House Prices and Rents in a Dynamic Spatial Equilibrium

Tuuli Vanhapelto*

November 2, 2022

Most recent version

Abstract

House prices and rents do not always comove across locations and over time. To study the causes and welfare consequences of rent and price variation, I set up a quantitative dynamic spatial equilibrium model of housing demand and supply. In the model, price-to-rent ratios can vary because of differences in expected rental growth or discounting, and data on prices, rents, migration and construction contain identification power to separate between the two. I take the model to data in the case of Finland, where house prices have been diverging across regions, even if rents have not. Through the lens of the model, the rapid price divergence between big and small cities can be rationalized as the combination of an increase in the value of living in cities and regionally diverging discount rates. These changes have led to an important regional divergence of both renter welfare and housing wealth across smaller and larger cities.

JEL: R13, R23, R30

Keywords: Housing, rents, dynamic spatial equilibrium, quantitative spatial models

*tuuli-maria.vanhapelto@tse-fr.eu, Toulouse School of Economics.

I am very grateful to Thierry Magnac, Martial Dupaigne, Essi Eerola, Christian Hellwig and Teemu Lyytikäinen for their help with this project. I also thank Gabriel Ahlfeldt, Christophe Bruneel, Patrick Coen, Fabrice Collard, Alexandre Gaillard, Jonas Gathen, Eugenia Gonzalez-Aguado, Martti Kaila, Carolina Kansikas, Markus Karlman, Nour Meddahi, Sophie Moinas, Niku Määttänen, Peter Neis, Tuukka Saarimaa, Clara Santamaria, Yuta Suzuki, Marko Terviö, Pablo Warnes, Miguel Zerecero, and participants in TSE workshops, Helsinki GSE Lunch seminar and Echoppe 2022 conference for helpful comments and discussion. I thank Yrjö Jahnsson foundation, Jenny and Antti Wihuri foundation and Nordea Pankin säätiö for funding. I thank VATT Institute for Economic Research for providing access to the data on rents from vuokraovi.com and prices from KVKL HSP.

1 Introduction

What can we learn about the distribution of welfare in space by comparing house prices and rents across locations? Regional variation in house prices and rents is important to understand, not only because a large share of household income is spent on housing consumption, but also because housing is the most important asset in household portfolios in OECD countries (OECD 2019, 2021). High prices often coincide with high rents, but recent evidence suggests that house prices and rents do not always comove over time and across locations. For example, in many countries, apartment prices in large cities have grown faster than rents, and price-to-rent ratios have diverged regionally (see, among others, Ahlfeldt et al. (2021) for evidence from Germany, Hilber & Mense (2021) for England, and Chapelle et al. (2022) for France). Yet, the canonical spatial equilibrium model used to study welfare differences in space does not distinguish between rents and prices, and therefore does not help us understand the welfare changes associated with diverging price-to-rent ratios (Rosen, 1979; Roback, 1982; Moretti, 2010).

In this paper, I study quantitatively *the causes* of price and rent divergence across regions and analyse their *welfare consequences* for those who rent and those who own housing in different locations. I address these questions through the lens of a dynamic spatial equilibrium model by incorporating housing markets into the spatial framework of Caliendo, Dvorkin, & Parro (2019), in which households make forward-looking location decisions. The novel feature of the model is a landlord sector which makes forward-looking housing investment decisions. The rental market clears on-the-spot as a function of the current housing demand and supply, but houses are priced as the net present value of future rent flows. As opposed to a standard static spatial model, this creates rationales for price-to-rent ratio variation. I use the model to understand *why* rents and prices evolve differently across regions by inverting the model in order to recover the underlying changes in location-specific economic fundamentals that are consistent with the observed changes in rents, prices, location decisions and construction. I then quantify *who has benefited* from those changes by assessing how the welfare gains from the changes are distributed between those who own houses and those who rent in different locations.

I take the model to data in Finland which provides a good laboratory for studying the separate roles of prices and rents in the housing market. First, while house prices diverge regionally faster than rents in many countries, Finland provides an extreme example of such evolution, as illustrated in Figure 1. Price-to-rent ratios have evolved in opposite directions across regions, with house prices growing faster than rents in large

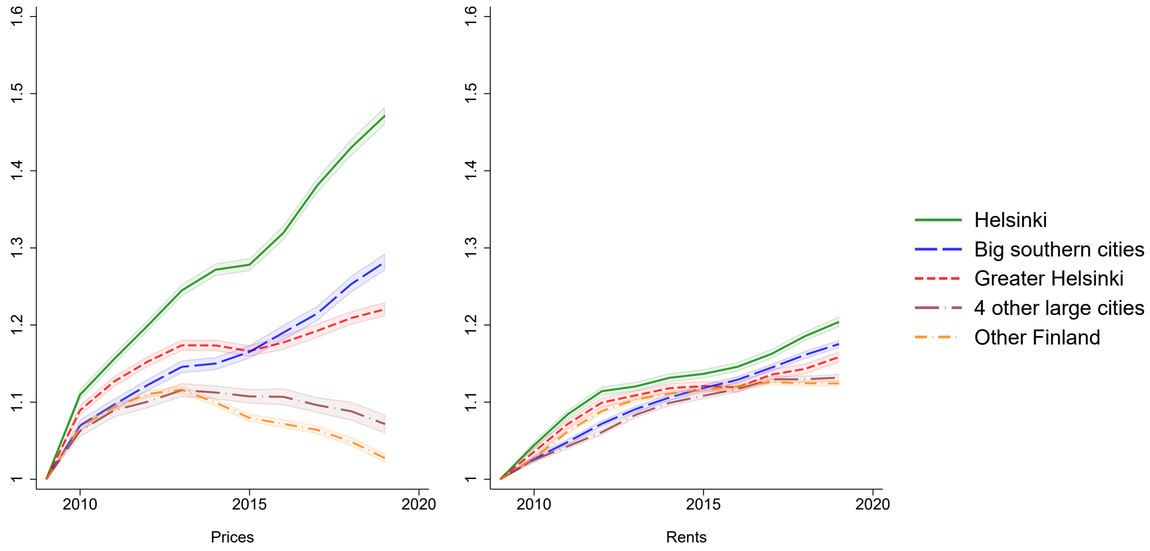


Figure 1: Hedonic indices of apartment prices and rents in Finland.

Notes. Prices are measured using transaction recodings by the Finnish Federation of Real Estate Agency (KVKL). Rents are measured using listing rents from the listings website vuokraovi.com. Both figures display hedonic indices constructed with apartment-level fixed effects and they can be interpreted as resales / relisting indices. Both samples are restricted to non-new apartments in multi-family housing. The shaded area indicates the 95% confidence interval. For details on the methodology, see appendix A.1, the data, see appendix A.2, and further measures of price and rent indices, see appendix A.3.

cities and the reverse occurring elsewhere. Second, Finland provides a relatively clean setting for studying the causes of regional divergence that are related to domestic housing demand and supply, allowing me to abstract away from international investment demand[1].

My first set of results concerns understanding the causes of the variation in price-to-rent ratios. The model nests two broad types of potential causes for price-to-rent variation in a unifying framework. First, price-to-rent ratios can vary if rental growth expectations vary, like in Hilber & Mense (2021) and Molloy et al. (2020). In the model, both construction and (internal) migration take time, so any changes affecting housing demand or supply will have immediate effects on prices following the new information, but more gradual effects on rents following the slow migration and construction responses. Second, changes in discount rates can also cause price-to-rent variation. For example, a decline in the risk-free rate can affect different regions differently if the

[1]Finland is a country where housing demand from foreign investors or short-term rental investors has remained low compared to many European capitals. According to the website "Airdna.co", in September 2022, Helsinki had 1 882 active listings for short-term rental apartments on websites Airbnb and Vrbo, compared to 6 787 in Copenhagen and 12 141 in Lisbon, which are both slightly smaller than Helsinki in terms of population. According to the real estate consulting company JLL, foreign real estate investment companies held approximately 15 000 apartments in Finland in 2019, relative to a total of 2 700 000 habited dwelling units.

regions also differ in housing supply responses or local risk premia, like in Karlman (2022) and Amaral et al. (2022). Since the model accommodates multiple possible explanations for rent-price divergence, it can be used to quantify the relative importance of the different mechanisms.

To do so, I *invert* the model to recover the changes to underlying location-specific economic fundamentals that rationalize the observed data for years 2012-2019. The inversion exercise allows me to recover changes in location-specific fundamentals related to housing demand (location-specific incomes and amenities), housing supply (location-specific land supply) and discounting (location-specific discount rates). As is common in urban economics, these fundamentals are recovered as residuals that rationalize the observed data. For example, amenities are recovered as residuals that rationalize the observed migration flows. Location-specific discount rates are recovered as residuals that rationalize the remainder of price-to-rent variation net of model-consistent rental growth. In a reduced-form way, these rates could capture for example differences in required rates of return by location, as in Levy (2021), or differences in housing depreciation.

I find that in Finland, two key mechanisms help understand the regional divergence of prices even in the absence of divergence in rents. First, the regional divergence in price-to-rent ratios coincides with an important increase in internal migration toward the biggest cities, which my model attributes mostly to amenity growth in cities. The divergence of amenities across locations is significant: in magnitude, the amenity change corresponds to a 20% change in incomes over 7 years. Since the migration responses to amenity changes are slow, future rents are expected to grow faster in large cities than elsewhere, causing some in price-to-rent ratios.

Second, the price-to-rent ratio divergence also suggests that there have been changes in location-specific discount rates. The regional divergence in rents anticipated by the model in response to changes in migration is not sufficient to fully account for the observed divergence of prices, and the remaining variation is interpreted via the residual characterizing location-specific differences discounting. The model-implied discount rates have declined in the 4 largest cities between 2012 and 2019 by on average 0.3 percentage points relative to a baseline of 1.8 percentages. This could reflect for example the simultaneous decline in risk-free rates. On the other hand, the model-implied discount rates have been *increasing* outside the large cities, on average by 0.4 percentage points relative to a baseline of 2.6 outside the 9 largest cities. This could reflect for example a simultanous increase in required rates of return or depreciation rates. Consistent with these hypotheses, discount rates are on an upward trend in locations where the age structure of the population is also changing: in Finland, the population outside the largest cities is aging fast.

The results from the inversion exercise are a key ingredient for the welfare analysis that follows because rents and prices are not alone informative about welfare changes. For example, rents could go up in response to a change that decreases renter utility, such as a negative housing supply shock, or in response to a change that increases renter utility, such as an amenity increase. The inversion exercise shows that other variables, including internal migration and construction, contain identification power to distinguish between these cases.

My second set of results addresses the welfare consequences of location-specific shocks in a spatial model. This is a central question of urban economics: location-specific shocks such as amenity or income increases could benefit either those who live in a specific location or those who own housing in that location, if the wage or amenity increases also push up rents and therefore house values. This incidence depends in part on household mobility and in part on housing supply responses (see discussion in Moretti (2010)). However, crucially, both migration and housing supply elasticities are different in the short run (neither housing construction nor migration can respond much) and in the long run (where both can respond), which suggests that accounting for transition dynamics for rents and prices can improve the understanding of the welfare effects.

The inversion exercise suggests that amenities have increased in large cities relative to other locations, and this translates to important regional variation in renter welfare. The difference in renter welfare between the location that benefited the most and the location that benefited the least is in the order of magnitude of 8 percentage points of consumption-equivalent variation. The welfare differences across locations, although important, are substantially mitigated by the possibility of migrating. When I exclude the possibility of migrating when computing the values, the dispersion in welfare differences more than doubles. Migration works as a buffer that reduces regional welfare differences after local shocks. For individuals to whom migration is not a possibility in reality, the welfare divergence can therefore be significantly larger than what my baseline estimates suggest.

The regional divergence of renter welfare is sizeable, but the divergence in landlord welfare, given by the value of their housing stock, is even greater. Those who own housing in big cities benefited from the higher anticipated future rents and the lower discount rates implied by the model relative to an equilibrium in which no changes would have taken place, and housing wealth in the biggest cities grew by 15-20% over seven years. On the other hand, those who own housing outside the largest 9 cities saw significant wealth declines relative to an equilibrium in which no changes would have taken place, and there, housing wealth declined by up to 20%. Jointly, this implies a dramatic regional divergence of housing wealth.

Finally, I evaluate the effects of a counterfactual policy that would lead to a positive (or negative) wage shock in some locations, but not others. One such policy is the proposed railroad tunnel that would connect the metropolitan areas of Helsinki and Tallinn to a single labor market area. I use the model to assess what would happen to rents and house prices as well as renter and landlord welfare if such a tunnel was constructed and it had a positive effect on household earnings in the Helsinki region. I find that the location-specific income shock would have asymmetric effects for renters and landlords across locations. Renters in all locations would benefit, but landlords outside the capital region would face capital losses. Thus the income increase in the capital region would amplify house price differences across regions.

