula:

$$n_{t,i}^{NBP} = \overline{NB_{t,i}^{D}} \times (1 + TNBB) \times p^{C} - n_{t,i}^{NB} \tag{A.34}$$

where $\overline{NB_{t,i}^{D}}$ is the average demand for newly built flats in bucket i in the past p^{NBD} periods, $TNBB$ is the target buffer rate of newly built flats, and p^{C} is the duration of construction expressed in periods. $NB_{t,i}^{D}$ is the number of households planning to buy a flat – let them be first-time buyers or movers – whose ideal fictive flat is a newly built flat in bucket i. For example, let us assume that the average demand in bucket i for newly built flats was 100, and there are 1400 flats under construction and 200 flats ready. If the buffer rate is 10 per cent, and construction takes 18 periods, then the construction sector aims to have $100 \times 18 \times 1,1 = 1980$ flats in its stock, and so it plans to start the construction of $1980 - 1400 - 200 = 380$ flats. If demand stays constant, it will be able to satisfy the accumulated demand (due to the buffer even a slightly higher demand), and in the long run it will start constructing 100 flats in each period. To avoid huge spikes, $n_{t,i}^{NBP}$ may not exceed a fraction r_{max}^{NB} of $\overline{NB_{t,i}^{D}}$.

In order to start building a new house, the construction sector must have free land area in the proper neighbourhood, which is a fraction r^{LA} of the area of the flat to build. Whenever the construction sector purchases a used flat, the flat area is added to its free land area in the neighbourhood,[21] and whenever the construction of a new flat f begins in the neighbourhood, this land area decreases by $r^{LA} \times size_{f}$. In each period t for each bucket i, the construction sector determines $n_{t,i}^{NBP}$, and determines the sizes of the flats to build randomly within the size interval of the given bucket. Summing the sizes of these flats for each neighbourhood n (A_{n}^{d}), the construction sector needs to have a free land area of $r^{LA} \times A_{n}^{d}$. If its land area is less, then it starts purchasing flats in the neighbourhood until the land area exceeds the needed volume. The construction sector buys flats sequentially, and every time it purchases the flat of the cheapest unit price (independently of its size) in the given neighbourhood. In this case the transaction price equals the list price.

A.5.2 PURCHASES FOR INVESTMENT PURPOSES

Purchases for investment purposes can be made either by the professional investor or buy-to-let investor households. Since the professional investor is a representative agent, it makes large-scale investment decisions in several segments of the market at the same time. Its decision making mechanism partly differs from that of the disaggregated household investors, since while single buy-to-let investors make binary decisions, the representative professional investor needs to make a decision on the aggregate investment volume as well.

The basic logic in the case of the professional investor is the following: Firstly, it decides how much to invest in each neighbourhood n ($PIV_{t,n}^{e}$), and it purchases flats sequentially until it reaches the desired investment volume. When choosing the next flat to buy, the professional investor randomly selects a bucket i with probability $pr_{t,i}^{b}$, and purchases the flat with the cheapest unit cost in the bucket. In contrast, single households may invest by purchasing only one flat at a time. For each household, we assign a neighbourhood n with *selection probability* $pr_{t,n}^{s}$ in which it tries to invest with *investment probability* $pr_{t,n}^{i}$. If the household tries to invest in neighbourhood n, it selects a bucket the same way as the professional investor (with probability $pr_{t,i}^{b}$). If it can manage to finance the purchase of the flat with the cheapest unit cost in the bucket, it will buy it, otherwise it will not invest in the given period.

Both $PIV_{t,n}^{e}$ and $pr_{t,n}^{i}$ depend on the expected return spread:

$$ERS_{t,n} = ER_{t,n} - IR_{t}^{b} \tag{A.35}$$

where IR_{t}^{b} is the base interest rate and $ER_{t,n}$ is the expected return on investment which is calculated as the average return in renting out flats in neighbourhood n in the past 12 months:

$$ER_{t,n} = \sum_{s=t-12}^{t-1} BTL_{s,n}^{r}/12 \tag{A.36}$$

[21] Demolishing the purchased flat to increase the free land area does not have a cost in the model.

where $BTL^{r}_{t,n}$ is the monthly return which can be realized by renting out flats in neighbourhood n in period t. It can be calculated as:

$$BTL^{r}_{t,n} = \sum_{f} RI_{t,f} / \sum_{f} PR^{m}_{t,f}, f \in R_{t,n}, \tag{A.37}$$

$$RI_{t,f} = \begin{cases} 0, \text{if } f \text{ is vacant} \\ R^{m}_{t,f} \text{if } f \text{ is rented} \end{cases} \tag{A.38}$$

where $PR^{m}_{t,f}$ is the market price, $R_{t,n}$ is the set of flats in the rental market (rented or vacant) in neighbourhood n, $RI_{t,f}$ is the rent income on flat f, and $R^{m}_{t,f}$ is the market rental price of flat f (see A.7).

After calculating $ERS_{t,n}$ using (A.35)-(A.38), the planned (expected) investment value by the professional investor in neighbourhood n is:

$$PIV^{e}_{t,n} = (c^{PIR}_{0} + c^{PIR}_{1} ERS_{t,n}) \times \sum_{f} PR^{m}_{t-1,f}, f \in R_{t-1,n}, \tag{A.39}$$

where c^{PIR}_{0} and c^{PIR}_{1} are exogenous parameters and $R_{t-1,n}$ is the set of flats of the rental market (rented or vacant) in neighbourhood n in the previous period.

Consequently, as the expected return spread increases, the professional investor plans to invest a higher fraction of the market value of the flats at the rental market. However, we also apply an upper constraint, $PIV^{e}_{t,n}$ may not exceed $PIV^{e,max}_{t,n}$, which is a fraction r^{max}_{PIMV} of the total market value of the flats in the neighbourhood:

$$PIV^{e,max}_{t,n} = r^{max}_{PIMV} \times \sum_{f} PR^{m}_{t-1,f}, f \in F_{t-1,n} \tag{A.40}$$

Now we can also define the *investment probability* of households in neighbourhood n, which is an increasing function of $ERS_{t,n}$:

$$pr^{i}_{t,n} = \begin{cases} 0, \text{if } ERS_{t,n} < 0 \\ \max\{pr^{i}_{max}, c^{i} ERS^{\alpha_i}_{t,n}\}, \text{otherwise} \end{cases}, \tag{A.41}$$

where pr^{i}_{max} is the cap of household investment probability and c^{i} and α_i are exogenously given parameters.

The *selection probability* of neighbourhood i is proportionate to $n^{mn}_{t,i}$, which is the difference between the number of flats needed to satisfy the estimated (fictive) rental demand ($FDR^{n}_{t-1,i}$, see A.7) with a target vacancy rate of TVR and the actual number of flats on the rental market in the neighbourhood ($n(R_{t-1,i})$):

$$n^{mn}_{t,i} = FDR^{n}_{t-1,i}/(1 - TVR) - n(R_{t-1,i}) \tag{A.42}$$

Finally, we need to elaborate on $pr^{b}_{t,n}$ for neighbourhood n. We designed the investment mechanism such that the series of single investment decisions tend to result in a more or less *even vacancy rate* ($EVR_{t,n}$) across buckets in a given neighbourhood (B_n). Let $n^{m}_{t,i}$ denote the number of flats missing in the rental market in bucket i to achieve the vacancy rate $EVR_{t,n}$. We calculate $n^{m}_{t,i}$ for every bucket in the neighbourhood and $EVR_{t,n}$ endogenously, by solving the following equation:

$$PIV^{e}_{t} + HIV_{t-1,n} = \sum_{i} (FDR^{b}_{t,i}/(1 - EVR_{t,n}) - R^{b}_{t,i}) \times PR^{m}_{t,sf_i} i \in B_n, \tag{A.43}$$

where $HIV_{t-1,n}$ is the investment volume of household investors in period $t-1$ in neighbourhood n and PR^{m}_{t,sf_i} is the market price of a *sample flat in bucket i with the mean size and state of the bucket*. The left hand side captures the expected investment volume of the professional investor and households together, using adaptive expectations for household investment. The right hand side captures the financing need of purchasing the missing flats to achieve a vacancy rate $EVR_{t,n}$.

Since investment decisions are directly governed by the expected return spread, the supply side can be temporarily detached from the demand side, managing to generate disequilibrium outcomes in some segments of the rental market. This mechanism is further amplified by the potential overshooting on the supply side coming from the asynchronous decision making of autonomous agents. Additionally, the endogenous change in rental markups (see A.7) ensures the long term convergence to an equilibrium state.

A.5.3 PURCHASES OF OWNER OCCUPIERS

While the construction sector and buy-to-let investors purchase flats at the ask price, owner occupiers engage in a bidding process. Every bid must be at least as high as the ask price, hence, owner occupiers on average will pay higher prices than the construction and the investment sector. This enables the model to reflect the weaker bargaining power of owner occupiers.

Owner occupiers may purchase flats in multiple (n^r) rounds. In each round households who are still willing to buy a home place bids on flats. Then we go through the flats sequentially (in a random order), and the household with the highest bid purchases the flat. If there is only one bidder, the transaction price equals the ask price, otherwise it equals the second highest bid, which corresponds to the logic of a Vickrey auction. If a household purchases a flat, all of its other bids will be withdrawn, and the household does not enter the next bidding round.

As discussed in Appendix A.2.2, each household has an ideal flat $f^*_{t,h}$ in mind. It serves as a reference when actually deciding which flats to bid on. In each round of bidding, households willing to buy a flat go through the list of F_i available flats they might be interested in (see Appendix A.4.3), and they evaluate at most n^E flats. For each evaluated flat, they calculate an adjusted reservation price $RPA_{t,f,i}$ as a function of $p^{wb}_{t,i}$ which is the number of periods since they try but do not manage to buy a flat:

$$RPA_{t,f,i} = \begin{cases} RP^a_{t,f,i}, & \text{if } p^{wb}_{t,i} = 1 \\ RP^a_{t,f,i} + (SP_{t,f^*,i} - SP_{t,f,i}) \times \text{min}\{c^{rpa}(p^{wb}_{t,i})^{\alpha_{rpa}}, RPA^{max}\}, & \text{otherwise} \end{cases}, \tag{A.44}$$

where c^{rpa}, α_{rpa} and RPA^{max} are exogenously given parameters. That is, we use the reservation price (considering also optimal renovation, see Appendix A.4.2), which can be increased as the household spends more and more periods on the market unsuccessfully. The maximum adjustment is an RPA^{max} fraction of the difference in the consumer surpluses of the ideal fictive flat and the flat in consideration.

The adjusted reservation price $RPA_{t,f,i}$ implies an *adjusted consumer surplus* $SPA_{t,f,i} = SP_{t,f,i} + RPA_{t,f,i} - RP^a_{t,f,i}$. If $RPA_{t,f,i} > PR^a_{t,f}$, then the household places a bid on flat f with a probability of $p^b_{t,f,i}$:

$$p^b_{t,f,i} = \text{max}\{1, (c^{PPB}p^{wb}_{t,i} + 1) \times SPA_{t,f,i}/SP_{t,f^*,i}\}, \tag{A.45}$$

where c^{PPB} is an exogenously given parameter. Hence, the closer the adjusted consumer surplus is to the consumer surplus in the case of the ideal fictive flat, the higher the probability for a household to bid on a flat. This probability increases as the household spends more and more periods on the market unsuccessfully. Placing more bids increases the possibility to be the highest bidder for one of the bidded flats. However, a household can place at most n^B bids. If the number of bids would exceed n^B, only the flats with the highest consumer surpluses will be considered.

The bid is the weighted average of the adjusted reservation price of the household and the ask price of the flat. Moreover, the bid can be limited by the available credit on the flat:

$$bid^f_{t,i} = \text{min}\{w^{RPA}RPA_{t,f,i} + (1 - w^{RPA})PR^l_{t,f}, D_{t,i} + CREDIT^{max}_{t,f,i}\}, \tag{A.46}$$

where w^{RPA} is the weight of the adjusted reservation price, $D_{t,i}$ is the amount of deposit and $CREDIT^{max}_{t,f,i}$ is the maximum available credit for flat f for household i. (See Appendix A.6).

By assigning a relatively low value to parameter w^{RPA}, in most cases all the bids will exceed the ask price only by a slight margin, however, the bidding mechanism is still capable of effectively allocating the flats to the bidders who value them the most. The increase of $RPA_{t,f,i}$ (after spending more and more periods on the market) will increase both the probability of placing a bid and the bid itself. Both mechanisms enhance the chance of purchasing a flat, though at the same time the household might have to settle for a lower consumer surplus.

A.6 CREDIT MARKET

In the model there are three types of loans: regular housing loan, bridge loan and renovation loan. Housing loans and bridge loans are mortgage loans, and each flat can serve as a collateral for at most one loan contract at a time. Regular housing loans and bridge loans are issued at the time of purchase while renovation loans are issued at renovation. Housing loans and renovation loans are used for different purposes, from the bank's perspective they are treated in the same way: both are annuity loans and need to comply with the DSTI and LTV regulation.

Bridge loans can be issued for marketable flats which is defined by two criteria in the model: (i) the market price should exceed a threshold thr^{BL} (eliminating cheap rural houses) and (ii) the unit price of the flat should exceed the renovation unit cost in the region of the flat. The bridge loan can be at most a ratio r^{BL}_{max} of the market price of the flat. Each month, the principal of the bridge loan is increased by the interest, based on the base rate ($IR^b/12$). The bridge loan principal is to be repaid at the sale of the collateral.

Regarding regular housing loans and renovation loans, the monthly installment for loan l ($MI_{t,l}$) is calculated as an annuity, which is affected by the interest rate (IR_l) and the maturity (p^D_l) of the loan:

$$MI_{t,l} = LOAN_f \times IR_l/(1-(1+IR_l)^{-p^D_l}) \tag{A.47}$$

The maturity of a newly issued loan is automatically p^{ML} or p^{RL} periods for housing and renovation loans, respectively. The interest rate of a newly issued loan is the sum of the base rate and a spread. The spread is set according to a regression on ln W^h and LTV and three age categories. The coefficients were estimated according to empirical data (Appendix B.4). Finally the spread is adjusted according to interest rate fixation.

When issuing a new loan (or extending an existing one because of renovation) households need to meet the two regulatory requirements (LTV and DSTI) and the bank's own prescription regarding a minimum consumption level. Regarding the LTV regulation, the principal for flat f for household i at the time of issuance ($LOAN_f$) may not exceed a fraction LTV^{max} of the ask price of the flat.[22] The compliance with the DSTI regulation is calculated by taking into account all monthly installments:

$$MI^i_{t,l_f} + \sum_l MI_{t,l}, < DSTI_{i,l_f} I^h_{t,i}, l \in L_{t,i} \tag{A.48}$$

where MI^i_{t,l_f} is the monthly installment provided that household i gets the loan for flat f and $DSTI_{i,l_f}$ depends on the length of interest rate fixation and on $I^h_{t,i}$. The length of interest rate fixation is generated according to the empirical distribution of the loan contracts in Hungary.[23]

The household pays the monthly installments in the order of issuance. If a household cannot fully pay the monthly installment of loan l, it pays as much as it can and the loan becomes nonperforming. By $p^{np}_{t,l}$ we denote the number of periods since loan l is nonperforming. If $p^{np}_{t,l}$ exceeds a threshold p^{npr}, the bank tries to restructure the loan, by increasing the duration at most to p^{ML} periods, until the debtor would meet the DSTI requirements. If the household becomes able to fully pay the restructured installment, the loan contract becomes performing again. However, if this intervention does not work and $p^{np}_{t,l}$ exceeds p^{npl}, the bank no longer tries to restructure the loan and the collateral will be liquidated.

If a loan contract reaches the stage of liquidation, the outstanding principal is accounted as a loss. If a loan is nonperforming, and the actual payment does not cover the interest payment, the outstanding principal will be increased by this missing interest payment. Until the collateral can be liquidated, the accumulation of interest payment continues, as well as the increase of the losses.

If the collateral can be liquidated, we adjust the bank losses by the income of the bank coming from this transaction. In this case the debtor only gets a r^{fs} fraction of the transaction price of the flat, then it pays all its obligations to the bank connected to the given loan.

A.7 RENTAL MARKET

Renters pay the ask rental price in the rental market, so the ask rental price and the market rental price can be used interchangeably. For flat f the ask (or market) rental price ($R^m_{t,f}$) is a fraction r^{RP} of the market price of the flat adjusted by markup

[22] In the case of renovation loans, we use the market price since the ask price is not available.

[23] We extrapolated the observed trend of the share of long term fixation converging to 1.

$\mu^r_{t,n}$, used throughout neighbourhood n of the bucket of flat f:

$$R^m_{t,f} = r^{RP} \times PR^m_{t,f} \times (1 + \mu^r_{t,n}) \tag{A.49}$$

$$\mu^r_{t,n} = c^\mu_r (1 - VR_{t-1,i})^{\alpha^\mu_r} \tag{A.50}$$

where c^μ_r and α^μ_r are exogenous parameters, and $VR_{t,i}$ is the vacancy rate of the rental market in neighbourhood n in period $t-1$, so the markup is a monotonically decreasing function of the vacancy rate.

If a household does not have an own home nor a rental contract, but needs to live in a separate flat, to select a flat to rent, it looks through the vacant flats in the rental market in its preferred region and selects the one with the highest surplus. The calculation of the rental reservation price ($RRP_{t,f,i}$) happens using the household's utility function as in (A.23), but to get the surplus, instead of the rental ask price we use $R^m_{t,f}/r^{RP}$ in order to match the appropriate magnitude. Households come in random order, and rent out a flat for p^r periods.

Similarly to the fictive ideal own flat for households, we also define an ideal rental flat for those who are active in the rental market. We determine the fictive rental demand for each bucket i ($FDR^b_{t,i}$) based on this calculation, which is a contributing factor to investment decisions (see Appendix A.5.2).

The probability of taking a vacant flat to the housing market in period t in neighbourhood i is $pr^{RS}_{t,i}$ (rent sale probability). The calculation of this consists of two components:

$$pr^{RS}_{t,i} = pr^{ERS}_{t,i} \times pr^{VR}_{t,i}, \tag{A.51}$$

where $pr^{ERS}_{t,i}$ is a component which linearly depends on the expected return spread (with constant c^{RS}_0 and coefficient c^{RS}_1). $pr^{VR}_{t,i}$ is an adjustment factor based on the vacancy rate in the given neighbourhood:

$$pr^{VR}_{t,i} = \frac{1 - TVR}{1 - VR_{t,i}}, \tag{A.52}$$

where TVR is the target vacancy rate (see (A.42)).