Literature. The literature on spatial equilibrium models (Rosen, 1979; Roback, 1982) views regional housing costs as a location-specific congestion mechanism. Households are mobile across locations, but housing is not perfectly elastic in supply: thus equilibrium housing costs will adjust to ensure that marginal households are indifferent between living in different locations. A small literature on dynamic spatial equilibrium models with housing markets highlights the importance of understanding house prices as assets (Davis et al., 2013, 2014; Glaeser et al., 2014; Halket & Vasudev, 2014; Yoon, 2017; Herkenhoff et al., 2018). Particularly relevant for my application is Van Nieuwerburgh & Weill (2010) who use a dynamic spatial equilibrium model to study the role of regional income divergence in explaining house price divergence. I contribute to these papers by studying both rent and price divergence in a quantitative dynamic spatial equilibrium.

To do so, I complement the quantitative spatial model of Caliendo et al. (2019) by modeling the accumulation of housing capital. Quantitative spatial models in this line of papers exploit tractability from closed-form expressions for value functions, which makes it possible to solve for equilibrium transitions in time differences. I incorporate housing accumulation in a similar way as Kaplan et al. (2020), and show that with this approach, I can apply the model solution technique from Caliendo et al. (2019). Previously, housing has been modeled in quantitative spatial models with a fixed supply (Zerecero, 2021; Warnes, 2020; Balboni, 2019). Two papers close to mine are Kleinman et al. (2021) and Suzuki (2021), who have shown how to model the accumulation of production capital in these frameworks. The key difference in modeling housing capital from production capital is that housing is produced using land, and land is fixed in supply, implying that the price of an additional unit of housing capital can vary over time and across locations.

I estimate the model using an inversion technique which is a common strategy in urban economics. Ahlfeldt et al. (2020) and Kleinman et al. (2021) suggest techniques for inverting dynamic quantitative spatial models. I contribute to the literature first by leveraging the fact that housing (or capital) prices are informative about the net present value of future returns. Second, I provide a fast algorithm for the numerical model inversion which allows me to implement the inversion even if there are unanticipated changes to economic fundamentals and without imposing restrictions on agents' choices.

I use the model to understand spatial variation in price-to-rent ratios. The classical decomposition of J. Y. Campbell & Shiller (1988) applied to housing tells us that price-to-rent ratios can fluctuate through three channels: future rental growth, future risk-free interest rates, or the housing risk premium (Goodman, 1988; S. D. Campbell et al., 2009). The roles of credit and expectations on the housing market are also explored in Kaplan et al. (2020); Favilukis et al. (2017); Greenwald & Guren (2021), but outside a spatial setup. In a spatial setup, Hilber & Mense (2021); Büchler et al. (2021); Molloy et al. (2020); Bischoff (2012); Howard & Liebersohn (2020) study location-specific rent growth expectations in explaining regional divergence in rent-price ratios, highlighting regional supply constraints and regional growth rates. Another strand of the spatial literature studies the role of declining interest rates in explaining regional housing market divergence (Karlman, 2022; Miles & Monro, 2021; Amaral et al., 2022). I contribute to this literature by setting up a model that nests the key suggested mechanisms for rent-price ratio divergence from these two sets of papers in a unifying framework, which then allows me to evaluate the relative importance of these different causes empirically.

I also contribute to the literature on the capitalization of local shocks to house prices in spatial equilibrium (Moretti, 2010, 2013; Hsieh & Moretti, 2019; Hornbeck & Moretti, 2022; Notowidigdo, 2020). Only a small set of papers addresses this question through the lens of a structural model, which is needed to make welfare analyses. Giannone et al. (2020) and Greaney (2022) study the effects of location-specific shocks on house prices with heterogenous-agent models. Ahlfeldt et al. (2020) study the effects of location-specific shocks on rents in a quantitative spatial model. As opposed to these papers, my setup can speak to rent and price differences, which is important for understanding welfare when the two do not comove. Cun & Pesaran (2022) study the effects of location-specific shocks on house prices and rents, and I contribute to their work by modeling dynamic renter utility. This is important because the option to migrate in the future is an important factor in mitigating regional differences in household utility after local shocks, as I will show.

Roadmap Section 2 presents the dynamic spatial model with housing markets. Section 3 presents the identification strategy of this paper which is the model inversion exercise. Section 4 presents the data sources as well as the sources of the externally calibrated parameters and provides descriptive evidence of regional divergence in Finland. Section 5 reports the implied economic fundamentals. Section 6 provides a discussion on the welfare implications of the main results and contains a counterfactual policy evaluation. Section 7 concludes. The Appendix contains further evidence of regional price and rent divergence in Finland, a descrpition of the data sources, computational details and proofs.

2 A dynamic spatial model of the housing market

This section presents a dynamic spatial equilibrium model of the housing market. The model encompasses two types of forward-looking behavior: Households make forward-looking location choices and landlords make forward-looking housing investment decisions. Housing markets are incorporated into a model in the spirit of Caliendo et al. (2019) by modeling housing as a location-specific investment that is produced using scarce land as an input, like in Kaplan et al. (2020). This is done in a way that preserves the original, tractable equilibrium solution technique in time differences.

The economy consists of a discrete set of L locations, indexed by l, d and k. Time is discrete and indexed by t, s and z. There are three types of agents: households, landlords and developers.

Households choose where they desire to reside among the discrete set of locations. They earn location-specific income and consume it on housing services and nonhousing consumption. Households rent housing services from landlords, giving rise to a demand curve for housing services.

Landlords operate as intermediaries: They own houses and supply rental housing services to households. Each period, they choose how much new housing to purchase from the development sector, taking into account the value of anticipated future rents, which gives the demand curve for buildings each period.

The development sector purchases land from the government, employs land and labor to build houses and sells the houses to the landlord sector in each location. The development sector faces a tradeoff between larger buildings and higher construction costs, leading to an upward-sloping construction supply curve.

Homeownership is not explicitly modeled. Modeling the ownership of houses via landlords allows to separate the household migration decisions from housing investment decisions. This assumption buys important tractability, but it implies that the model cannot be used to quantify welfare implications of shocks to owner-occupiers (who live in some location, derive utility from location-specific incomes and amenities in that location, and also own housing in that location). The welfare effects they experience are a combination of the welfare effects on renters and landlords in the model.

The production sector of the nonhousing consumption good operates outside the model, using labor as an input and producing the consumption good as an output. Wages are exogenously given by location-specific productivities[2]. Moreover, I abstract away from location-specific prices of nonhousing consumption as I view these prices of second-order importance for regional migration decisions relative to regional rent differences[3]. The tradable nonhousing consumption is elastically supplied for a unit price in each location.

Government sets the quantities of building permits issued each period and location. The government also collects a capital gains tax on rental income from landlords. Government spends their revenues outside the model.

There are two types of exogenous parameters in the model. First, there are a set of deep, time-invariant parameters describing the behaviour of agents, which are labeled as "structural parameters". Second, there are a set of location- and time-specific exogenous variables (location-specific incomes, amenities, buildable land and discount rates used by the landlord sector). They are labeled as "economic fundamentals".

Agents have perfect foresight regarding all the all exogenous and endogenous aggregate quantities of the model. This assumption implies that the model cannot be used to make structural claims about investment risk, which will be captured in a reduced-form sense. The only uncertainty is related to households' idiosyncratic location preferences. However, there can be unanticipated changes to the sequences of exogenous variables.

All derivations are presented in appendix D.

2.1 Households

Households living in location l in time period t supply inelastically one unit of labor and earns a location-specific, exogenous income $w_{l,t}$. They derive utility from the consumption of housing services $h_{l,t}$ and from nonhousing consumption $c_{l,t}$, as well as

[2]Modeling labor market clearing is not necessary as there is another congestion mechanism, rents, which ensures that the economy is in a spatial equilibrium.

[3]Thus, I abstract away from congestion effects in other markets than the rental market.

location-specific amenities $A_{l,t}$. Within a period, they solve

$$\max_{c,h} \log A_{l,t} + \phi \log(c) + (1-\phi)\log(h) \tag{1}$$

$$\text{s.t. } w_{l,t} = c + r_{l,t} h \tag{2}$$

where ϕ denotes the budget share of nonhousing consumption, and $r_{l,t}$ is the price paid for one unit of housing services in one time period (the rent). The price of the consumption good (the numeraire) is 1 uniformly across locations. The indirect utility of a household living in location l in time t is

$$u(w_{l,t}, r_{l,t}, A_{l,t}) = \log A_{l,t} + \log(w_{l,t}) \; - \; (1-\phi)\log r_{l,t} \; + \; \tilde{\phi} \tag{3}$$

where $\tilde{\phi}$ is a constant. Households do not have access to a savings technology so they live hand-to-mouth.

Households are mobile across locations, but changing locations requires paying a migration cost, which makes the location choice a dynamic problem. A household living in location l in time t makes a forward-looking decision in choosing whether they wish to move to a different location in the next time period. Household i's value of being in location l in time period t is

$$v_{l,t}(\epsilon_{i,t}) = u(w_{l,t}, r_{l,t}, A_{l,t}) + \max_{d \in 1,\ldots,L} \left[\beta \mathbb{E}_{\epsilon}\big(v_{d,t+1}\big) - \tau^{l,d} + \eta \epsilon^{d}_{it}\right] \tag{4}$$

where households optimize over the next period location d. Households discount the future value with discount factor β. In the following period, a household obtains the value of living in the optimal location d in the next period, net of moving costs from l to d, $\tau^{l,d}$. The location decision is also affected by idiosyncratic, household-i specific preference shocks ϵ^{d}_{it}, which follow a type-1 extreme value distribution with a scale parameter 1 and a location-parameter $-\gamma$, where γ is the Euler's constant (so ϵ^{d}_{it} has mean 0). The idiosyncratic shocks are scaled with parameter η, reflecting the elasticity of migration to incentives to move.

Information structure In period t, agents anticipate with perfect foresight the sequences of location-specific, time-varying exogenous variables, referred to as the "economic fundamentals". They consist of the exogenous location-specific sequences of amenities, $\{A_{l,t+s}\}_{s=0}^{\infty}$, $\forall l$ and wages, $\{w_{l,t+s}\}_{s=0}^{\infty}$ $\forall l$, as well as land supply constraints and interest rates (to be specified later). Denote by Θ_t the sequences of economic fundamentals across all locations anticipated by agents in period t.

Migration probabilities Following the dynamic discrete choice literature and Caliendo et al. (2019), the distributional assumption on the idiosyncratic preference shocks implies that household ex-ante value function $V_{l,t}$ (expectation over the idiosyncratic preference shocks) can be expressed as

$$V_{l,t}(\Theta_t) = \mathbb{E}_\epsilon(v_{l,t}) = u(w_{l,t}, r_{l,t}, A_{l,t}) + \eta\big[\log \sum_{k=1}^{L} \exp(\beta V_{k,t+1}(\tilde{\Theta}_{t+1}) - \tau^{l,k})^{1/\eta}\big] \tag{5}$$

where $\tilde{\Theta}_{t+1} = \mathbb{E}(\Theta_{t+1}|\Theta_t)$. This notation is used to highlight that in period t, agents decide their locations for $t+1$, and these decisions are made using period-t anticipated values of economic fundamentals. Given perfect foresight, if $\Theta_t = \{x_{t+s}\}_{s=0}^{\infty}$ then $\tilde{\Theta}_{t+1} = \{x_{t+s}\}_{s=1}^{\infty}$.

Moreover, the probability that a household migrates from location k to location d can be expressed as

$$\mu_t^{k,d} = \frac{\exp(\beta V_{d,t+1}(\tilde{\Theta}_{t+1}) - \tau^{k,d})^{1/\eta}}{\sum_{l=1}^{L} \exp(\beta V_{l,t+1}(\tilde{\Theta}_{t+1}) - \tau^{k,l})^{1/\eta}} \tag{6}$$

Household allocations Let us denote the number of households residing in location l in period t by $N_{l,t}$. The law of motion for households is then given by the migration probabilities

$$N_{l,t+1} = \sum_{k=1}^{L} \mu_t^{k,l} \, N_{k,t} \tag{7}$$

Demand for rental housing From the household within-period optimization, housing demand of a single household is given by $h_{l,t} = (1-\phi)\frac{w_{l,t}}{r_{l,t}}$. The aggregate demand for housing in location l in time period t is then given by

$$H_{l,t}^{demand}(N_{l,t}, w_{l,t}, r_{l,t}) = N_{l,t} \cdot h_{l,t}(w_{l,t}, r_{l,t}) = N_{l,t}\,(1-\phi)\,\frac{w_{l,t}}{r_{l,t}}$$

2.2 Landlords

In each location, there are immobile landlords who own location-specific housing. They purchase durable houses from developers and rent periodic housing services to households. Both rental and construction markets are competitive, and landlords take the sequences of current and future prices and rents as given.

A unit of housing yields rent $r_{l,t}$ each period, and landlords take the future sequence of rents $\{r_{l,t+s}\}_{s=0}^{\infty}$ as given. Landlords can purchase new housing, $q_{l,t}$, for price $p_{l,t}^{Q}$, and they take this price as given. Landlords in location l in period t discount the future revenues from period s with discount factors $\{\rho_{l,t,t+s}\}_{s=0}^{\infty}$, which are exogenous to the model. Landlords are financially unconstrained.