A.8 SAMPLE MODEL - EMPIRICAL MODEL DIFFERENCES

The sample model has less features than the empirical model and we had to implement the generation of initial individuals, households and flats. In the sample model there is only one region (and so first-time buyers do not consider moving to another region). This region contains only four neighbourhoods which is enough to implement neighbourhood quality heterogeneity, but within a neighbourhood we have the same buckets regarding size and state. This implies that even a model with 20-30 thousand households can have enough transactions in a month to have a properly functioning model (though we have seen in 3.4 that increasing the number of households leads to more accurate results). We tried to keep input generation (individuals, households and flats) as simple as possible while minimizing the effect of demographic changes and trying to make the housing stock more or less match the preferences of the initial population. Regarding individuals, the age distribution is even, every woman gets married and they marry at the same age and give birth to two children at the same age, and every individual passes away at the same age. We keep educational heterogeneity, but regarding income, the starting real wage distribution is normal.

In the beginning we match a husband to every woman above marriage age with a proper number of children and then we calculate the lifetime income of households. We assign utility function parameters according to trimmed normal distributions regarding the six parameters. Given the lifetime income and the utility function parameters, we can assign a flat to every household. To do so we generate fictive flats for every bucket in the model which are linearly priced. Out of these flats we assign the flat with the highest surplus to a given household. The initial deposit of households is a specific fraction of their lifetime income.

In the sample model we turn the rental market off, so every household is an owner occupier (if it has already managed to purchase a flat). Since buy-to-let investors rent out flats, this simplification eliminates investment purchases as well. We keep the housing stock constant as well, and so the sample model does not contain renovation nor a construction sector and newly built flats. Finally, we keep GDP and the price level constant, so we disregard economic growth. (However, these features can be activated in the code.)

Using a lower degree of heterogeneity, we can investigate some basic mechanisms with lower number of agents. This enables us to run the model in a longer time horizon, and to investigate the cyclical behaviour of the model. And so despite the many simplifications of the sample model, we add one feature, which is an additional cyclical rule. According to the Expenditure Cascades Hypothesis Levine et al. (2010), when prices go up (down), households may decide to spend more (less) on housing, mimicking the formation of habits or because they try to live in flats similar to that of their peer groups.

We implement this feature by generalizing the utility function by including a cyclical adjuster CA_t (which is set to 1 in the empirical model):

$$RP^f_{t,i} = (FC^{rp}_{t,f,i} + NEIGH^{rp}_{f,i}) * LTI^h_{t,i} * P_t * CA_t + NB^f_{t,i} \tag{A.53}$$

$$CA_t = c^{ca} \times \left(\frac{PI_t/(P_t GDP_t)}{PI_b/(P_b GDP_b)} \right), \tag{A.54}$$

where c^{ca} is an exogenous parameter, and PI_t is the quarterly price index in period t. So the cyclical adjuster captures the change in the price index to the nominal GDP or the change in the affordability of house prices. We can construct the aggregate price index PI_t by calculating and aggregating the quarterly price indices of the neighbourhoods. We can calculate the quarterly price index for a given neighbourhood by regressing the price of the transactions of the 3 previous months and those of a base quarter using the following regression function:

$$\log PRICE = \beta^{PI}_0 + \beta^{PI}_1 \log SIZE + \beta^{PI}_2 \log SIZE^2 + \beta^{PI}_3 QUARTER, \tag{A.55}$$

where $QUARTER$ is 1 for transactions of the last 3 previous months, and 0 for the base quarter. The quarterly price index ($PI^q_{t,i}$) for neighbourhood i is ($PI^q_{t,i} = \exp(\beta^{PI}_3)$). To get the aggregate quarterly price index (PI^q_t), we use the weighted average of the price indices of the neighbourhoods with the number of transactions as weights.

A.9 PARAMETERS

Table A.1: Parameter values of the empirical model

Parameter name	Notation	Value
minimal length of unemployment (in periods, for three educational categories)	p^{um}_c	3, 5, 6
threshold for size deviation ratio for closest neighbour calculation	th^{cns}	0.1
threshold for state deviation ratio for closest neighbour calculation	th^{cnq}	0.1
weight of size deviation ratio	w^{ds}	0.5
weight of state deviation ratio	w^{dq}	0.5
monthly price decrease	c^{mpd}	0.015
discount factor for forced liquidation	DC	0.8
augmentation term to the discount factor for forced liquidation	DC^a	0.5
shock threshold for the activation of the augmentation term	thr^{dc}	0.95
maximum state of flat for land price calculation	$state^{max}_{LP}$	2.07

base construction unit cost	CUC^b	150,000
base construction unit cost markup for Budapest, Pest county	θ^{CUC}	0.5, 0.2
average wage coefficient for construction unit cost	c^{CUC}	0.5
constant of construction markup	$\mu_{C,0}$	-0.1
sold ratio coefficient of construction markup	$\mu_{C,1}$	0.5
monthly price convergence of newly built flats	γ	0.15
number of periods after completion to start price converge	p^{nba}	4
number of fictive flats in a bucket	n^f	5
monthly depreciation rate of the state of flats	δ	0.004
monthly probability of home renovation	pr^r	0.02
base renovation unit cost	RUC^b	100,000
base renovation unit cost markup for Budapest, Pest county	θ^{RUC}	0.5, 0.2
average wage coefficient for renovation unit cost	c_1^{RUC}	0.5
renovation volume coefficient for renovation unit cost	c_2^{RUC}	1.0
maximum lifespan in months	a_{max}	1080
beginning of working age in periods, by educational level	a^w	[260; 240; 216]
retirement age in periods, by educational level	a^p	[780; 780; 780]
pension repalcement rate, by educational level	prr	[0.85; 0.85; 0.85]
number of periods to look ahead to calculate lifetime income	p^{LTI}	240
deposit inheritance ratio	r^{di}	0.3
constant of target savings rate	c_0^{tsr}	-4.5113
wage coefficient of target savings rate	c^{tsr}	0.3125
minimum consumption rate	r^{mc}	0.5
number of periods to reach susbsistence consumption	p^{mc}	48
age in periods to count as adult in subsistence per capita consumption	a^{aspc}	164
weight of additional adults in subsistence per capita consumption	w^{aspc}	0.75
weight of children in subsistence per capita consumption	w^{cspc}	0.5
lower per capita subsistence level	c^{spcl}	60,000
threshold for lower per capita subsistence level	th^{spcl}	100,000
upper per capita subsistence level	c^{spcu}	90,450
threshold for upper per capita subsistence level	th^{spcu}	160,000
constant in additional utility gain of newly built flats	c^{nba_0}	0.4

neighbourhood quality coefficient in additional utility gain of newly built flats	c^{nba_1}	0.06
ratio of households which can consider renovation when making bids	r^r	0.5
renovation interval	int^r	0.1
renovation cap	$state^{max}_{used}$	2.71
additional nonmaterial renovation cost ratio	ν	0.1
share of first-time buyers evaluating a second region (Budapest/other regions)	p^{fb}	[0.25; 0.5]
probability of considering moving at age of mandatory moving decision	pr^m_f	0.7
age in periods of mandatory moving decision	a^m_f	540
probability of considering moving for young households	pr^m_0	0.006
threshold age in periods for lower probability of considering moving	a^m_0	540
moving probability decay parameter	MPD	0.95
threshold deviation rate for moving	thr^m	0.4
number of periods for the calculation of average newly built demand	p^{NPD}	12
target buffer rate of newly built flats	$TNBB$	0.2
duration of construction in periods	p^C	18
cap of new construction to demand	r^{NB}_{max}	2
land area need ratio	r^{LA}	0.2
constant in planned professional investment value	c^{PIR}_0	0.0
coefficient of expected return spread in planned professional investment value	c^{PIR}_1	0.14
cap of professional investment value ratio	r^{max}_{PIMV}	0.01
cap of household investment probability in neighbourhood	pr^i_{max}	0.1
coefficient of expected return spread for household investment probability in neighbourhood	c^i	10
power of expected return spread for household investment probability in neighbourhood	α_i	2
target vacancy rate	TVR	0.05
number of bidding rounds in the housing market	n^r	10
number of flats to evaluate	n^E	400
coefficient in adjustment regarding surplus difference	c^{rpa}	0.05
power in adjustment regarding surplus difference	α_{rpa}	1.25
maximum adjustment as a proportion of surplus difference	RPA^{max}	0.75
adjuster coefficient in probability of placing bid	c^{PPB}	1
maximum number of bids placed per household	n^B	10

weight of adjusted reservation price in bid	w^{RPA}	0.05
threshold market price for bridge loan	thr^{BL}	5,000,000
maximum bridge loan to value	r^{BL}_{max}	0.75
duration of housing loans	p^{ML}	300
duration of renovation loans	p^{RL}	60
loan-to-value cap	LTV^{max}	0.8
number of nonperforming periods to restructure loan contract	p^{npr}	3
number of nonperforming periods to liquidate collateral	p^{npr}	6
income ratio of forced sale	r^{fs}	0.25
rent to price	r^{RP}	0.004
coefficient of utilization ratio in rental markup	c^{μ}_{r}	0.3
power of utilization ratio in rental markup	α^{μ}_{r}	3
length of rental contracts in periods	p^{r}	12
constant in investment sale probability	c^{RS}_{0}	0.0208
coefficient in investment sale probability	c^{RS}_{1}	-0.5208

Appendix B Data generation

B.1 GENERATING THE HOUSING STOCK

Representativeness of the housing stock

The housing stock of Hungary in the model has been generated with the motivation of approximating reality as closely as possible. We attempted to achieve this by using several data sources, most importantly a realtor dataset consisting of 172-214 thousand transactions depending on the year between 2016 and 2020. These records were multiplied in a way to match not only the size of the actual housing stock but also to mimic the empirical distributions of the flat characteristics which are the most relevant for our model. (Although these distributions are based on micro-census data, the Hungarian Central Statistical Office publishes only the aggregated statistics and not the granular data.) To accomplish our objective, the observed flats were scaled up following an elaborated weighting strategy. As a first step, we assessed the differences between the micro-census statistics and the realtor dataset based on the distributions of the variables observed in both cases, and we selected those with the highest deviations. Ideally, one could use all the variables in the data to reweight the observations, however, this would result in countless categories with a very small number of observations in each of them. In the end, we applied iterative proportional fitting, i.e. raking procedure based on the region, the settlement type, the size and the price of the flats[24].