In period t, a representative landlord in location l decides on the quantity of new housing investment by maximizing the net present value of future profits

$$\max_{q_{l,t}\geq 0} \sum_{s=0}^{\infty} \rho_{l,t,t+s}\pi_{l,t+s} \tag{8}$$

$$\pi_{l,t} = r_{l,t}^{N} h_{l,t} - p_{l,t}^{Q} q_{l,t} \tag{9}$$

$$r_{l,t}^{N} = (1-\Upsilon)(1-\xi_l) r_{l,t} \tag{10}$$

$$h_{l,t+1} = (1-\delta) h_{l,t} + q_{l,t} \tag{11}$$

where π is the linear profit function. Landlord revenue is given by their current housing stock $h_{l,t}$ times a net rent r^N, which is the rent paid by households net of operating costs ξ_l and a tax of rate Υ[4]. The housing stock depreciates at rate δ. Landlords' choice variable is the amount of new housing that they purchase, $q_{l,t}$.

The discount factor $\rho_{l,t,t+s}$ is given by

$$\rho_{l,t,t+s} = \prod_{z=0}^{s} \frac{1}{1+i_{l,t+z}}$$

where $i_{l,t}$ is the discount rate in location l in time t. The location-specific discounting is a reduced-form way of capturing possibly different required returns on investment in different locations and different times[5].

Demand for housing structures Landlord utility is linear in profits, so under a transversality condition

$$\lim_{s\to\infty} \rho_{l,t,t+s} r_{l,t+s} = 0$$

housing is valued at the net present value of future rents, net of depreciation. Given a sequence of rents, landowners are indifferent between purchasing any quantity of new

[4]The operating costs refer to costs such as the heating and cleaning of the public spaces of the building, and they do not therefore correspond to maintenance charges. As shown in section A.2, these operating costs vary across locations but not over time, and therefore they are modeled as time-invariant.

[5]Since the model does not have uncertainty related to housing investment, these location-specific discount rates do not have a structural interpretation as stochastic discount rates.

housing from developers for price p^Q, given by the net present value of future rents

$$p_{l,t}^{Q,Demand} = \sum_{s=1}^{\infty} \rho_{l,t,t+s} (1-\delta)^{s-1} r_{l,t+s}^N \tag{12}$$

where the first rent from new housing purchased in period t is received in $t+1$.

The value of holding h units of housing in location l in time t is linear in in the quantity h. The only difference between a unit of new housing, purchased from the construction sector, or a unit of pre-existing housing is that new housing which is purchased in t yields the first rent in period $t+1$, whereas existing housing yields rent also in period t. Therefore, landlords are indifferent between trading existing houses with each other for a unit price that equals the unit value of housing capital

$$p_{l,t} = \sum_{s=0}^{\infty} \rho_{l,t,t+s} (1-\delta)^{s} r_{l,t+s}^N \tag{13}$$

Supply for rental housing Since the net rent $r_{l,t}^N$ is always positive, landlords find it optimal to let out their entire housing holdings, and the supply of rental housing is equal to the current housing stock:

$$H_{l,t}^{supply} = H_{l,t}$$

where $H_{l,t}$ is the current aggregate housing stock in location l.

Law of motion for housing stock New housing purchased in period t can be rented out to households for the first time in period $t+1$. The aggregate housing stock in location l evolves according to

$$H_{l,t+1} = (1-\delta) H_{l,t} + Q_{l,t} \tag{14}$$

where $Q_{l,t}$ is the aggregate amount of new housing investment by landlords in location l.

2.3 Developers

There is a competitive development sector which develops houses in each location using land and labor as inputs. Each period, the government auctions an exogenously given amount of land $\overline{L}_{l,t}$ for development (this quantity can be thought of as building permits). Developers bid for the permits so that in equilibrium, the sector makes zero profits. Developers employ land and labor to produce housing with a CRS technology, and then sell the houses to landlords. The land permits must be used in the period in which they were acquired, implying that the development decision is static.

The representative development firm in location l solves

$$\max_n p^Q_{l,t} Q - w_{l,t} n - p^L_{l,t} \overline{L}_{l,t} \tag{15}$$

$$\text{s.t. } Q = n^\gamma \overline{L}^{1-\gamma}_{l,t} \tag{16}$$

where $p^L_{l,t}$ is the price of land, $p^Q_{l,t}$ the price of new housing (paid by the landowner sector) and $w_{l,t}$ the cost of labor n. Q is the amount of new housing built, and γ is the CRS construction technology parameter.

Supply of housing structures The supply of new housing depends on the house price, wage and land supply as

$$Q_{l,t} = \Big(\frac{\gamma p^Q_{l,t}}{w_{l,t}}\Big)^{\frac{\gamma}{1-\gamma}} \overline{L}_{l,t} \tag{17}$$

meaning that the price elasticity of housing construction w.r.t. house prices is $\frac{\gamma}{1-\gamma}$, or conversely the inverse housing supply curve writes

$$p^{Q,Supply}_{l,t}(Q_{l,t}, w_{l,t}, \overline{L}_{l,t}) = \frac{w_{l,t}}{\gamma}\Big(\frac{Q_{l,t}}{\overline{L}_{l,t}}\Big)^{\frac{1-\gamma}{\gamma}} \tag{18}$$

2.4 Market clearing

Construction market clearing Equating the flat inverse housing demand from the landlords and the upward-sloping inverse housing supply from the development sector, new construction is equal to

$$Q_{l,t} = (w_{l,t}^{-\frac{\gamma}{1-\gamma}})\big(\gamma\, p^Q_{l,t}\big)^{\frac{\gamma}{1-\gamma}} \overline{L}_{l,t} \tag{19}$$

Therefore, conditional on the current construction price (which depends on the sequence of future rents) and current wages and land supply, the house price that clears the construction market is given by a static market clearing condition.

Rental market clearing Equating household demand for housing with the housing stock, the rent that clears the rental market is given

$$r_{l,t}(N_{l,t}, H_{l,t}, w_{l,t}) = \frac{(1-\phi) w_{l,t} N_{l,t}}{H_{l,t}} \tag{20}$$

which is a static market-clearing condition conditional on current allocations $N_{l,t}$, $H_{l,t}$ and current wages $w_{l,t}$.

2.5 Equilibrium

There are two endogenous location-level state variables in the economy, the distribution of households across locations L and the distribution of the housing stock H. The time-invariant exogenous parameters of the model consist of the structural parameter vector $\theta = (\beta, \eta, \phi, \delta, \gamma, \Upsilon)$ as well as the time-invariant moving costs $\tau^{d,k}\ \forall\ d, k$ and the operating cost shares $\xi_l\ \forall\ l$. The time-varying exogenous location-specific variables consist of the location-specific sequences of wages $\{w_{l,t}\}_{l=1,t=0}^{L,t=\infty}$, amenities $\{A_{l,t}\}_{l=1,t=0}^{L,t=\infty}$, land supplies $\{\overline{L}_{l,t}\}_{l=1,t=0}^{L,t=\infty}$ and discount rates $\{i_{l,t+s}\}_{t=0,s=0}^{t=\infty,s=\infty}$. Θ_t summarizes the sequences of "economic fundamentals" that are anticipated by agents in time t for all locations
$\Theta_t = \{\ A_{l,t+s},\ w_{l,t+s},\ \overline{L}_{t+s},\ i_{l,t+s}\ \}_{s=0,\,l=1}^{\infty,\,L}$.

Definition 1. [Sequential competitive equilibrium] Given the initial distribution of labor and housing across locations, $N_{l,t-1}$ and $H_{l,t-1}\ \forall l$, the time-invariant parameters of the model (θ, $\tau^{d,k}\ \forall\, d, k$, $\xi_l\ \forall l$) and the anticipated sequences of the time-varying, location-specific economic fundamentals Θ_t (amenities, wages, land supplies, discount rates), the sequential equilibrium consists of the sequences of endogenous variables $\{N_{l,t}, H_{l,t}, Q_{l,t}, \mu_t^{l,d}, V_{l,t}, p_{l,t}^{Q}, r_{l,t}\}_{t=0}^{\infty}\ \forall l, d$, which solve, for all t, l : the household dynamic problem given by equation 4, landlord problem given by 8, developer problem given by 15, and the corresponding laws of motion given by 7 for households and 14 for the housing stock, as well as within-period market clearing given by 19 for the construction market and 20 for the rental market.

Definition 2. [Stationary equilibrium] A stationary equilibrium is a sequential competitive equilibrium such that all time-variant fundamentals, $\{A_{l,t},\ w_{l,t},\ \overline{L}_t,\ i_{l,t+s}\}$ as well as the endogenous variables $\{N_{l,t}, H_{l,t}, Q_{l,t}, \mu_t^{l,d}, V_{l,t}, p_{l,t}^{Q}, r_{l,t}\}\forall l, d$ are constant for all t.

2.6 Solving for the equilibrium

I solve the model on a transition path towards a steady-state. This is important on the one hand because I want to be able to analyse the short- and long-run effects of shocks, and on the other because I want to be able to analyse nonstationary divergence patterns like the one in Figure 1.

In this subsection, I show that I can apply the so-called "dynamic hat algebra" results from Caliendo et al. (2019) to solve the model with housing markets. The idea is that conditioning on observed allocations, the model can be solved in time differences. Propositions 2.1 and 2.2 are adapted from propositions 2 and 3 in Caliendo et al. (2019) to the model with housing markets. Wherever possible I use notation similar to Caliendo et al. (2019). Start by denoting time differences of a variable by $\dot{x}_t = x_t/x_{t-1}$. Let us define a convergent sequence of changes to a variable as a sequence of changes, $\{\dot{x}_t\}_{t=0}^{T}$, which converges to 1 by time T at the latest. In other words, after time period T, variable x is constant. Define also some additional notation: denote $u_{l,t} = \exp(V_{l,t}(\Theta_t))$ and $\omega_{l,t} = \exp(u(A_{l,t}, w_{l,t}, r_{l,t}))$.

Proposition 2.1. (*Solving the sequential equilibrium given an anticipated sequence of discount rates and an anticipated sequence changes in fundamentals.*)

Given

1. *an initial observed allocation of the economy* $N_{l,0}$, $H_{l,0}$, $Q_{l,-1}$, $\mu_{-1}^{l,k}$, $r_{l,0}$, $p_{l,0}$, $\forall l, k$
2. *sequences of location-specific discount rates* $\{i_{l,s}\}_{s=0}^{s=\infty}$ $\forall l$
3. *a convergent sequence of changes in economic location-specific fundamentals* A, w, $\overline{L}$, *relative to their initial levels under perfect foresight,* $\{\dot{A}_{l,t+s}, \dot{w}_{l,t+s}, \dot{\overline{L}}_{l,t+s}\}_{s=0}^{s=\infty}$ $\forall l$

the sequential equilibrium solves equations

$$\mu_{t+1}^{k,d} = \frac{\mu_t^{k,d} \cdot \dot{u}_{d,t+2}^{\beta/\eta}}{\sum_{l=1}^{L} \mu_t^{k,l} \cdot \dot{u}_{l,t+2}^{\beta/\eta}} \tag{21}$$

$$\dot{u}_{l,t+1} = \dot{\omega}_{l,t+1}\Big(\sum_{l=1}^{L} \mu_t^{k,l} \cdot \dot{u}_{l,t+2}^{\beta/\eta}\Big)^{\eta} \tag{22}$$

$$N_{l,t+1} = \sum_{k=1}^{L} \mu_t^{k,l}\, N_{k,t} \tag{23}$$

$$\dot{Q}_{l,t} = \dot{\overline{L}}_t \cdot \dot{p_{l,t}^{Q}}^{\frac{\gamma}{1-\gamma}} \cdot (\frac{1}{\dot{w_{l,t}}})^{\frac{\gamma}{1-\gamma}} \tag{24}$$

$$Q_{l,t} = \dot{Q}_{l,t} Q_{l,t-1} \tag{25}$$

$$H_{l,t+1} = (1-\delta) H_{l,t} + Q_{l,t} \tag{26}$$

$$\dot{r}_{l,t} = \dot{w_{l,t}} \cdot \dot{N_{l,t}} \cdot \frac{1}{\dot{H_{l,t}}} \tag{27}$$

$$p_{l,t}^{Q} = \rho_{l,t,t+1} r_{l,t+1}^{N} + \rho_{l,t,t+1} p_{l,t+1}^{Q} \tag{28}$$

Proof: Appendix E.1.

Proposition 2.1 tells us that we do not need to know the *levels* of the economic fundamentals A, w, $\overline{L}$ in order to solve the sequential equilibrium, only their changes relative to the baseline economy. Proposition 2.1 tells us also that we can solve the sequental equilibrium in time differences by solving a set of nonlinear equations without having to, for example, solve numerically for any value functions, making the computational burden very small relative to most common dynamic models.

Consider next what would happen if agents were anticipating given sequences of economic fundamentals in period $t-1$, but they learned in period t about changes to these anticipated sequences. Suppose that in period $t-1$, agents were anticipating sequences given by x, and denote the sequences anticipated in t by primes x'. Denote the new anticipated sequences of changes by $\dot{x}'_s = x'_s/x'_{s-1}$. Denote the ratio between the new anticipated sequences of changes and the previously anticipated sequences of changes by $\hat{x}_s = \dot{x}'_s/\dot{x}_s$. In other words, $\hat{x}_s = 1$ means that whatever change agents anticipated in period $t-1$ to happen for variable x in period s is the same as what they anticipate now in period t, after the new information.