State and neighbourhood characteristics

Besides the above mentioned traits, we also assigned state and neighbourhood characteristics (representing the quality of the flats and their environment) for all the generated flats. To represent each of these with one continuous variable respectively, we had to reduce the high-dimensional space of various characteristics of the flats. In the case of the *state*, we compressed the year of construction, the type of the building, the type of the heating system and the condition of the flat characteristics into one categorical variable, and we divided the flats into 21 categories. Regarding the *location* of flats, we divided the country into actual, interpretable neighbourhoods based on postal codes, the administrative category of the settlements, whether they are in the agglomeration of a larger town, whether they are a touristic destination, the distance from the capital of the county and the distance from Budapest[25]. We created 124 neighbourhoods among which 40 can be found in the capital, and 84 cover the rest of the country. The main principle in this process was to divide the country as homogeneously as possible from the point of view of pricing while keeping the number of observations sufficiently high everywhere to be able to calculate neighbourhood-level price indices.

To obtain continuous values, firstly we estimated a quality measure for both of these categorical variables by running a logarithmic regression on the price of the flats:

$$ln(price)_i = \beta_0 + \beta_1 \times ln(size_i) + \beta_2 \times ln(size_i^2) + \beta_3 \times ln(state_i) + \beta_4 \times ln(nh_i) + \epsilon \qquad \text{(B.1)}$$

As a second step, we estimated another variant of this regression using the continuous version of the variables. (We obtained these continuous values using the coefficients estimated for each of the categories in the first stage.):

$$price_i = \beta_0 + \beta_1 \times size_i \times nh_i + \beta_2 \times size_i^2 \times nh_i + \beta_3 \times size_i \times state_i \times nh_i + \beta_4 \times nh_i + \epsilon \qquad \text{(B.2)}$$

[24] Since the size distribution is very different between the countryside and urban areas, we applied different size categorization depending on the settlement type.

[25] We had to make several data cleaning steps in this process. Most importantly, the agglomeration of some towns were mixed up and had to be disentangled based on postal codes; at some locations we could observe only the name of the settlements, but not their postal codes; the touristic areas in the Danube Bend and at Lake Velencei were too broadly defined, so we narrowed these to contain only the areas directly connected to the water.

To enhance the data quality, we restored the missing location information for 2,000 cases based on the neighbourhood of the flats. Furthermore, based on the results of the regression, we could impute the missing state observations too. As further corrections, we modified the prices using the housing market price indices we calculated.

To make the results more stable we conducted the following outlier filtering steps: (i) We omitted the observations with the top 5-5% error expressed both nominally and in relative terms (9% of the observations altogether), and we rerun the estimation on the remaining sample; (ii) we also omitted the observations where the state values were smaller than 0 or larger than 3.25.

Representativeness of the transactions

To ensure that the transactions on the housing market in the model happen in a realistic way, we assessed the time of the deal, the price, the size, the state and the neighbourhood of the flats in the transactions based on the empirical observations in a dataset coming from the National Tax Authority between 2016 and 2017. As we cannot observe the *state* information for these observation, we calculated it based on its distribution in the 2017 realtor dataset.

We also ensured the representativeness regarding the age of the parties in the transactions. For this, we could observe the year of birth information between 2016 and 2018 in around 80% of the observations, i.e. for 411,000 transactions. Additionally, we observed the age distributions of the inheritance events and the amount of savings.

Disaggregated house price indices

Our granular setup made it possible to calculate not only aggregated, but also county and even neighbourhood and settlement type level housing price indices. We produced monthly and quarterly indices between 2014 and 2018 for three settlement types in the 19 counties (villages, small towns and regional centers); and 22 indices for the districts of Budapest[26]. Furthermore, we also calculated the price indices at the level of neighbourhoods.

The data used in these estimations comes from the National Tax Authority. The methodology of the calculation is a full time horizon, hedonic regression with time dummies to ensure the sufficient number of observations even at disaggregated levels. We applied a two-stage outlier detection: (i) in the first phase we determined absolute lower and upper bounds for the size, the CPI adjusted transaction price and the adjusted price per squared meter ratio; (ii) while in the second step we applied 4 statistical extreme value filtering techniques based on the predicted values of the regression.

Representativeness of rented flats

We determine in the model which households live in rented flats based on empirical observations coming from the above described realtor dataset. To see the ratio of households living in rented flats at a disaggregated level, we use the distribution of the following household and flat characteristics: region, settlement type, flat size category, price category and the income decile of the households. We decide whether a flat will be rented or owned by going over them in multiple rounds until we approximate the empirical distributions sufficiently. During this procedure we introduce a few restrictions: (i) only young households live in rented flats; (ii) and households with housing loans do not rent.

We also consider short-term renting activities to take into account the effects of the Airbnb-type business activities. We determine this external demand expressed in the number of flats in each neighbourhood and in each size and state category based on Boros et al. (2018), Dudás et al. (2018) and Jancsik et al. (2018).

B.2 GENERATING INDIVIDUALS

Similarly to the housing stock, the objective of the individual generating process is to match the size and the empirical distributions of the Hungarian population. To complete this endeavor, our starting point was to use the database of the Central Administration of National Pension Insurance (CANPI) containing information about individuals' age, sex, place of living (at NUTS 1-3 and LAU 1 level) and starting wage. As the CANPI system stores information only about the part of the population

[26] Due to the low number of observations we handled together district 22 and 23.

which is active on the labor market (or receive social benefits after children), we contrasted its aggregated numbers with the yearly demographic statistics coming from the Hungarian Central Statistical Office (HCSO). This way we could make corrections to match the empirical data about the population pyramid of Hungary and we could also match the demographics of each of the counties in Hungary.

Additionally, we had to make some adjustments to account for the fact that young adults tend to start working earlier in the more underdeveloped regions of the country, so their regional distribution is imbalanced in the CANPI statistics. To correct for this, we re-weighted the young population based on the data of more representative age cohorts.

Although we also observe the occupational classification (ISCO-88) of the workers in the CANPI data, the number of employees working in each category changes across cohorts. To stabilize this, we fix the proportions of the occupational classes based on the 34 years old cohort (in the case of the 2018 CANPI observations), because we assume that people usually reach their final occupation at this age.

Finally, we had to determine the age-dependent mortality rates of the individuals, for which we used the 2018 Demographic Yearbook of the HCSO.

B.3 GENERATING HOUSEHOLDS

Additionally to the marital information of the CANPI, we can also identify couples empirically using the Central Credit Register (CCR) of Hungary based on the co-debtors in the loan contracts. If the two debtors of a housing loan contract belong to the opposite sex, they live in the same area and their age difference is at most 10 years we consider them a couple.

However, for that part of the population where we cannot directly observe the relationships, we must match individuals using different strategies. Firstly, we assign a husband to all the women who are having children. We do this by assessing the attributes (location of the place of living, salary, occupation, and age difference) of the members of married couples. We create new couples by adhering to the empirical observations along these characteristics. As a next step, we turn to the individuals who do not have children and younger than 30 years[27]. In these cases, we assign these individuals to parent households (located in the same county), which are selected such that we follow the empirical distribution of the number of children per household. Thirdly, we also create couples without children following similar principles as before but in an ordered way based on the individuals' age. (This way we ensure that the probability of being a single-individual household is higher for older people.) In this step we also utilized the empirically observed probabilities of women becoming married, which we calculated based on the differences of the ratio of married women between consecutive cohorts. Finally, we make sure that there is the same number of households as the number of flats in each county[28].

Besides marital information, the CANPI data also contains records about several social welfare benefits, most notably the baby-care allowance (CSED), childcare allowance (GYES) and child-care benefit (GYED) based on which one can identify birth giving events. If the length of these social transfers exceeds the time interval available after the birth of a child, we assume the birth also of a second child, thus we can also consider the birth order of the children. (However, we disregard the birth of twins or disabled children for whom the length of the social benefits is extended.)

B.4 CLEANING AND IMPUTATION OF THE LOAN CONTRACT DATA

Although we can observe the loan contracts in the CCR data, there are some important missing information and also several data errors which we had to handle based on the observed variables. Regarding the cleaning of the data, we made many corrections, e.g. in the case of maturity we observed some negative numbers; we corrected unrealistic installment and outstanding principal numbers using annuity calculation; sometimes there were also inconsistencies between the loan-to-value (LTV) numbers and the loan characteristics or between the installments and the incomes of the debtors.

However, most importantly we had to calculate the interest rates of the loans as this information is not part of the CCR dataset. We did this in two distinct ways, firstly by using the amount of the outstanding principal, and secondly by using the original loan

[27] In the model we assume that individuals live with their parents until they have children, or until they become 30 years old.

[28] To ensure this, we erase single-individual households where we could not observe the occupational classification.

amount in an annuity formula. Since we know only the year but not the exact date of signing the loan contract (we assume it to be in June in every case) the first version yields more accurate results, but the number of observations is significantly higher in the case of the original loan amount. If we could obtain both results, we took the minimum of the two (but in the end we only kept the values if they are below 20%).

As an alternative approach, we also performed a regression analysis, for which we used not only the CCR data, but also information about the income of the debtors coming from income tax data and additional information about the loan contracts based on the L11 data reporting at the Central Bank of Hungary. The predicted values of this regression can also be used to impute the missing observations, but also we need them to be able to determine the interest rate of newly issued loans in the model.