Let us call the sequential equilibrium under the new anticipated sequence of changes (period t and after) to exogenous variables the *counterfactual* equilibrium. Let us call the sequential equilibrium under the old anticipated sequences (period $t-1$ anticipated sequences) the *previous* equilibrium.

Proposition 2.2. (*Solving the sequential equilibrium given an unanticipated change to economic fundamentals.*)

Given

1. *a previous economy from period t onwards,* $\{N_{l,t+s}, H_{l,t+s}, Q_{l,t+s}, \mu^{l,k}_{t+s}, r_{l,t+s}, p_{l,t+s}\}_{s=0}^{\infty}$ $\forall l, k$
2. *previous and counterfactual sequences of location-specific discount rates,* $\{i_{l,t+s}\}_{s=0}^{s=\infty}$ $\forall l$ *and* $\{i'_{l,t+s}\}_{s=0}^{s=\infty}$ $\forall l$
3. *the initially anticipated sequence of changes for fundamentals,* $\{\dot{A}_{l,t+s}, \dot{w}_{l,t+s}, \dot{\overline{L}}_{l,t+s}\}_{s=0}^{s=\infty} \forall l$
4. *counterfactual convergent sequences of changes in* A, w, $\overline{L}$, $\{\dot{A}'_{l,t+s}, \dot{w}'_{l,t+s}, \dot{\overline{L}}'_{l,t+s}\}_{s=0}^{s=\infty}$ $\forall l$

the solution to the counterfactual sequential equilibrium, $\{N'_{l,t+s}, H'_{l,t+s}, Q'_{l,t+s}, \mu'^{l,k}_{t+s}, r'_{l,t+s}, p'_{l,t+s}\}_{s=0}^{\infty}$

$\forall l, k$ *solves equations*

$$\mu_t'^{k,d} = \frac{\mu_{t-1}'^{k,d} \cdot \dot{\mu}_t^{k,d} \cdot \hat{u}_{d,t+1}^{\beta/\eta}}{\sum_{l=1}^{L} \mu_{t-1}'^{k,l} \cdot \dot{\mu}_t^{k,l} \cdot \hat{u}_{l,t+1}^{\beta/\eta}} \tag{29}$$

$$\hat{u}_{l,t} = \hat{\omega}_{l,t} \Big(\sum_{l=1}^{L} \mu_{t-1}'^{k,d} \cdot \dot{\mu}_t^{k,d} \cdot \hat{u}_{d,t+1}^{\beta/\eta} \Big)^{\eta} \tag{30}$$

$$N_{l,t+1}' = \sum_{k=1}^{L} \mu_t'^{k,l} N_{k,t}' \tag{31}$$

$$\dot{Q}_t' = \dot{\overline{L}}_t' \cdot \dot{p}_{l,t}^{Q\,\prime\frac{\gamma}{1-\gamma}} \cdot (\frac{1}{\dot{w}_{l,t}'})^{\frac{\gamma}{1-\gamma}} \tag{32}$$

$$Q_t' = \dot{Q}_t' \cdot Q_{t-1}' \tag{33}$$

$$H_{l,t+1}' = (1-\delta) H_{l,t}' + Q_{l,t}' \tag{34}$$

$$\dot{r}_{l,t}' = \dot{w}_{l,t}' \cdot \dot{N}_{l,t}' \cdot \frac{1}{\dot{H}_{l,t}'} \tag{35}$$

$$p^{Q'}{}_{l,t} = \rho_{l,t,t+1}' r^{N'}{}_{l,t+1} + \rho_{l,t,t+1}' p^{Q'}{}_{l,t+1} \tag{36}$$

Proof. Appendix E.2.

Similar to proposition 2.1, proposition 2.2 tells us that we can solve the counterfactual equilibrium in time differences without knowing the levels of A', w', $\overline{L}'$. Notice that period-$t-1$ migration probabilities and construction values are *not* needed, which is useful, because this will allow us to solve for the equilibrium even if there is news arriving in subsequent periods.

2.7 Welfare measures

I will use household ex-ante value functions to measure renter welfare. The value function 5 can be rewritten as

$$V_{l,t}(\Theta_t) = u(w_{l,t}, r_{l,t}, A_{l,t}) + \eta \log \frac{1}{\mu_t^{l,l}} + \beta V_{l,t+1}(\tilde{\Theta}_{t+1}) \tag{37}$$

$$= \sum_{s=0}^{\infty} \beta^s \Big[u(w_{l,t+s}, r_{l,t+s}, A_{l,t+s}) + \eta \log \frac{1}{\mu_{t+s}^{l,l}} \Big] \tag{38}$$

(see, for example, Arcidiacono & Miller (2011) and Caliendo et al. (2019)), highlighting that indirect utilities and migration probabilities summarize the relevant information for computing household values.

Consider a baseline equilibrium value $V_{l,t}$ and a counterfactual equilibrium value in the same location and the same time period, $V'_{l,t}$. I will define compensating variation as the scalar χ_l such that:

$$V'_{l,t}(\Theta'_t) = \sum_{s=0}^{\infty} \beta^s \left[u(w'_{l,t+s}, r'_{l,t+s}, A'_{l,t+s}) + \eta \log \frac{1}{\mu'^{l,l}_{t+s}} \right] \tag{39}$$

$$= \sum_{s=0}^{\infty} \beta^s \left[u((1+\chi_l) \cdot w_{l,t+s}, r_{l,t+s}, A_{l,t+s}) + \eta \log \frac{1}{\mu^{l,l}_{t+s}} \right] \tag{40}$$

In other words, the compensating variation measures the permanent proportional change in incomes that households should receive in the baseline equilibrium, holding constant the equilibrium allocation, to be indifferent between the baseline allocation and the counterfactual allocation.

I also decompose the welfare effects to the part of utility increase that is related to the current location, and to the part that is located to the migration value. To do so, I compute a compensating variation for "stayers", which is the same as equation 39 but where migration is "shut down" when computing the value

$$V'^{stayer}_{l,t}(\Theta'_t) = \sum_{s=0}^{\infty} \beta^s \left[u(w'_{l,t+s}, r'_{l,t+s}, A'_{l,t+s}) \right] \tag{41}$$

$$= \sum_{s=0}^{\infty} \beta^s \left[u((1+\chi^{stayer}_l) \cdot w_{l,t+s}, r_{l,t+s}, A_{l,t+s}) \right] \tag{42}$$

Landlord welfare is given by their wealth, the value of their housing stock $p_{l,t}h_{l,t}$, where $p_{l,t}$ is the unit value of capital.

3 Inverting the model

In this section, I show how to identify location-specific "economic fundamentals" from location-specific observables by inverting the model. The inversion exercise seeks to understand *why* are rents and prices evolving as they do by recovering changes in unobserved economic fundamentals from observed variables. The inversion is similar in spirit to Ahlfeldt et al. (2020) and Kleinman et al. (2021), but I contribute to the existing approaches in two ways. First, I lever information in house prices as a part of the inversion, with the idea that house (or capital) prices reflect the net present value of future returns. Second, I propose a fast algorithm for cases where the inversion needs to be done numerically (for example, if there are unanticipated changes to economic fundamentals). This improves the flexibility of the model inversion, and for example, in my application, restrictions on agents' choice sets are not necessary.

Overview of the model inversion I treat as observed the location-specific endogenous quantities from the model: prices, rents, construction, migration, housing stock and household allocations. Moreover, in the inversion exercise, the structural parameter vector
$(\beta, \eta, \phi, \delta, \gamma, \Upsilon, \xi_l)$ is treated as known, and it will be parameterized later. I show that these observables can be mapped to changes in the exogenous location-specific fundamentals of the model: wages, amenities, land supply and discount rates in each location.

The economic fundamentals are recovered as residuals (or *wedges*) that rationalize the observed engogenous outcomes. The residuals are useful for distinguishing between the two different sources of price-to-rent ratio variation: future rental growth and discounting. Intuitively, increasing rent growth expectations are related to a positive housing demand shock or a negative housing supply shock, which are reflected in data on migration and construction. Location-specific discount rates are recovered as a residual that rationalizes the remaining price-to-rent variation.

In a first step, I show that in each period, market-clearing equations can be inverted to recover information on wages and land supplies. Model-consistent wages are recovered as a residual that rationalizes observed rents, because households consume a constant share of their income on housing. Model-consistent land supply is recovered as the residual that rationalizes observed construction quantities.

In a second step, I show that using techniques from the conditional choice probability literature, household values and migration costs can be recovered from migration probabilities. The intuition for migration costs is that if flows between two locations are high, then the migration costs are low. The intuition for household values is that if many households choose a specific location, it must provide a high value.

In a third step, I recover amenities as a residual that rationalize the values that are recovered from migration flows. Intuitively, the part of utility that cannot be accounted for by rents or wages is explained with amenities.

In a fourth step, I recover location-specific discount rates as the residuals that rationalize the observed house prices, given all other parameters of the model.

Timing The timing assumed in the inversion exercise is as follows: We observe an initial allocation of the economy in some period t, such that in period t, agents anticipate the same sequences of exogenous economic fundamentals as they did in period $t-1$. This implies that no new information arrives in period t. This period is referred to as the "initial period" and the equilibrium implied by this allocation as the "baseline equilibrium". From period $t+1$ onwards in each period agents might (or might not) receive new information in the beginning of each period about wages, amenities, land supplies and landlord sector discount rates.

3.1 Inverting the static market clearing equations

The two static market clearing equations for each location can be inverted to recover current wages and land supply constraints in a given location and any time period.

By inverting the rental market clearing equation, for a given ϕ, model-consistent wages can be recovered as

$$w_{l,t} = r_{l,t} \frac{H_{l,t}}{(1-\phi)N_{l,t}} \tag{43}$$

where rent $r_{l,t}$, housing stock $H_{l,t}$ and number of households $N_{l,t}$ are observed[6]. Since the budget share of housing consumption is constant, rents and the quantity of housing consumed by households are directly informative about their incomes.

From the market clearing equation for the construction market, for a given value of γ, current model-consistent supply of building permits can be recovered as

$$\overline{L}_{l,t} = \left(\frac{\gamma p^{Q}_{l,t}}{w_{l,t}}\right)^{-\frac{\gamma}{1-\gamma}} Q_{l,t} \tag{44}$$

where the prices of new houses $p^Q_{l,t}$ and new construction $Q_{l,t}$ are observed, and wages $w_{l,t}$ backed out previously. If in two locations, construction prices and wages are the same but in one of them, the observed construction is higher in the other one, this location is interpreted to have a higher land supply.

The above equation treats the price of new buildings, $p^Q_{l,t}$, as observed, while it might not necessarily be observed in the data. I treat the price of *existing* buildings, $p_{l,t}$, as observed, together with rents, and recover the price of *new* buildings $p^Q_{l,t}$ from

$$p_{l,t} = r^{N}_{l,t} + (1-\delta)p^{Q}_{l,t} \tag{45}$$

(see appendix D.4).

[6]Since regional wages are also directly observed, we could use them directly in the empirical exercise as well. Similarly, for H and N we observe both empirical counterparts and model-consistent counterparts, which might not be equal (for example, since our depreciation estimate δ might not be the correct one), and either one might be used. If we want the model to reproduce exactly the prices and rents observed in the data, we should use the model-consistent estimates for both.

3.2 Recovering household values

This subsection shows how migration costs and values of living in different locations can be identified from migration flows.

Assumption 3.1. (*Symmetry of migration costs.*) *Migration costs are symmetric* ($\tau^{l,k} = \tau^{k,l}$) *and there is no migration cost to not changing locations* ($\tau^{l,l} = 0$).

Proposition 3.1. (*Recovering migration costs.*) *Under assumption 3.1 and conditional on time-invariant parameters β and η, migration costs can be recovered from migration flows, and they are given by*

$$\tau^{k,d} = \frac{1}{2} \log \left[\left(\frac{\mu^{k,d} \mu^{d,k}}{\mu^{k,k} \mu^{d,d}} \right)^{-\eta} \right] \tag{46}$$

Proof: Appendix F.1.

From the symmetry assumption on migration costs it follows that the cost of migrating from k to d can be pinned down simply by the migration probabilities from k to d and d to k as well as the probabilities of remaining in these locations (as previously shown by Bryan & Morten (2019) and Zerecero (2021)). Intuitively, high migration flows between two locations are associated with low migration costs.

From equation 46 it becomes clear that the migration costs could be recovered separately for each year, since they depend on migration probabilities, which are observed annually. However, I treat migration costs as time-invariant, as they seem to have barely changed over my estimation period[7].

Proposition 3.2. (*Recovering values.*) *Given migration cost parameters $\tau^{k,d}\,\forall k, d$, conditional on time-invariant parameters β and η, differences in values of living across locations can be recovered from migration flows, and they are given by*

$$V_{k,t+1}(\tilde{\Theta}_{t+1}) - V_{d,t+1}(\tilde{\Theta}_{t+1}) = \frac{\eta}{\beta} \left[\ln(\mu_t^{l,k}) - \ln(\mu_t^{l,d}) \right] + \frac{1}{\beta} \left[\tau^{l,k} - \tau^{l,d} \right] \tag{47}$$

Proof: Appendix F.2.

Proposition 3.2 builds on the symmetry assumption 3.1 in that we can proceed sequentially and use the migration cost estimates obtained from 3.1 to then apply 3.2. Once the migration costs are recovered, then the differences in values of living in location k and location d are also identified solely from the choice probabilities, as in Hotz & Miller (1993).

[7] In practice, I estimate migration costs separately for each year in the data and then use the average across years as the time-invariant measure.

Proposition 3.2 suggests that migration probabilities in a dynamic spatial model have a similar interpretation to market shares in static discrete choice models: If a particular alternative is chosen by many individuals, then it must provide a high mean utility. If more people choose location k than location d, then it must be that either migrating to k is more affordable than migrating to d, or that households expect to get a higher utility in k. Moreover, since the migration costs are modeled as time-invariant, then if migration to k increased by more than migration to d from one year to the other, it must also be that the anticipated value of living in k increased by more.

3.3 Recovering initial amenities

In order to recover period-t wages and land supply quantities as well as migration costs and differences in household values, we did not have to assume anything about what agents anticipated in period t regarding the evolution of the economic fundamentals. However, this is not the case for recovering amenities. In order to recover current amenities, we must specify whether some new information arrives in the beginning of period t or not.

Start by considering the case in which agents hold the same beliefs about the sequences of economic fundamentals in period t as in period $t-1$.

Assumption 3.2. (*No unanticipated changes.*) *Suppose we observe migration probabilities in some periods $t-1$ and t s.t. there are no news arriving in period t.*

Proposition 3.3. (*Recovering amenities under no unanticipated changes.*) *Given migration costs $\tau^{k,d}$ $\forall k, d$, location-specific incomes $w_{l,t} \forall l$ and time-invariant parameters β, η and ϕ, under inversion assumption 3.2, location-specific amenities relative to a reference location in period t can be recovered from observed migration probabilities in periods $t-1$ and t and observed rents in period t as*

$$\frac{A_{k,t}}{A_{d,t}} = \exp\Big(u(w_{k,t}, r_{k,t}, A_{k,t}) - u(w_{d,t}, r_{d,t}, A_{d,t}) \\ -\big[\log(w_{k,t}) - \log(w_{d,t})\big] \;+\; (1-\phi)\big[\log r_{k,t} - r_{d,t}\big]\Big) \tag{48}$$

where

$$u(w_{k,t}, r_{k,t}, A_{k,t}) - u(w_{d,t}, r_{d,t}, A_{d,t}) = \\ V_{k,t}(\tilde{\Theta}_t) - V_{d,t}(\tilde{\Theta}_t) - \eta\big[\log\frac{1}{\mu_t^{k,k}} - \log\frac{1}{\mu_t^{d,d}}\big] - \beta\big[V_{k,t+1}(\tilde{\Theta}_{t+1}) - V_{d,t+1}(\tilde{\Theta}_{t+1})\big] \tag{49}$$

Proof: Appendix F.3.

Proposition 3.3 tells us that if there is no unanticipated information arriving in period t, we can recover period-t amenities as a residual which rationalizes the observed migration flows, given the observed rents and previously backed-out wages. The information assumption 3.2 is important in recovering the indirect intra-period utility differences across locations, because from these differences it is simple to back out amenities given wages and rents. This information assumption ensures that in period t, agents get exactly the utility they were expecting to get, implying that the migration flows, which are governed by the utility that households expect to get, are directly informative about amenities.

Amenities are only identified relative to some reference location amenities, as highlighted in equation 48. This is consistent with the idea that in the model, amenities operate as a utility shifter. As usual, we can only recover *differences* in utilities across alternatives. Thus, in reporting the estimated amenities, they are always expressed relative to some reference location, in which their level is normalized.

3.4 Recovering amenities if there are unanticipated changes.

Identification challenge Consider next a situation in which agents might receive news about economic fundamentals in the beginning of period $t+1$. The inversion strategy from proposition 3.3 no longer works because it is no longer the case that migration flows are directly informative about household values: Households might have made optimal migration choices in period t given the information they had at hand in period t, but this is no longer optimal given the new information that arrives in $t+1$.

Can we still recover the new level of amenities period $t+1$, even if amenities or some other fundamentals of the model might have changed? The sequential structure of the economic fundamentals causes an identification issue. Agents could, in principle, receive news about any future fundamentals - they could, for example, learn in period $t+1$ that a specific location will benefit from a productivity shock in period $t+s$, where s could be any period. On the other hand, in order to back out period-$t+1$ beliefs, we can only use data from period-$t+1$, since agents could, in principle, receive again some new information in period $t+2$.

Thus, in order to back out period-$t+1$ beliefs about economic fundamentals, we can only use data from $t+1$, whereas news could arrive about any future period values of exogenous variables, so that there are an infinite number of observationally equivalent news that could arrive.

Restricting agents' expectations about economic fundamentals To overcome the identification challenge, I need to restrict the beliefs that agents might have about the evolution of economic fundamentals. In order to identify what agents anticipate in period $t+1$, I can only use data about values in the $L-1$ locations in period $t+1$, so the object to be recovered cannot have more than $L-1$ dimensions.

I adopt the following convention. Economic agents do not anticipate future changes to economic fundamentals. Instead, they observe the current realizations of $A_{l,t}$, $w_{l,t}$ and $\overline{L}_{l,t}$ and assume them to remain constants. Similarly, I also assume that landlords' discount rates $i_{l,t+s}$ do not depend on s, and thus landlords use $i_{l,t}$ to discount all future revenues from any period [8]. Moreover, agents do not anticipate changes to $i_{l,t}$.

These assumptions imply for example that nobody migrates today to a specific location because they anticipate a productivity increase to take place in that location in some future date. However, economic agents are sophisticated enough that after observing the current realizations of exogenous variables, they do then correctly anticipate the future evolution of endogenous variables (in particular, rents), if no further shocks were to take place[9]. Ahlfeldt et al. (2020) adopts a similar timing convention[10].

This simplified time structure for location-specific economic fundamentals implies that equation 37 simplifies to

$$V_{l,t}(\Theta_t) = \sum_{s=0}^{\infty} \beta^s u(A_{l,t}, w_{l,t}, r_{l,t+s}) - \eta \sum_{s=0}^{\infty} \beta^s \log \mu_{t+s}^{l,l}$$

where $A_{l,t}, w_{l,t}$ denote the period-t realizations of amenities that agents then project into the future.

Note that this restriction on agents' beliefs is adopted in order to recover *identification*. The *solution* to the model can be computed under any arbitrary, convergent sequences of changes to these fundamentals.

To see whether we can recover the identication of amenities, suppose that all other parameters and variables of the model are unknown but the amenities in one location, $A_{l,t}$.

Assumption 3.3. (*Anticipated evolution of economic fundamentals*) *In period t, agents anticipate all location-specific economic fundamentals $(A_{l,t}, w_{l,t}, \overline{L}_{l,t}, i_{l,t})$ to remain constant at their current levels from period t onwards.*

[8]It would be possible to allow for a known term structure of interest rates, too.

[9]Other assumptions about agents' beliefs would also be possible as long as they depend on only one unknown parameter: We could, for example, let agents anticipate constant growth rates in fundamentals for a given amount of years.

[10]Previously, Ahlfeldt et al. (2020) have suggested a model inversion exercise that builds on a similar timing convention but that puts additional structure on agents' choices (they can only move once). In the setup that I am using such restrictions are not necessary.

Proposition 3.4. (*Recovering amenities in a single location*) *Under assumption 3.3, conditional on all other variables of the model, including amenities in all other locations relative to some reference location k, then $A_{l,t}$, the period-t level of amenities in location l can be recovered from period-t migration probabilities as long as $V_{l,t} - V_{k,t}$ is monotone in $A_{l,t}$.*

To see this, write the difference in the values across two locations in period $t+1$ (which can be recovered from the period-t migration probabilities) as

$$V_{l,t+1}(\tilde{\Theta}_{t+1}) - V_{k,t+1}(\tilde{\Theta}_{t+1}) = \sum_{s=1}^{\infty} \beta^s \log\left(\frac{A_{l,t}}{A_{k,t}} \frac{w_{l,t}}{w_{k,t}} \left(\frac{r_{k,t+s}(A_{l,t})}{r_{l,t+s}(A_{l,t})}\right)^{(1-\phi)} \left(\frac{\mu^{k,k}_{t+s}(A_{l,t})}{\mu^{l,l}_{t+s}(A_{l,t})}\right)^{\eta}\right)$$

where the notation $f(A_{l,t})$ is used to highlight that the rents and migration probabilities depend on the value $A_{l,t}$. Since all other exogenous variables are fixed, the left-hand side is observed (from the migration probabilities), and the right-hand side is increasing in $A_{l,t}$ (by assumption), a unique value of $A_{l,t}$ will satisfy this equation.

The intuition of proposition 3.4 is the same as in proposition 3.3: We back out location-specific amenities as the residuals that rationalize the observed migration flows. The difference to proposition 3.3 is that this time we cannot directly back out indirect intra-period utilities from observables: Instead, we need to solve for the sequential equilibrium to compute the values of living in different locations under the new information. Proposition 3.4 suggests that as long as the amenities in location l increase the value of living in l more than the value of living in another location k, then the equation on value differences can be inverted. In other words, as long as the direct effects of amenities dominate any indirect equilibrium effects, then amenity levels can be recovered by inverting the migration probabilities. The inversion needs to be done numerically but the solution is unique.

3.5 Recovering discount rates

Next, suppose that all other variables of the model are known but one of the discout rates. Given assumption 3.3, the discount factor used by the financial sector is given by $\rho_{l,t} = \frac{1}{1+i_{l,t}}$.

Proposition 3.5. (*Recovering the discount rates.*) *Suppose that agents' beliefs are given by assumption 3.3 and that all other parameters and all other discount rates are known exept the discount rate in location l. Suppose that the house price in location l is declining in the discount rate in location l, $\frac{\partial p_{l,t}}{\partial i_{l,t}} < 0$. Then, there is a unique discount rate that $i_{l,t}$ rationalizes the observed price in location l in time t.*

To see this, write the house price as the NPV of future cash flows using equation 12

$$p_{l,t}(i_{l,t}) = \sum_{s=0}^{\infty} (\frac{1}{1+i_{l,t}})^s (1-\delta)^s r^N_{l,t+s}(i_{l,t}) \tag{50}$$

where the notation $r^N(i_{l,t})$ is used to highlight that the net rent, which is an equilibrium outcome, also depends on the unknown interest rate. The left-hand side is observed and the right-hand side depends on a single unknown $i_{l,t}$ in a monotone way, so that a unique $i_{l,t}$ satisfies the equation.

Proposition 3.5 tells us that holding constant interest rates in all other locations as well as all other variables, we can recover the interest rate in l that would rationalize the observed price in location l. Thus, the interest rate serves as the residual allowing to rationalize the remaining variation in observed prices that is not accounted for by the model via expected rental growth, given the other parameters of the model.

The condition $\frac{\partial p_{l,t}}{\partial i_l} < 0$ requires simply that interest rates do not have an unusual effect on prices such that the decline of the interest rate would have a negative effect on prices. While this is not a strong assumption in general, it could be violated for example in a model with very strong congestion forces: if a decline in interest rates in location l would increase prices and increase housing supply, thereby causing in-migration, which would then result in lower willingness to pay for housing in that location due to congestion, it could be possible that the interest rates had an adverse effect on prices. In the housing market model that I present, which has neither congestion nor agglomeration forces, such effects seem unlikely.

3.6 Implementation

Inversion algorithm In practice, to implement the model inversion exercise, we can start by inverting the static market clearing conditions, after which we need to solve a system of $2L-1$ equations in $2L-1$ unknowns simultaneously (interest rates in all locations and amenities in all but one location). This needs to be done while taking into account that the observed migration probabilities and prices are determined in equilibrium and depend on all the $2L-1$ unknowns simultaneously. The uniqueness of this inversion and the practical implementation are discussed in appendix C. I also propose a fast algorithm for implementing the inversion procedure.

Interpretation The discount rates are recovered as the structural residuals that allow rationalizing exactly the observed prices, given the anticipated rental growth implied by the other economic fundamentals in the model. Similarly, the amenities are recovered as structural residuals that rationalize observed migration. Therefore, any factors that are absent from the model but affect migration or prices will be loaded on these structural residuals. For example, if a part of the price divergence between big cities and the rest of the country was due to changes in international migration, which is not modeled, we would interpret this as changes in discount rates. Thus, the amenities and discount rates have a residual interpretation instead of a structural one.

Identification of amenities In typical location models of location choice, amenities operate as utility shifters. However, like in any discrete choice model, the *level* or the *dispersion* of utility cannot be identified from observed choices. Thus, amenities are also only identified up to normalisations *location* and *scale* normalisations. In my application, these concern fixing the level of amenities in some reference location, and setting the utility function so that utils are measured in log incomes. A discussion on these normalisations is provided in appendix F.4.