The regression was estimated by using data only on housing loans in 2017. Firstly, we had to identify individuals belonging to the same household (for which we used the income dataset), and collapse the data to the household level. Furthermore, we filtered the extreme values in the case of the loan amount, the interest rate, the market value of the flat, the income, the age and the LTV variables. In the end we only kept observations where the debtors' age were between 18 and 50; their income exceeded the minimum wage, but remained below monthly HUF 3 million (EUR 9000); the interest rate of the loan is fixed at least for 10 years; the interest rate is between 2-8%. For the sake of efficient applicability in the model, we opted for a simple specification consisting of only three explanatory variables: LTV, the logarithm of the net income of the household, and the age of the debtors.

B.5 GENERATING UTILITY FUNCTION PARAMETERS

To determine the parameters of the function representing households' preferences in choosing flats, we utilized the fact that we can observe many households which are present simultaneously in the CANPI, in the CCR and in the realtor dataset as well. This way we can see these households' income, the educational level and age of the debtors, and also the flats belonging to these households. There are around 9500 households which are observable in all the three above mentioned datasets. However, after cleaning these observations, we could use around 8500 households. (E.g. we removed the observations where the life-time income did not exceed the price of the purchased flat by at least 30%. In these cases at least one of the debtors in the household had zero wage, and the remaining income could not cover consumption expenses sufficiently.)

Based on these pieces of information, we could calibrate each of these households' reservation price function by finding the parameter set for each household, with which the assumed function gives the highest consumer surplus for exactly the flat that they own in reality (or for one which is very similar to it). By assigning unique parameters to every household, they can compare flats in the model realistically, and make optimal decisions for themselves.

The exact reservation price functions of the households depend on their life-time income (I_h^{lt}) and the characteristics of the given flat f the following way:

$$RP_{h,f} = \left(c_h^{rps} \times size_f^{\alpha_h^{rps}} \times (1 + c_h^{rpq} \times state_f^{\alpha_h^{rpq}}) + \frac{\sigma_{1,h}}{(1+\exp(-Q_{N_f}))^{1/\sigma_{2,h}}} \right) \times I_h^{lt} \tag{B.3}$$

where the parameters to calibrate are c_h^{rps}, c_h^{rpq}, α_h^{rps}, α_h^{rpq}, $\sigma_{1,h}$, $\sigma_{2,h}$.

This way we can calculate the maximum amount of money a given household is willing to pay for a flat. By subtracting the actual price of this flat one obtains the consumer surplus, which the households aim to maximize during the calibration process.

To perform this calculation, firstly we determined the expected life-time income of the households. This depends on the starting wage and the educational level of the members in a given household. We assumed that households do not want to spend on house purchases out of their pension income (not even in the form of installments). Consequently, we have taken into account only the incomes which are due until an individual reaches the retirement age (65 years). Furthermore, we also assume that households can take out loans, and the amount available from this source corresponds to the sum of the household members' income for the next 20 years (or for the years until the retirement age, if it is less than 20 years). Hence, households can afford more and more expensive flats as their members get older, and the whole amount of their life-time income becomes available at the age of 45 due to the option of taking out a loan.

We imposed some additional restrictions regarding the potentially available flats in the calibration process for a given household based on three criteria: (i) year of the purchase, (ii) price and (iii) location:

- We only considered flats which were sold in the same time period as the observed purchase of a given household. This way we could recreate the supply-side market conditions to some extent.
- The price range of the flats among which households could choose during the calibration depends on their condition. If the actually purchased flat of the household has low state value, the price deviation cannot be more than 10% downwards and 30% upwards. In the case of flats with average state value both of these thresholds are 20%, while for flats in excellent condition the price deviation can be at most 30% downwards and 10% upwards. This way the reservation price function contains information also about the preferences between a cheaper flat which requires renovation or a more expensive flat which is in better condition.
- If the empirically observed purchase happened in the capital, by default we consider flats in the same district, otherwise we consider flats in the same county. To account for the possibility of migration within the country, we extended the available flat list by adding 500 randomly sampled flats from the capital (within the parameters of the first two constraints above). This represents the typical moving pattern in Hungary from rural areas to the capital.

During the calibration process households choose the flat with the highest consumer surplus from this list. As we want to find the parameters which capture the observed behaviour of the households, we need to evaluate how close the characteristics of the flat chosen during the calibration are to the empirically observed flat of a given household. We calculate the maximum of the percentage differences between the two flats' price, size, state value and neighbourhood quality score, and we minimize the value of this function during the parameter search. We consider the calibration successful if the value of the objective function is less than 20%. (The zero value implies that the household picked the same flat after the optimization as it did in reality.)

To perform this task, we used a stochastic optimization tool called *simulated annealing*, which is an adaptation of the Metropolis–Hastings algorithm. It is well suited to find the global maximum of non-linear, non-differentiable, multimodal functions such as the one in our optimization problem. When choosing the hyperparameters of this process, we usually used the default values of the *optimization* package in R. The lower and upper limits of the parameters were set to enable for extreme preferences.

In the end, the calibration was successful for 3 775 households (i.e. the success rate was 44.6%), among which we had perfect matching in 454 cases.

B.6 MERGING THE DIFFERENT DATA SOURCES

After generating all the agents and other objects, we had to create connections among them. Most importantly, we connected the households to flats, loan contracts, and utility function parameters. We observed some of these connections in the empirical datasets, but we had to generate the rest of the links following the procedure described below.

There were many loan contracts in the CCR data for which there was no CANPI id, i.e. we could not identify the household to the given mortgage contract. To establish these links, we divided the unmatched loan contracts and households into categories which can correspond to each other. Households have been categorized based on the age and income quartiles they belong to and the county[29] they live in. Loans have been divided based on the year of making the contract, the quartile of the amount of the loan and the county of the debtors. Finally, we randomly matched the elements of the corresponding categories. (E.g. a mortgage contract with a high loan amount which was made recently with a client from Fejér county has been assigned to a wealthy household with young adult members in Fejér county.)

After matching the housing loans to households, we had to add the flats and the corresponding reservation price function parameters to these households as well. First we do this for those cases where we observe the value of the flats. Both the housing loan contract data and the housing market transaction dataset contains information about the county and the type of

[29] Although Budapest is not classified as part of any county, in this context we divided it into three pseudo-county parts: city center, outer urban areas and suburbs.

the settlement where the flats are located. Furthermore, for these cases we could create quartile categories in both datasets regarding the value of the flats. (We observe the transaction value in the transaction dataset, and the market value in the housing loan dataset.) This would be enough to form categories (by using these three variables) and perform a random assignment between flats and households. However, in order to match also the unique parameters of the households which govern their reservation price formation we had to calculate households' life-time income. (Consumption is calculated simply as the difference between the life-time income and the value of their flat).

In the next step we assigned flats and parameters to those households who still had loan contracts, but there was no information about the value of their flats. In these cases we tried to find the most suitable matching by choosing the best option for the household among 100 randomly chosen flat and 5 randomly chosen parameter sets. Although this is a crude approach, we had to match almost 700,000 households this way, which made it infeasible to apply more sophisticated solutions.

Finally we dealt with those households which were not present in the CCR data. Here we only used the county and the income quartile versus flat price quartile information to form categories for the random assignment process. If there was not enough flats in a county, we relaxed this constraint and allowed to pick flats anywhere in the region. If even this was too restrictive, we only kept the income and price quartile requirement, and eventually we even had to make some completely random matching.

Appendix C Sensitivity analysis

C.1 ROBUSTNESS ANALYSIS

Figure C.1
Results of the Welch's test statistics for 40 realizations of the main output variables of the model between two parameter sets: original values and 10% perturbed values.

	2018 Q1	2018 Q2	2018 Q3	2018 Q4	2019 Q1	2019 Q2	2019 Q3	2019 Q4
New Loans	0.0001	0.0006	0.0082	0.0033	0.1695	0.0686	0.8244	0.3395
Purcases with loans	0.9222	0.3760	0.2514	0.6962	0.4745	0.2971	0.1141	0.2510
Prices								
Budapest	0.0004	0.0059	0.0620	0.0122	0.0561	0.0353	0.5181	0.1872
Pest county	0.0001	0.0099	0.0208	0.0102	0.7126	0.2480	0.0604	0.0016
N. Hungary	0.0004	0.0156	0.0089	0.1816	0.5042	0.0875	0.2890	0.0890
N. Great Plain	0.0163	0.3039	0.5466	0.8998	0.7897	0.3536	0.8945	0.4206
S. Great Plain	0.0000	0.0059	0.0025	0.1352	0.1787	0.4484	0.6746	0.8353
C. Transdanubia	0.0000	0.0033	0.2829	0.5822	0.8578	0.3797	0.1982	0.0883
S. Transdanubia	0.0017	0.0046	0.0254	0.1045	0.4089	0.0751	0.5358	0.5068
W. Transdanubia	0.0071	0.0163	0.1616	0.3066	0.0456	0.5054	0.5421	0.6770
Transactions								
Budapest	0.0003	0.7851	0.1186	0.0705	0.5573	0.6332		
Pest county	0.0000	0.0015	0.0018	0.0307	0.2902	0.8072		
N. Hungary	0.0001	0.0037	0.0397	0.3903	0.0011	0.0003		
N. Great Plain	0.0003	0.0005	0.0227	0.2450	0.0017	0.0011		
S. Great Plain	0.0000	0.0262	0.0159	0.6277	0.0576	0.0122		
C. Transdanubia	0.0000	0.0057	0.0011	0.0057	0.5411	0.7538		
S. Transdanubia	0.0004	0.0002	0.0009	0.0005	0.4071	0.9955		
W. Transdanubia	0.0001	0.0001	0.0018	0.0025	0.1232	0.6652		
Total	0.0001	0.0000	0.0003	0.0035	0.1005	0.0782		
of newly built flats	0.0001	0.1416	0.0137	0.0293	0.0071	0.5006	0.2251	0.3994
LTV -50	0.2206							
LTV 50-70	0.1158							
LTV 70+	0.8423							
DSTI -20	0.0163							
DSTI 20-30	0.2322							
DSTI 30+	0.0158							
IncomeDecile -7	0.0151							
IncomeDecile 8-9	0.0125							
IncomeDecile 10	0.2891							

Figure C.2
Results of the Welch's test statistics for 40 realizations of the main output variables of the model between two parameter sets: original values and 20% perturbed values.

	2018 Q1	2018 Q2	2018 Q3	2018 Q4	2019 Q1	2019 Q2	2019 Q3	2019 Q4
New Loans	0.0001	0.0043	0.1707	0.0742	0.9387	0.8088	0.2213	0.0318
Purcases with loans	0.0829	0.4940	0.9115	0.9440	0.9840	0.4510	0.1994	0.4356
Prices								
Budapest	0.0042	0.0040	0.1352	0.0847	0.8672	0.6568	0.1423	0.0065
Pest county	0.0028	0.0103	0.0092	0.3716	0.7747	0.6752	0.0318	0.0435
N. Hungary	0.0001	0.0032	0.0121	0.2062	0.4946	0.7479	0.9787	0.6715
N. Great Plain	0.0058	0.6450	0.2856	0.0331	0.7965	0.2252	0.2981	0.1824
S. Great Plain	0.0001	0.0044	0.0160	0.1706	0.1465	0.4696	0.6719	0.7868
C. Transdanubia	0.0029	0.0816	0.2452	0.0909	0.0260	0.0324	0.0059	0.0081
S. Transdanubia	0.0016	0.0214	0.0278	0.9890	0.4136	0.7183	0.2401	0.0967
W. Transdanubia	0.0043	0.1641	0.1290	0.9615	0.8424	0.2715	0.0695	0.0010
Transactions								
Budapest	0.0003	0.0373	0.1960	0.9660	0.1686	0.6003		
Pest county	0.0002	0.7719	0.6712	0.3303	0.1122	0.1420		
N. Hungary	0.0002	0.0514	0.7981	0.1382	0.0000	0.0000		
N. Great Plain	0.0010	0.5845	0.8855	0.2986	0.0003	0.0000		
S. Great Plain	0.0001	0.3648	0.4288	0.5549	0.0249	0.0220		
C. Transdanubia	0.0003	0.3767	0.8166	0.8406	0.0489	0.1873		
S. Transdanubia	0.0016	0.1735	0.6671	0.8677	0.1133	0.0025		
W. Transdanubia	0.0001	0.4357	0.2623	0.5358	0.0720	0.0569		
Total	0.0002	0.8482	0.8744	0.5949	0.0045	0.0056		
of newly built flats	0.0022	0.3208	0.7633	0.0035	0.0052	0.8394	0.0077	0.0140
LTV -50	0.8579							
LTV 50-70	0.2014							
LTV 70+	0.3857							
DSTI -20	0.0265							
DSTI 20-30	0.7096							
DSTI 30+	0.0049							
IncomeDecile -7	0.0118							
IncomeDecile 8-9	0.1116							
IncomeDecile 10	0.0119							

Figure C.3
Results of the median test statistics for 40 realizations of the main output variables of the model across three parameter sets: original values, 10% perturbed values and 20% perturbed values.

	2018 Q1	2018 Q2	2018 Q3	2018 Q4	2019 Q1	2019 Q2	2019 Q3	2019 Q4
New Loans	0.0001	0.0082	0.0247	0.0672	0.7408	0.1496	0.6703	0.3012
Purcases with loans	0.2019	0.3012	0.3012	0.6703	0.2725	0.4966	0.0608	0.2725
Prices								
Budapest	0.0004	0.1225	0.9048	0.6703	0.9048	0.1225	0.2725	0.0273
Pest county	0.0055	0.1496	0.4966	0.4066	0.9048	0.6703	0.0022	0.0018
N. Hungary	0.0608	0.3012	0.3012	0.9048	0.7408	0.0450	0.0136	0.0074
N. Great Plain	0.3012	0.4966	0.4966	0.1225	0.6703	0.0608	0.0074	0.0022
S. Great Plain	0.0006	0.2725	0.2019	0.4966	0.4966	0.7408	0.2725	0.7408
C. Transdanubia	0.0273	0.4966	0.4966	0.4966	0.1496	0.3012	0.0672	0.0608
S. Transdanubia	0.0005	0.0202	0.0821	0.1496	0.1225	0.4066	0.4966	0.1225
W. Transdanubia	0.0608	0.2019	0.0608	0.7408	0.1225	0.2725	0.0672	0.0247
Transactions								
Budapest	0.0074	0.6703	0.2725	0.7408	0.0672	0.0608		
Pest county	0.0000	0.0202	0.0247	0.2019	0.0247	0.3012		
N. Hungary	0.0082	0.0264	0.0247	0.4966	0.0000	0.0000		
N. Great Plain	0.0082	0.0029	0.4966	0.9048	0.0012	0.0000		
S. Great Plain	0.0000	0.0003	0.0450	0.2725	0.3931	0.2019		
C. Transdanubia	0.0005	0.0033	0.1225	0.0821	0.1225	0.1952		
S. Transdanubia	0.0055	0.0821	0.0821	0.0450	0.4066	0.1225		
W. Transdanubia	0.0005	0.0007	0.4066	0.0608	0.1225	0.4966		
Total	0.0000	0.0002	0.1496	0.1496	0.0033	0.0005		
of newly built flats	0.0004	0.9048	0.6703	0.0055	0.0672	0.7408	0.0821	0.0247
LTV -50	0.1225							
LTV 50-70	0.2019							
LTV 70+	0.4966							
DSTI -20	0.0074							
DSTI 20-30	0.7408							
DSTI 30+	0.0004							
IncomeDecile -7	0.0022							
IncomeDecile 8-9	0.0608							
IncomeDecile 10	0.2725							

C.2 SENSITIVITY ANALYSIS

Figure C.4
Number of significant regression estimation results for the considered parameters. (The type of the selected parameters is color-coded.)

	5% significance level			
	Prices	Transactions	Credit market	Total
Number of variables	8	10	11	29
upper per capita subsistence level	4	9	8	21
probability of considering moving at age of mandatory moving decision	3	7	6	16
number of nonperforming periods to restructure loan contract	8	3	4	15
target vacancy rate	3	2	10	15
shock threshold for the activation of the augmentation term	6	4	5	15
age in years of mandatory moving decision	0	5	9	14
minimum consumption rate	4	1	9	14
threshold for lower per capita subsistence level	2	3	6	11
lower per capita subsistence level	1	5	4	10
power of exp. return spread for household investment prob. in neighbourhood	1	2	7	10
constant in investment sale probability	4	0	5	9
threshold for upper per capita subsistence level	1	0	7	8
duration of construction in periods	2	4	2	8
monthly price decrease	1	3	4	8
number of fictive flats in a bucket	0	6	2	8
threshold deviation rate for moving	1	4	3	8
duration of renovation loans	1	3	4	8
coefficient in investment sale probability	0	4	4	8
neighbourhood quality coefficient in additional utility gain of newly built flats	1	2	5	8
threshold age in periods for lower probability of considering moving	1	3	3	7
cap of new construction demand	0	5	1	6
coefficient of expected return spread in planned professional investment value	3	0	3	6
maximum number of bids placed per household	0	2	4	6
monthly price convergence of newly built flats	0	4	1	5
renovation volume coefficient for renovation unit cost	0	3	2	5
power of utilization ratio in rental markup	0	1	4	5
deposit inheritance ratio	0	1	3	4
constant in additional utility gain of newly built flats	0	1	3	4
constant of construction markup	0	0	4	4
power in adjustment regarding surplus difference	1	3	0	4
monthly probability of home renovation	0	4	0	4
threshold for state deviation ratio for closest neighbour calculation	0	2	2	4
maximum adjustment as a proportion of surplus difference	0	4	0	4
probability of considering moving for young households	0	2	1	3
adjuster coefficient in probability of placing bid	0	3	0	3
cap of household investment probability in neighbourhood	0	1	2	3
sold ratio coefficient of construction markup	0	2	1	3
number of periods for the calculation of average newly built demand	0	2	1	3
land area need ratio	0	1	1	2
share of first-time buyers evaluating a second region (other regions)	1	0	1	2
cap of professional investment value ratio	0	1	1	2
weight of adjusted reservation price in bid	0	1	1	2
target buffer rate of newly built flats	1	1	0	2
coeff. of exp. return spread for household investment prob. in neighbourhood	0	1	0	1
maximum state of flat for land price calculation	0	1	0	1
additional nonmaterial renovation cost ratio	0	1	0	1
weight of state deviation ratio	0	1	0	1
share of first-time buyers evaluating a second region (Budapest)	0	0	1	1
coefficient in adjustment regarding surplus difference	0	0	1	1
duration of housing loans	0	1	0	1
number of bidding rounds in the housing market	0	0	1	1
renovation interval	0	1	0	1
threshold for size deviation ratio for closest neighbour calculation	0	0	1	1
weight of size deviation ratio	0	1	0	1
augmentation term to the discount factor for forced liquidation	0	0	0	0
constant in planned professional investment value	0	0	0	0
number of periods after completion to start price convergence	0	0	0	0
number of periods to reach subsistence consumption	0	0	0	0
coefficient of utilization ratio in rental markup	0	0	0	0

Figure C.5
Range of parameter values for the sensitivity analysis.