Even after these normalisations, relative amenity *levels* in quantitative spatial models are not very informative, because they are sensitive to the definitions of spatial units. It is common in urban economics to invert equations predicting utility, given observed rents and wages, to recover underlying amenities, and if two cities offer the same wages and rents, then the more populated city is interpreted to have higher amenities (Diamond, 2016; Albouy, 2008). However, this implies that the implied amenity measures are sensitive to the definition of a location. Consider, for example, the city of New York: We could define the city of New York to consist of *i*) "Manhattan" and "Other New York", or we could define it to consist of *ii*) "Manhattan", "Brooklyn", "Queens", "The Bronx" and "Staten Island". Ideally, we would like it to be the case that the amenities we would estimate for in the first scenario for "Other New York" were a convex combination of the amenities estimated for the Brooklyn, Queens, The Bronx and Staten Island in the second scenario. However, I show in appendix F.4 that this is not the case - instead, the definition of geographic units affects amenity estimates in undesirable ways. For this reason, I will treat amenity levels as uninformative, and focus only on changes in amenity levels over time, holding constant the geographic units.

An additional concern is related to normalising amenity levels in a dynamic model. As highlighed by propositions 3.3 and 3.4, amenity levels are recovered only up to a normalization on the amenity levels in some reference location, since utility levels are not identified. In a dynamic model, however, it is not innocuous to make normalizations in multiple periods (see appendix F.4 for details). Therefore, we cannot identify amenities in a specific location relative to their values in a previous period $\frac{A_{l,t}}{A_{l,t-1}}$, only the differences in amenities relative to some reference location $\frac{A_{l,t}}{A_{l,t-1}} \frac{A_{reference,t-1}}{A_{reference,t}}$. Thus, using estimates from the inversion exercise, we cannot make statements about changes in time in utility levels, only about changes in utility relative to a reference location.

4 Quantification

The empirical application of this paper concerns Finland and uses two types of information: data on the endogenous outcomes of the model and externally calibrated structural parameters. I start this section by describing the institutional context of the empirical exercise and the data sources for the endogenous outcomes of the model. I then discuss the choices of the structural parameters. Finally, I provide some descriptive evidence on regional divergence in Finland, summarizing the evolution of prices, rents, migration and construction across locations.

4.1 Data and institutional context

Quantifying the model requires two types of data. First, in order to implement the solution technique suggested by propositions 2.1 and 2.2, we need data on the initial allocation of the economy: initial distribution of households and housing units across locations, initial prices and rents, and initial construction and migration. Second, I will also use data on the realized equilibrium allocations in 2012-2019 to interpret the realized allocations through the lens of the model.

Institutional context: Geography Finland is a large country in terms of surface, but also very sparsely populated, with 5.5 million inhabitants. The population is strongly concentrated in the south of the country, with more than a million people living in the larger Helsinki area alone. However, there are relatively large cities also in the east and the north, possibly due to regional policies aimed at opening universities throughout the country (see Suhonen & Karhunen 2019). The long distances suggest that migration costs are likely to be important.

Institutional context: Housing markets Homeownership is common in Finland. Homeownership rate is 62% for the total population but in smaller apartments, the rate is significantly lower: 14% in 1-room apartments (excluding kitchen) in blocks of flats and 35% in 2-room apartments[11].

The rental sector consists of an important publicly subsidised sector (public housing and semi-public housing) as well as a private sector (for details, see Eerola & Saarimaa

[11]Statistics Finland, see appendix B.

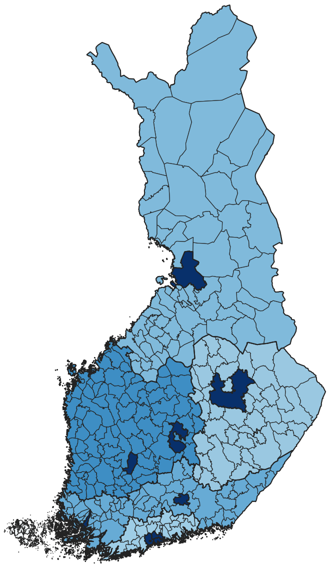


Figure 2: Aggregation of Finnish municipalities to regions for the quantitative exercise.

Notes. The shapefile with Finnish municipal boundaries is obtained from Statistics Finland, see appendix B.

2018). My interest is in the private sector. Among all households living in 2-room apartments (excluding kitchen) in blocks of flats, 40% live in rental units in the private sector, suggesting that the private rental sector is important in this housing market segment[12]. Rents in new rental contracts in the private sector are unregulated.

The rental property is held in part by individual landlords and in part by real estate investment companies. Statistics Finland inofficial estimates suggest that in 2019, in approximately 20% of all housing transactions the purchaser was a private individual who became a landlord (not moving to the apartment themselves), and in almost 10% of the transactions, the purchaser was a real estate investment company[13].

Geographic units In order to implement my empirical exercise in practice, I need to split Finland into geographically distinct regions. I will define the regions as follows: The cities with a population of above 100 000 in 2021 will all constitute separate locations, these cities being Helsinki, Espoo, Tampere, Vantaa, Oulu, Turku, Jyväskylä, Kuopio and Lahti. In 2021, approximately 40% of the total population live in these cities. The rest of the country will be allocated to 5 regions according to the 2021 NUTS 2 classification, excluding the 9 largest cities: Other Northern, Other Eastern, Other Western, Other Southern and Other Uusimaa regions. This aggregation is illustrated in Figure 2.

[12]Statistics Finland, see appendix B.

[13]Otto Kannisto, Martti Korhonen, Anu Rämö, Elina Vuorio, Statistics Finland Tieto & Trendit, "Yli puolet viime vuonna myydyistä yksiöistä meni sijoittajille", published 29.10.2020.

Timing I focus on years 2012-2019, when the regional divergence in house prices has been strong in Finland, as illustrated by Figure 1. I choose the initial year of my analysis to be 2012 in order to ensure that I am not accidentally capturing a housing market recovery from the 2009 financial crisis. I choose the final year to be 2019 in order to exclude the turbulence on the housing market caused by the global pandemic. I refer to year 2012 as the baseline economy.

House prices House prices are measured using microdata on transaction recordings from the Finnish Federation of Real Estate Agency (KVKL). The advantage of using this dataset instead of administrative alternatives is that the agencies collect transaction price information together with a high number of covariates. This dataset also allows me to measure the monthly fees that owners of apartments in blocks of flats must pay to building co-operatives in order to participate in building-level operating costs such as cleaning, heating, et cetera. These operating costs (ξ_l) are therefore treated as observed and they are deducted from rents to compute the net rents that landlords receive. A more detailed description of this data is provided in Appendix A.2.

Rents Rents are measured using rental listings data from the online listings service `vuokraovi.com`. As rental agreements are private transactions, the universe of rental contracts are not registered by administrative sources, leaving researchers with limited alternatives for measuring rents. Listing rents have the downside that they need not perfectly reflect the realized rents. However, they provide significantly better data coverage and covariate quality than available survey sources. Since I use listing rents, and there are no rent controls on new rental contracts in Finland, my rents measure should not suffer from a "sticky rents" measurement problem. A more detailed description of this data is provided in Appendix A.2.

Rent and price indices In my main empirical application, I measure regional apartment price and rents indices using hedonic regressions with strict sample selection in order to ensure that I measure rents and prices of comparable apartments (instead of comparing, say, detached home prices to studio rents). For both the rent and price regressions I only use apartments in blocks of flats, with the number of rooms equal to two (typically, a living room and a bedroom), in good condition, and excluding new buildings. In running the hedonic regressions, I control among other things for apartment floor area, building age and zipcode fixed effects. The details are provided in appendix A.2. Appendix A.3 contains further descriptive evidence on the regional divergence of rents and prices in Finland. As suggested by other figures in appendix

A.3, the measure that I use in my main specification is well in line with other possible measures of regional house price divergence.

I also need to measure rents, prices and operating costs in levels for the baseline year. For this purpose, I keep the same sample as for the main hedonic regressions, and compute simply averages of rents, prices and operating costs per square meter by year and by location.

Other data To measure population at the municipality level, I use Statistics Finland Vital Statistics. To measure migration from one region to another, I use Statistics Finland matrices on intermunicipal migration. To measure new housing construction, I use administrative data from the Finnish Population Information System, publicly available via the Liiteri service. To measure the initial housing stock in each location, I use Statisctics Finland household-dwelling statistics which is based on administrative data. Finally, to describe regional incomes, I use Statistics Finland Income Distribution statistics, and to describe the age structure and the number of workplaces in the service occupations by location, I use Statistics Finland Municipal Key Figures database. For a full description of these data sources as well as the full list of references, see appendix B.

4.2 Structural parameters

For the empirical exercise, we also need information on the structural parameters of the model $(\beta, \eta, \phi, \delta, \gamma, \Upsilon, \xi_l)$. This section provides a discussion for how these values are chosen.

Discount rates I think of a time period as approximately one year. I fix the discount factor $\beta = 0.95$ accordingly.

Consumption shares I set the consumption share of housing at 0.3 and thereby non-housing consumption share at $\phi = 0.7$. This is consistent with Statistics Finland Household consumption expenditure survey (see appendix B).

Depreciation The depreciation of the housing stock is set at $\delta = 0.015$. Typical estimates for housing depreciation, if there is no maintenance, are between 1% per annum (Wilhelmsson (2008)) and 2.5% per annum (Harding et al. (2007)). I choose a value from the lower end of this range, firstly since I measure prices of 2-room units in blocks of flats, where I expect depreciation to be lower than in detached homes, and secondly because to explain prices, I measure rents net of operating costs, which do cover some building-level maintenance (see section 4.1).

Rental revenue tax The tax rate on rental revenue is set at $\Upsilon = 0.3$ in order to reflect the capital gains taxation in Finland. Since 2012, the capital gains tax rate in Finland has been is 30% on incomes below 30 000 euros per annum.

Operating costs The operating costs of landlords by location, ξ_l, are observed (see appendix A.2). They range approximately from 20% to 30%.

Migration elasticity The (inverse) migration elasticity parameter η governs the importance of the idiosyncratic shocks for migration decisions relative to the systemic component in utility, and thereby governs the magnitude and speed of migration responses to location-specific shocks. All else equal, η should be higher if the model time periods are shorter. Caliendo et al. (2019) use a quarterly value for η of 5.3, consistent with an annual elasticity of 2, and Kleinman et al. (2021) use a 5-year elasticity of 2.3, both from the USA. Ahlfeldt et al. (2020) use an annual value of 3.4 for Germany. As these are the only benchmarks in the literature, I pick a number close to the German estimate and set $\eta = 3$ to reflect that Europeans might be on average slightly less responsive to migration incentives than U.S. residents. See Ahlfeldt et al. (2020) for a discussion on why the state-of-art estimation of the migration elasticity parameter is weak.

Construction technology The parameter of the housing production function, γ tells what is the technology via which land and labor can be combined to produce housing. In the model, all else equal (holding fixed the location-specific land supply and wages), γ also determines the price elasticity of housing supply

$$\frac{\partial Q_{l,t}}{\partial p^Q_{l,t}} \frac{p^Q_{l,t}}{Q_{l,t}} = \frac{\gamma}{1-\gamma}$$

I set γ to be 0.5 such that the housing supply elasticity following this choice is 1 to be consistent with a recent OECD estimate for the long-run housing supply elasticity in Finland (Cavalleri et al., 2019)[14]. An alternative to calibrating this parameter externally would be to estimate the parameter as in Ahlfeldt et al. (2020).

The selected parameter values and their sources are summarized in table 1.

[14]The OECD number is slightly higher than estimates from Oikarinen et al. (2015) for the largest Finnish cities, which could be used as an alternative parameterization for γ.

Parameter	Symbol	Value	Source
Discount rate	β	0.95	
Migration elasticity	η	3	Caliendo et al. (2019), Ahlfeldt et al. (2020)
Housing depreciation rate	δ	0.015	Harding et al. (2007)
Housing consumption share	$1-\phi$	0.3	Statistics Finland
Housing construction technology	γ	0.5	Cavalleri et al. (2019)

Table 1: A summary of the exogenously set structural parameter values.

4.3 Descriptive evidence on regional divergence in Finland

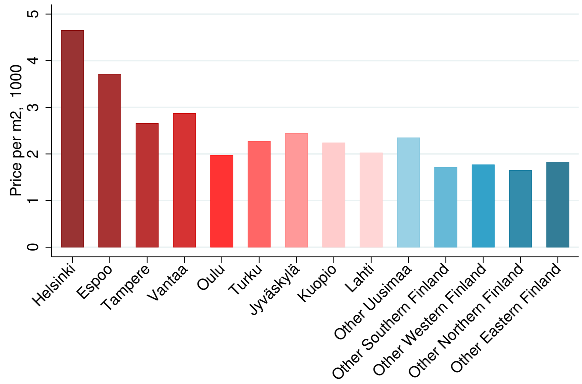
In this section, I provide descriptive evidence on regional divergence in Finland.

I start by characterizing the data summarizing the "initial allocation" of the economy as ment by proposition 2.1. Figure 3 summarizes the levels of rents, prices, number of households and housing stocks in the initial allocation, set to year 2012. The capital Helsinki is the largest among the 9 large cities (denoted in red). However, the 5 regional groups that pool together the regions outside the 9 cities (denoted in blue) are also large in terms of population. Helsinki is also the most expensive, both in terms of prices and rents.