	Range of values in the the sensitivity analysis
Minimum consumption parameters	
upper per capita subsistence level	72,360-108,540
minimum consumption rate	0.4-0.6
threshold for lower per capita subsistence level	80,000-120,000
lower per capita subsistence level	48,000-72,000
threshold for upper per capita subsistence level	128,000-192,000
Moving behavior parameters	
probability of considering moving at age of mandatory moving decision	0.56-0.84
age in years of mandatory moving decision	36-54
Foreclosure parameters	
number of nonperforming periods to restructure loan contract	5-7
shock threshold for the activation of the augmentation term	0.76-1
Investment behavior parameters	
target vacancy rate	0-0.24
power of exp. return spread for household inv. prob. in neighbourhood	1.6-2.4
constant in investment sale probability	0.017-0.025

Figure C.6
Box plots of 40 realizations for the average prices in different regions with respect to three parameter sets: original values, 10% perturbation, 20% perturbation.

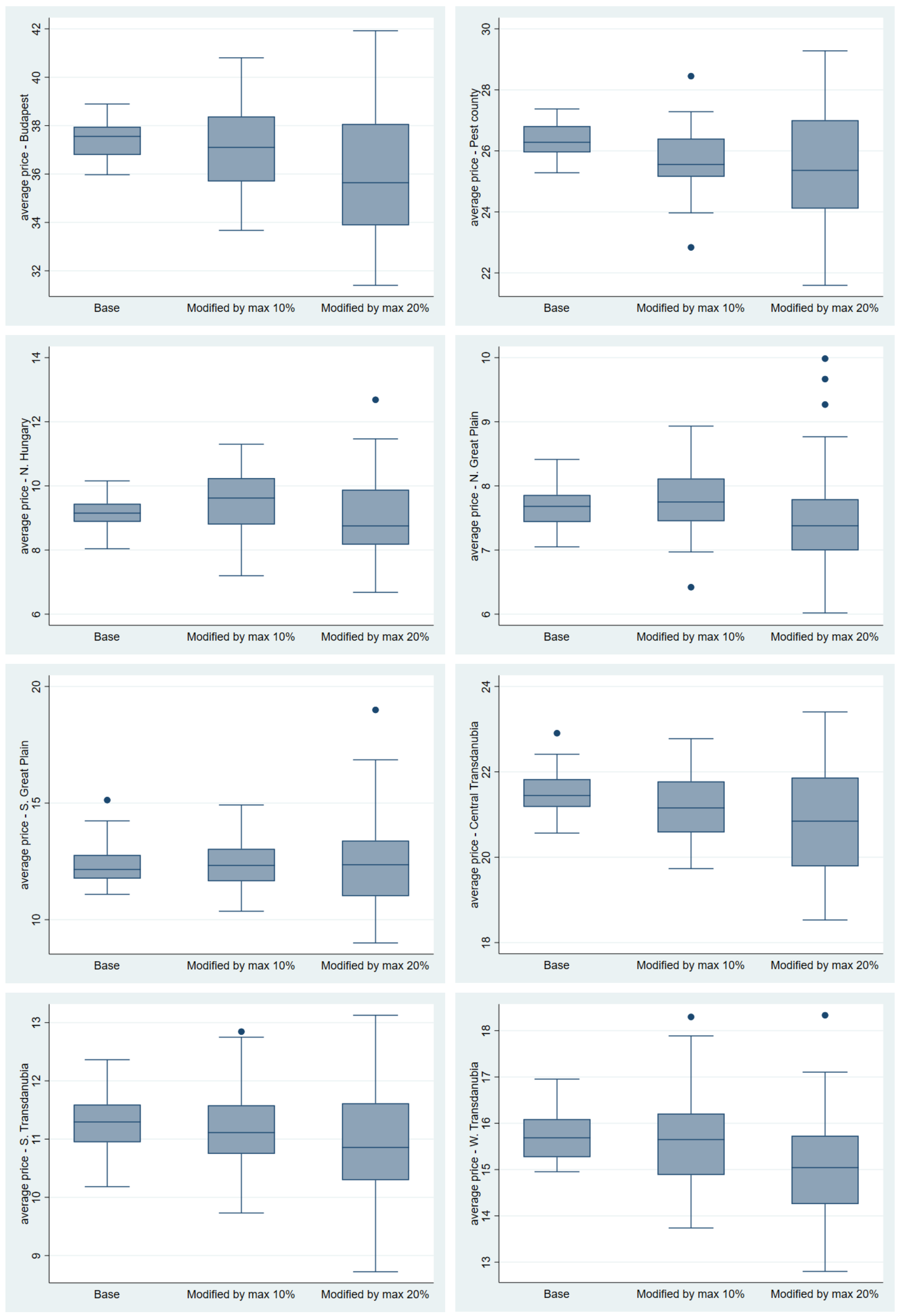


Figure C.7
Box plots of 40 realizations for the average number of quarterly transactions in different regions with respect to three parameter sets: original values, 10% perturbation, 20% perturbation.

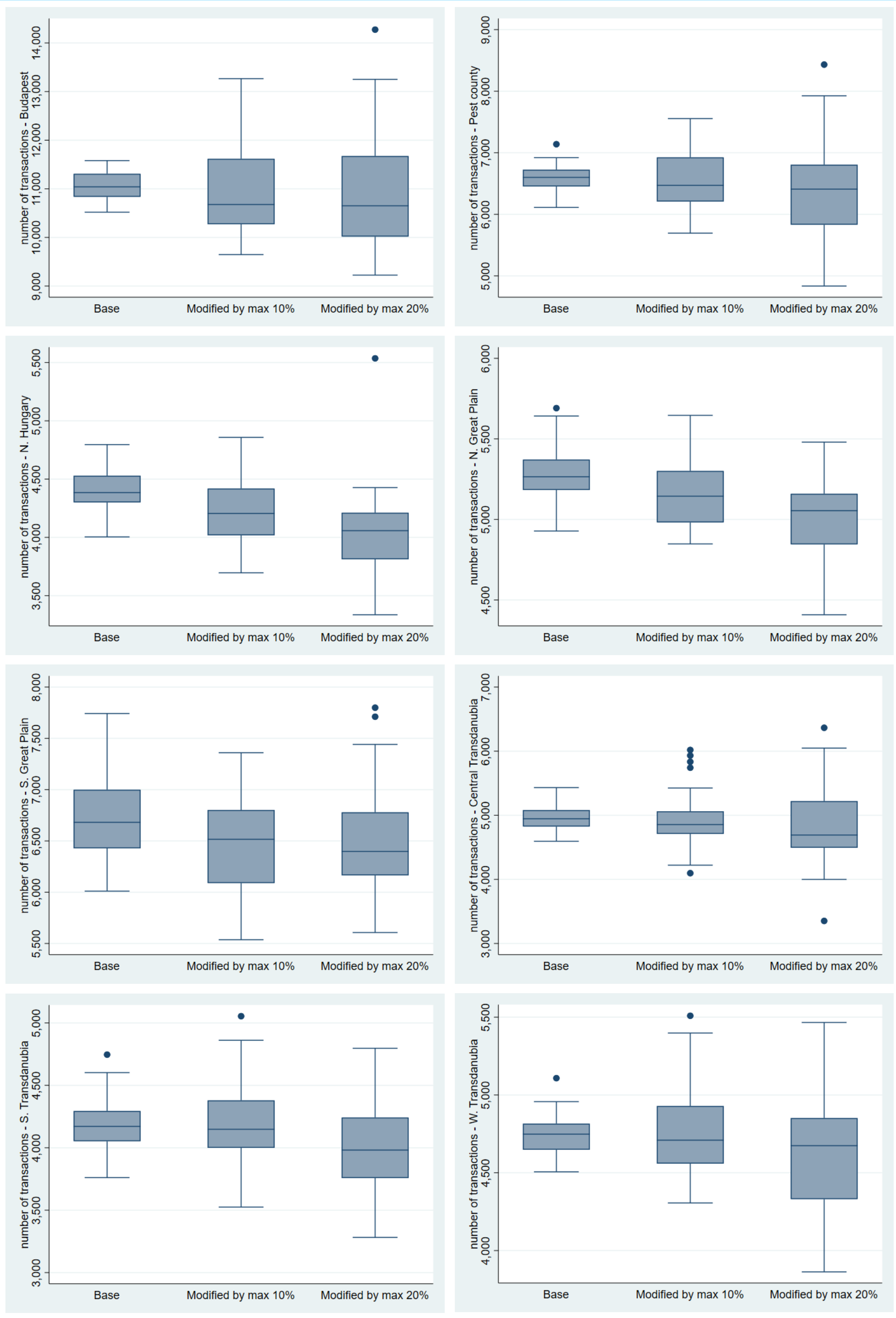


Figure C.8
Box plots of 40 realizations for the share of LTV categories with respect to three parameter sets: original values, 10% perturbation, 20% perturbation.