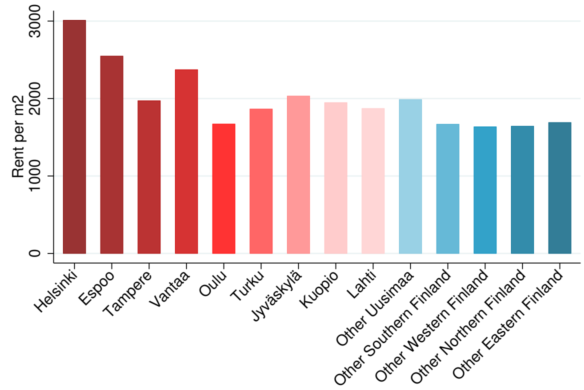
Figure 4 summarizes the evolution of prices and rents across regions, relative to their 2012 levels, in real terms. Price growth has been clearly strongest in the big cities: the capital Helsinki, the two other large southern cities Tampere and Turku, as well as the two cities neighboring Helsinki (Espoo and Vantaa). In contrast, prices are stagnating or even declining outside the 9 largest cities. For rents, there is less divergence in general and moreover, the ranking of cities by price growth and rental growth differs importantly. For example, rents in Helsinki have stayed virtually at their 2012 levels, but real prices have increased by more than 20%. Jointly, these findings imply that price-to-rent ratios have increased in the large cities and declined outside the large cities, as described in the bottom panels of figure 4. A more detailed analysis of the price and rent divergence is provided in appendix A.3.

Figure 5 provides descriptive evidence on regional incomes, construction and migration. The top-left panel depicts the evolution of wages across Finnish regions as seen in the data. The regional wage divergence is very modest[15]. The top-right panel plots the evolution of new construction by location. New construction has been more

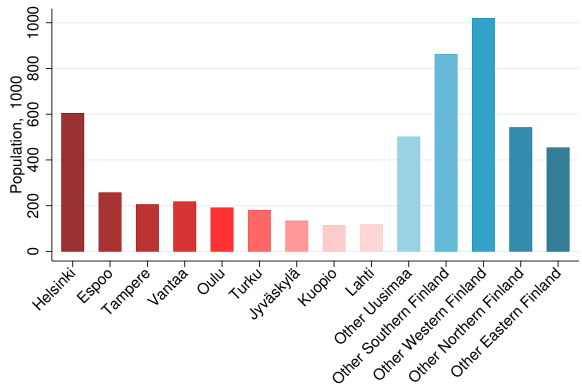
[15]While existing literature does suggest that small differences in regional incomes could translate to important price dispersion (Van Nieuwerburgh & Weill (2010)), there does not seem to be a systematic pattern in wage changes - wages increase in "other Finland" by more than in many of the big cities. Moreover, increased wages can be consistent with higher prices but they should also have similar effects on rents.



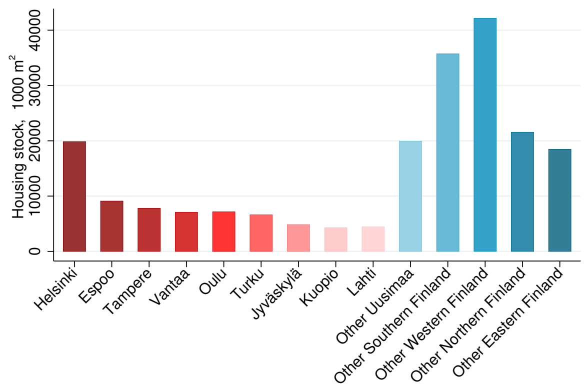
(**a**) Mean transaction prices, 2012



(**b**) Mean rents, 2012



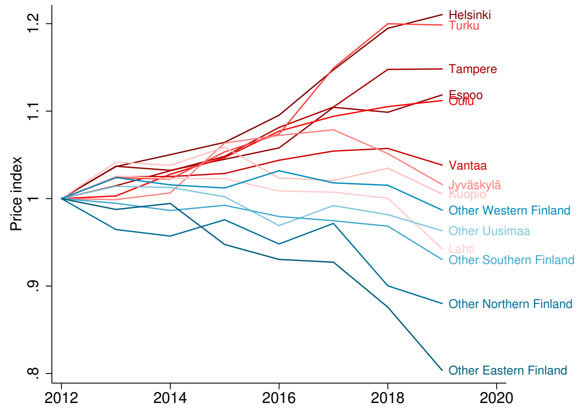
(**c**) Population, 2012



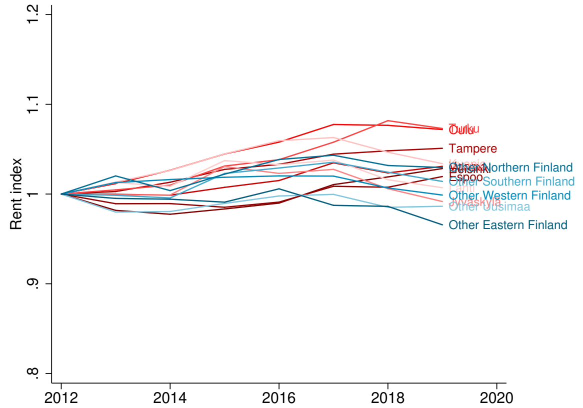
(**d**) Housing stock, 2012

Figure 3: Initial allocation of the economy in 2012.

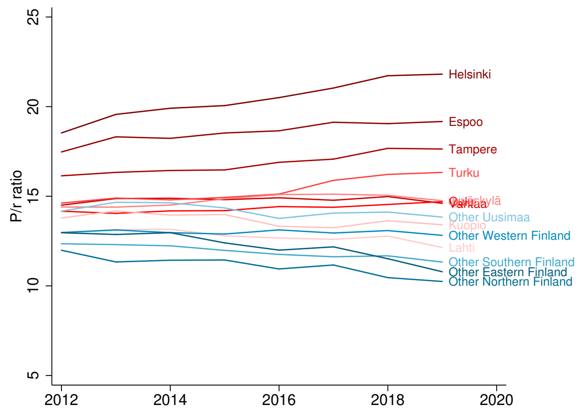
Notes. Mean transaction prices refer to mean prices in 2-room apartments in blocks of flats in good condition, in transactions registered by the Finnish Federation of Real Estate Agency (KVKL). Mean rents refer to mean rents in 2-room apartments in blocks of flats in good condition, in rental listings on the website vuokraovi.com. Population and housing stock are measured using Statistics Finland statistics. For details on the data, see section 4.1.



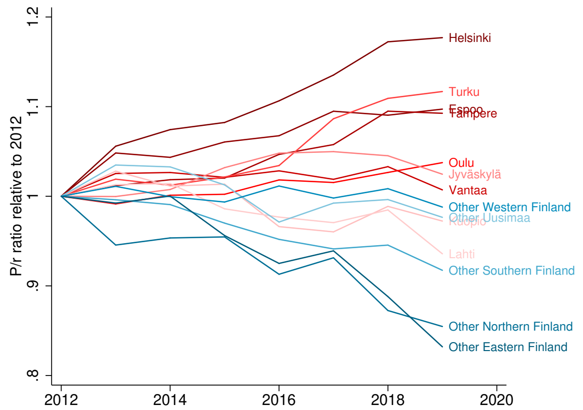
(**a**) Hedonic price index



(**b**) Hedonic rent index



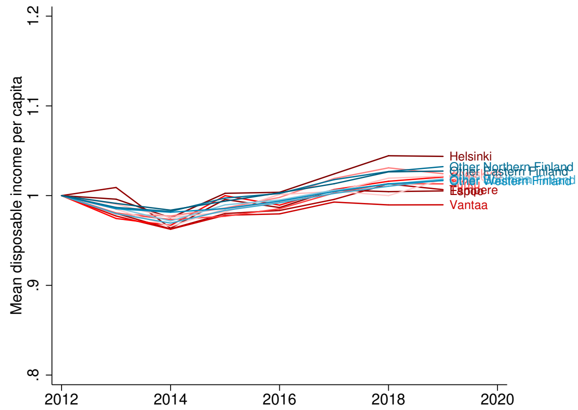
(**c**) Price-to-rent ratios



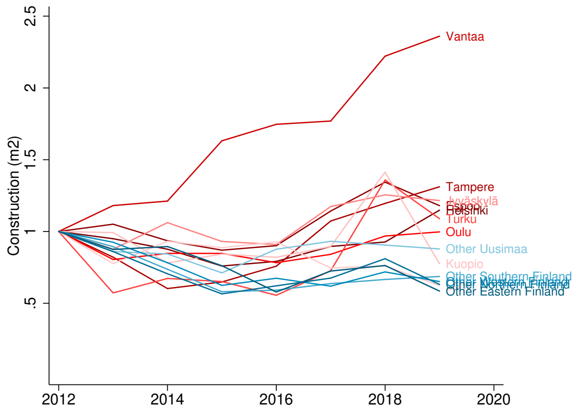
(**d**) Price-to-rent ratios relative to 2012

Figure 4: Hedonic indices for apartment rents and prices in 2-room apartments in blocks of flats (top panels) and the implied price-to-rent ratio as well as the price-to-rent ratio change relative to 2012.

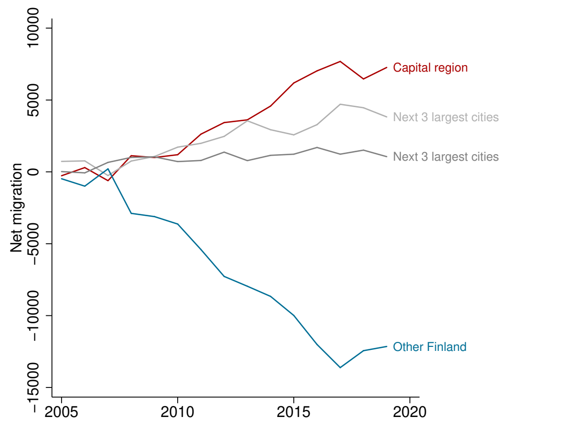
Notes. House prices are measured with a hedonic regression using Finnish Real Estate Agencies transaction dataset (KVKL HSP). Rents are measured with a hedonic regression using listings data from website vuokraovi.com. Both samples are restricted to apartments in multi-family blocks of flats, number of rooms equal to two (typically, a living room and a bedroom), in good condition, and excluding new buildings. The price regression includes controls floor area, floor area squared, age, age squared, floor number, maintenance charge (building operating costs) and a dummy characterizing land lot ownership status as well as zipcode fixed effects. The rent regression controls for floor area, floor area squared, age, age squared, an indicator for whether the apartment is immediately available, an indicator for the owner type (individual or a company), and zipcode fixed effects. Price-to-rent ratios in 2012 are measured by the ratio of mean prices to mean rents, and for 2013-2019 by using the price and rent changes implied by the hedonic indices relative to 2012 means, as described in appendix A.2. Rents and prices are deflated to 2020 euros.



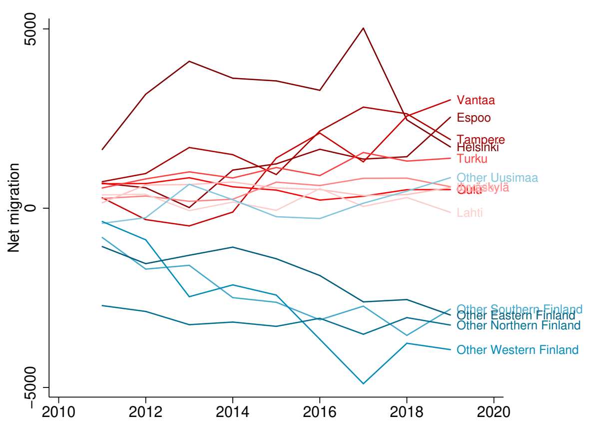
(**a**) Real per-capita incomes



(**b**) Construction



(**c**) Net migration 2005-2019



(**d**) Net migration, 2011-2019

Figure 5: Descriptive evidence on regional incomes, construction, and net internal migration across Finnish municipalities.

Notes. Incomes are measured as the mean per-capita incomes by municipality from the Statistics Finland Income Distribution Statistics. They are measured in 2020 euros. New construction is measured as the total new construction in m^2, provided by the Finnish Population Information System via the Liiteri service (Syke). Migration is measured using the Statistics Finland matrices on intermunicipal migration and population. Panel ... reports net migration from 2005 to 2019 across Finnish municipalities grouped into 4. In panel ... , the red line comprises of the capital region (Espoo, Helsinki, Vantaa), light gray line of the 3 next largest cities (Tampere, Turku, Oulu), dark gray line of the next 3 largest cities (Jyväskylä, Kuopio, Lahti), and finally the blue line comprises of all other Finland, corresponding to Other Northern, Other Eastern, Other Western, Other Southern and Other Uusimaa in my regional classifications. Panel B provides net migration numbers for years 2011-2019 using the same regional classification as the main analysis. Cross-municipality migration is measured with Statistics Finland Intermunicipal migration matrices.

important in particular in some of the larger cities. Finally, the bottom panels in figure 5 depict patterns of internal migration across Finnish regions. The left panel plots net migration across municipalities grouped to 4 larger groups. This figure suggests that while around 2005, net migration between the largest cities and elsewhere was close to zero, these differences have rocketed during 2010's. Net annual migration away from "Other Finland" into the 9 largest cities has gone from approximately 0 to more than 10 000 individuals annually. The bottom-right panel splits this up to the regional classification used in the analysis. Jointly, these graphs depict important changes over time in migration patterns. This is the variation that is going to be identifying differences in values of living in different locations.