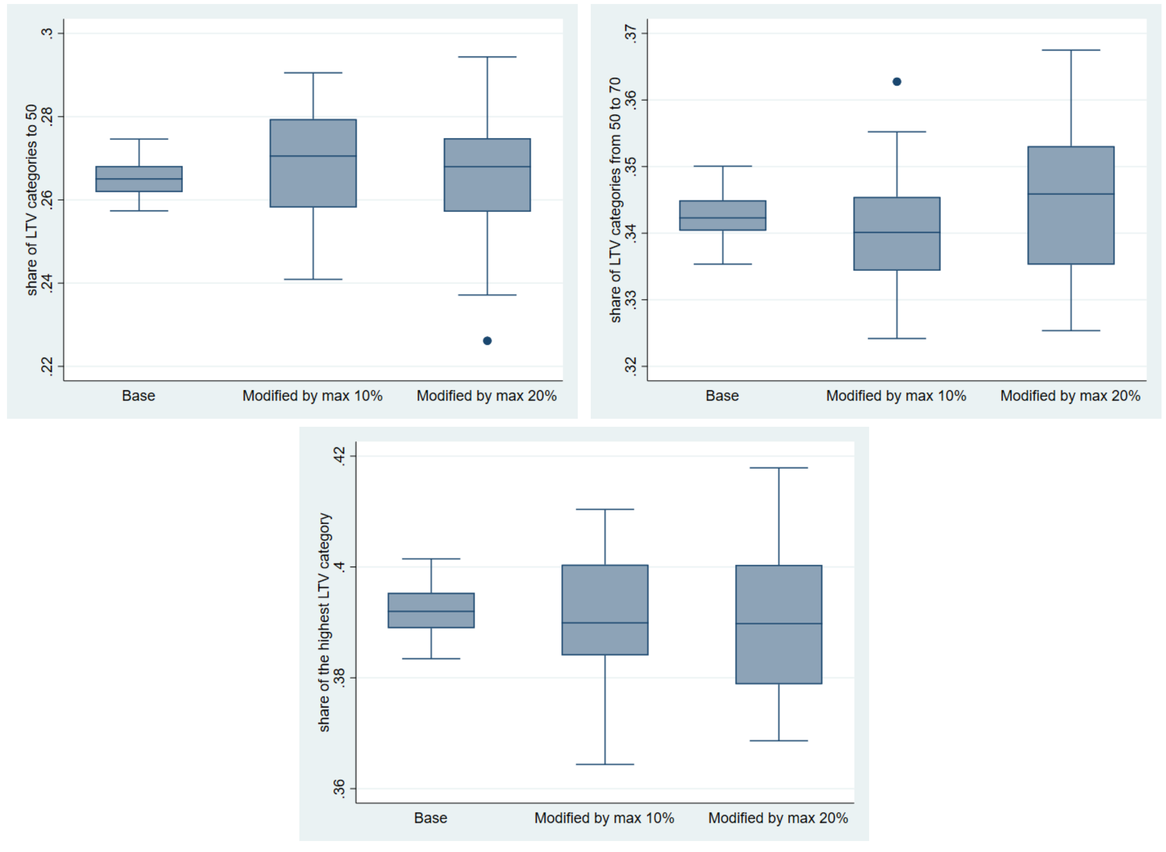


Figure C.9
Box plots of 40 realizations for the share of DSTI categories with respect to three parameter sets: original values, 10% perturbation, 20% perturbation.

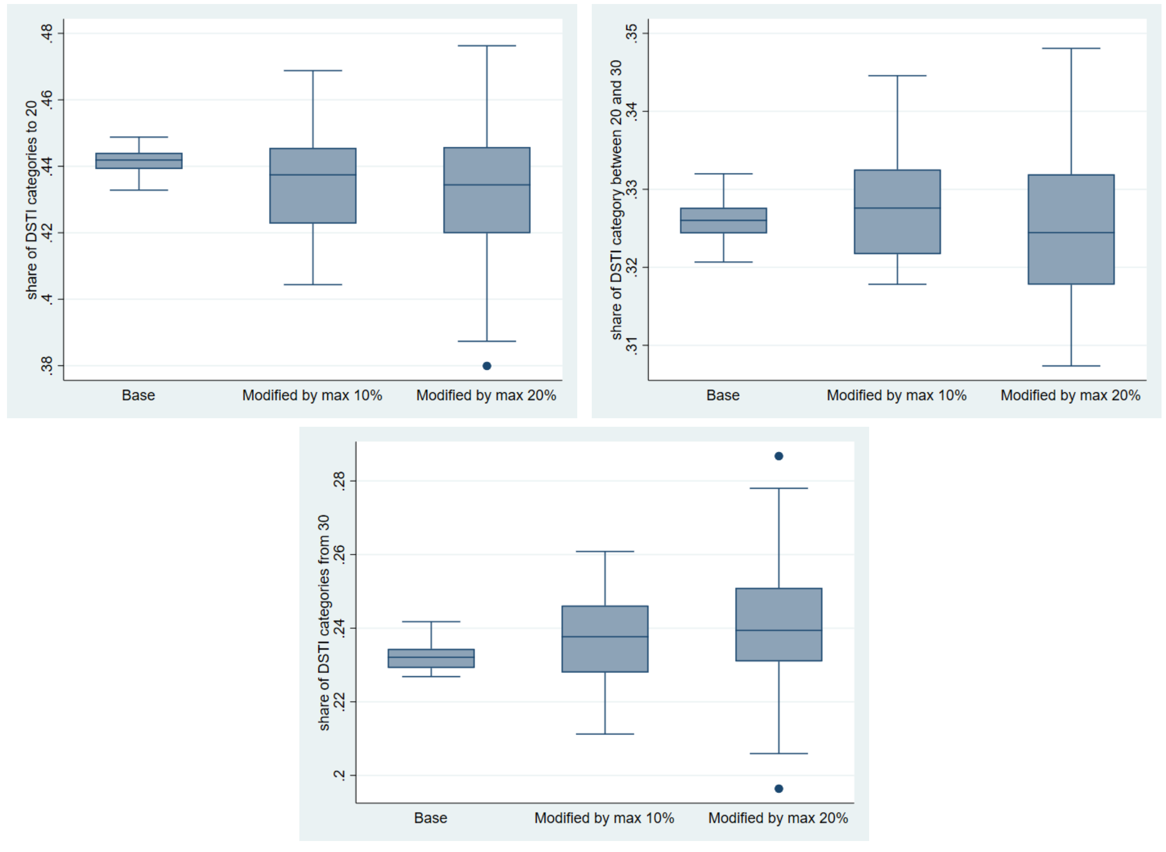
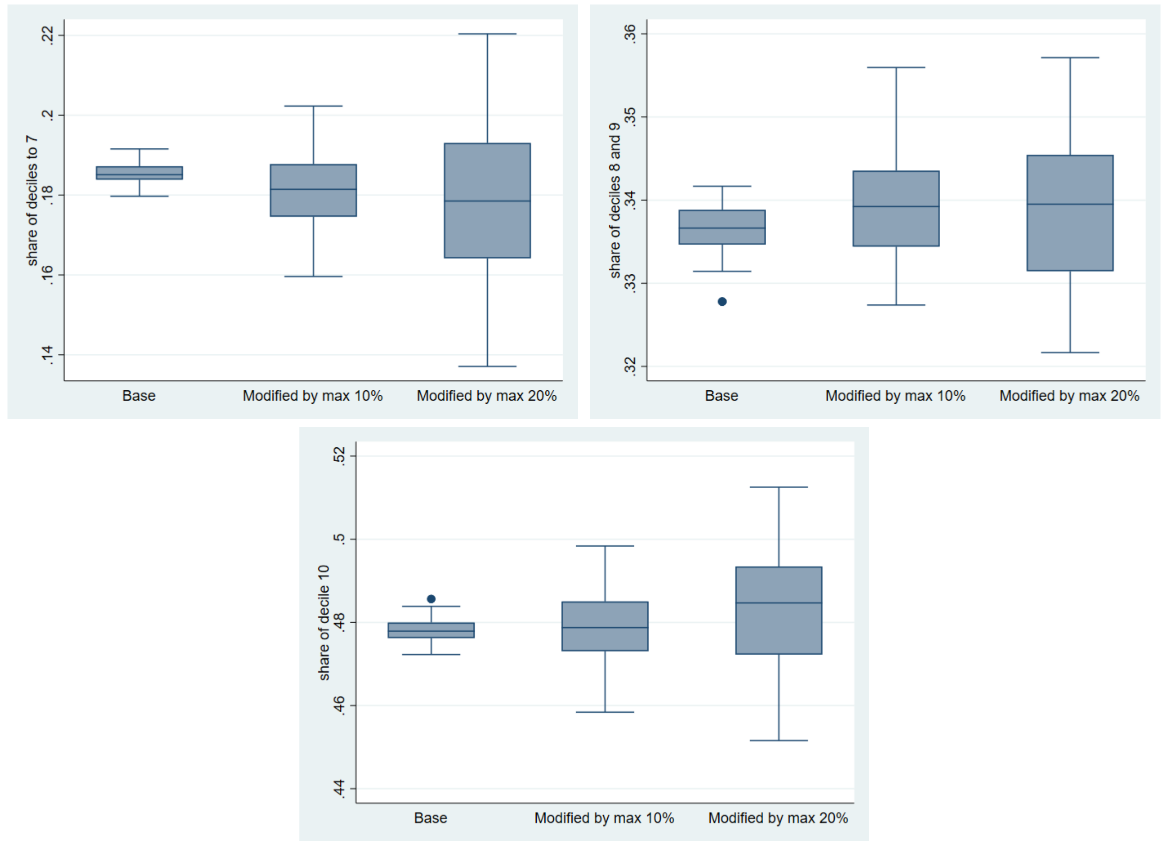


Figure C.10
Box plots of 40 realizations for the share of income deciles with respect to three parameter sets: original values, 10% perturbation, 20% perturbation.



MNB Working Papers 2022/7
A High Resolution Agent-based Model of the Hungarian Housing Market
Budapest, October 2022

mnb.hu

©MAGYAR NEMZETI BANK

1013 BUDAPEST, KRISZTINA KÖRÚT 55.