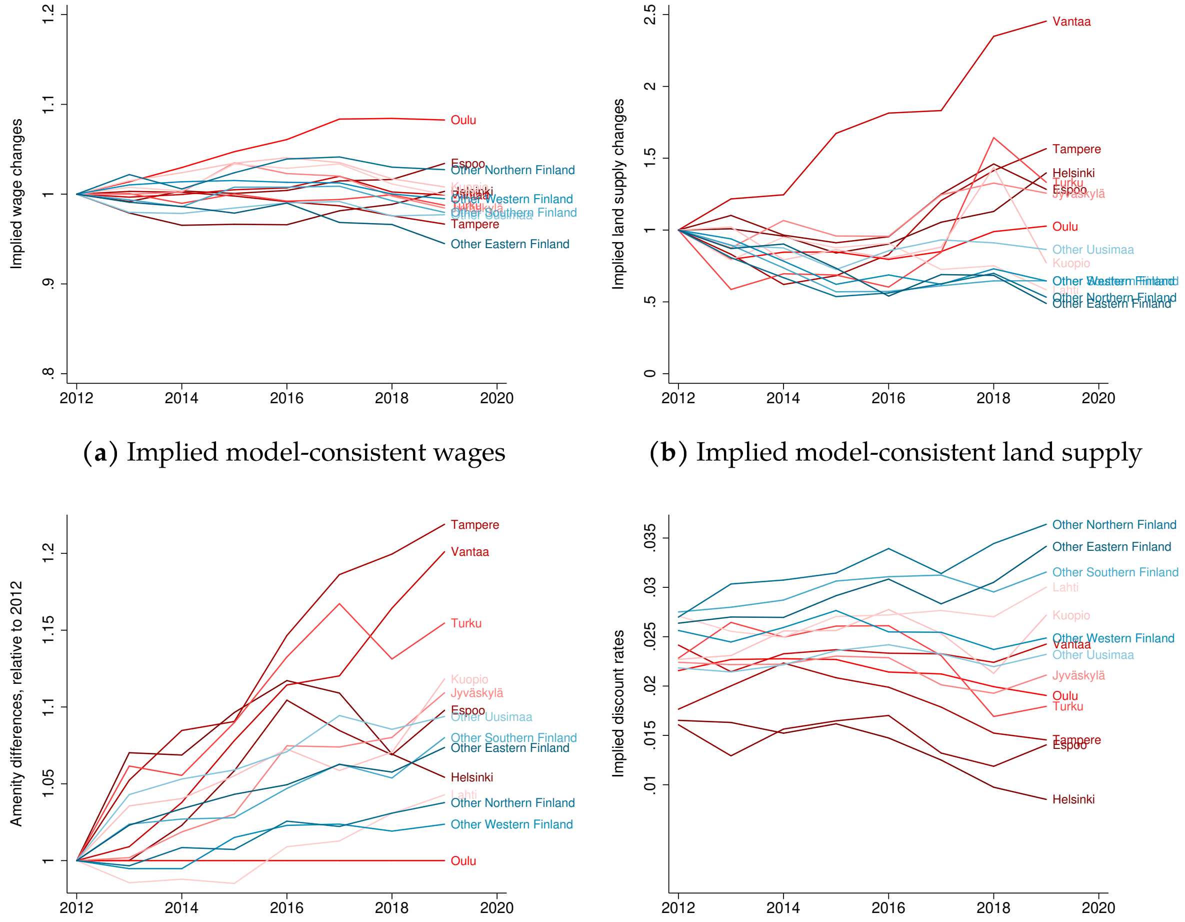
5 Results from model inversion

This section reports results from the model inversion exercise. The model inversion, described in section 3, allows me to recover changes in economic fundamentals (wages, amenities, land supply constraints, discount rates) that would rationalize the oberved data on prices, rents, migration and cosntruction in Finland. I first document the implied economic fundamentals. Second, to assess the relative importance of these different fundamentals in explaining price-to-rent ratio divergence, I report counterfactual price-to-rent ratios which are obtained by assuming that the economy would have evolved otherwise as estimated, but shutting down the underlying changes one at a time. Third, I study how the implied economic fundamentals are associated with obervable regional characteristics.

5.1 Implied fundamentals

The top-left panel in Figure 6 reports the the model-consistent wages, backed out from information on rents as well as housing consumption. Consistent with the real evolution of mean wages by region, depicted in Figure 5, the model-implied wages also diverge at most slightly across regions. However, the inversion exercise does slightly overestimate wage dispersion, in particular by overestimating wages in cities like Oulu. This is because for such cities, rents have grown faster than incomes while the quantities of housing consumption have not changed, and the model inversion exercise interprets this as increased wages.

The top-right panel in figure 6 reports the land supply quantities implied by the inversion. This figure effectively mirrors the construction quantities. For example, in the city of Vantaa, construction increases more than twofold in the time period, and the



(**a**) Implied model-consistent wages

(**b**) Implied model-consistent land supply

(**c**) Implied model-consistent amenities

(**d**) Implied model-consistent discount rates

Figure 6: Implied economic fundamentals across locations, obtained from the model inversion.

Notes. Wages and land supply are backed out from the static market clearing equations. They are reported relative to their 2012 levels. Amenities and discount rates are backed out using the numerical procedure described in appendix C. Amenities are reported as the relative change in amenity differences between location l and the reference location (Oulu, where amenities are normalized to 1 each period), relative to the initial differences. Discount rates are reported in levels.

model interprets this as a twofold increase in land availability. This is consistent with the modest changes in wages and relatively modest changes in house prices across regions. For example, in the city of Vantaa, house prices and wages barely changed from 2012 to 2019, so the model loads the over 200% change in construction on the land supply variable.

Next, the lower left panel in Figure 6 reports the amenities obtained by the inversion procedure suggested in propositions 3.3 and 3.4. As explained in section 3, amenities cannot be identified in levels - instead, they are identified relative to amenities in some reference location in that same time period. I fix the reference location to be the city of Oulu, which is the location where the amenity growth appears to be the lowest in my sample period, and amenities in this reference location are normalized to 1 every period. If, for example, amenities in some other location stay constant relative to the reference location, this implies that they have been growing or decreasing in the same way as amenities in the reference location[16]. Moreover, as discussed in section 3, amenity levels are not very informative as they are sensitive to the definition of the geographic units. Therefore, in the figure, I plot amenity differences between each location and the reference location Oulu, relative to the same difference in 2012.

Figure 6 highlights that amenities in big cities have been growing faster than amenities in "Other Finland". The value 1.2 for the city of Vantaa in 2019 has the interpretation that amenity differences between Oulu and Vantaa are 20% larger in 2019 than in 2012:

$$\frac{\mathrm{A}_{Vantaa,\,2019}}{\mathrm{A}_{Oulu,2019}} = 1.2\,\frac{\mathrm{A}_{Vantaa,2012}}{\mathrm{A}_{Oulu,\,2012}}$$

This is a significant increase - the implied increase in the difference in utilities is in the same order of magnitude as if Vantaa would have seen a 20% wage increase while wages would have not grown at all in Oulu.

The observed fast amenity growth in the big cities is consistent with the migration patterns in the data. As depicted by Figure 5, net migration from Other Finland to the big cities has been increasing in 2012-2019, while wages nor rents did change much in either groups of locations. The model therefore interprets this change in migration as relative amenity improvement in the cities. Similarly, the observed amenity growth for the city of Helsinki is consistent with the positive net migration to Helsinki: For years 2012-2016, amenity growth in the city of Helsinki is in the same order of magnitude as for other large cities. However, in the last years of the sample period, amenities in Helsinki lag behind amenities in other large cities - consistent with the drop in the net in-migration to Helsinki depicted in figure 5.

The lower right panel in Figure 6 reports the implied interest rates used for discounting that would be consistent with thee observed prices and the future evolution of rents, as predicted by the model. This figure suggests an important regional divergence of discount rates. The anticipated divergence of rents, implied by today's migration and construction flows, is not sufficient to account for the full regional divergence of prices in the data, and the reminder is loaded on these discount rates as a residual.

The implied discount rates in Helsinki and other large cities have been declining, which is consistent with the decline in the risk free rates over the sample period (12-month Euribor rates have declined from $\approx$ 1.5% in January 2012 to $\approx$ -0.25 % in December 2019). However, the implied discount rates in "Other Finland" do not reflect

[16] As the amenity levels are not identified, we cannot distinguish between a scenario in which amenities in the reference location change or not. It could be, for example, that amenities in the reference location Oulu would have been declining over the sample period 2012-2019.

this decline - to the contrary, they display an upward trend. These residuals could reflect multiple factors: It could be, for example, that the increasing discount rates reflect growing required returns to housing investment in these locations, or that there are regional differences in housing depreciation. Finally, as these discount rates are recovered as residuals, they might also capture differences in rental growth expectations that are not correctly reflected in the structural model (it could be, for example, that the model does not correctly estimate the extent to which rents are anticipated to decline in Other Finland due to reasons that do not affect migration and construction, such as higher population aging).

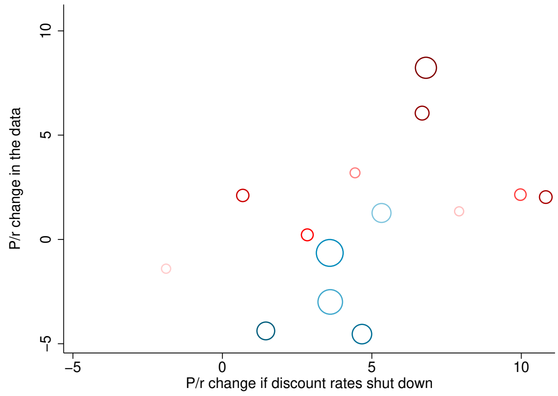
5.2 Relative importance of different channels

As suggested by the model inversion exercise, there have been two important drivers of the regional price and rent evolution in Finland: *i*) a regional divergence of amenities, suggesting that future rental growth rates will differ across locations, and *ii*) a regional divergence of discount rates. I next wish to assess how important are each of these two channels for understanding the regional divergence of price-to-rent ratios in Finland.

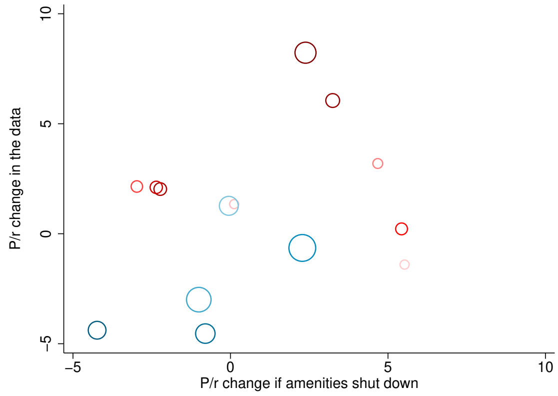
In order to quantify these roles separately for the price divergence, I conduct two experiments. First, I take the baseline equilibrium in 2012 and assume that all other exogenous variables would have evolved as suggested by my inversion exercise, but no amenity changes took place. Second, I conduct a similar exercise, letting all other changes take place as suggested by model inversion exercise, but shutting down the changes in interest rates. The price-to-rent ratios obtained from this exercise are portrayed in Figure 7.

The upper panels of Figure 7 describe price-to-rent ratios in 2015 relative to 2012, the left panel shutting down discount rate changes and the right panel shutting down amenity changes. These plots suggest that both amenities and interest rates have played important roles in explaining the regional divergence of price-to-rent ratios in the data in the early sample period. Shutting down either channel will miss out on an important part of the price-to-rent variation. The correlation between price-to-rent changes in the data and the model is 0.45 if discount rate changes are shut down, and 0.36 if amenity changes are shut down, suggesting that amenity changes are slightly more important than discount rate changes in the first half of the sample period.

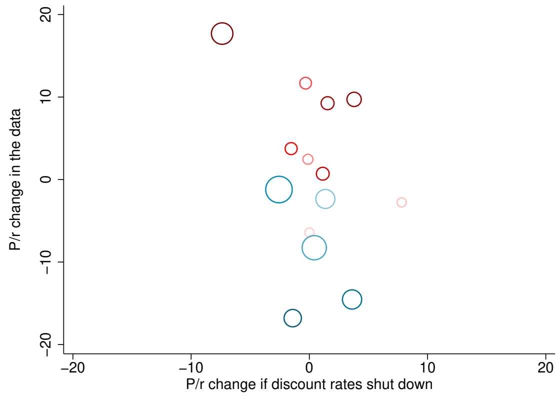
On the other hand, the lower panel plots respective figures but in 2019 relative to 2012. These figures indicate that the story is different for the latter half of the sample. Shutting down amenity changes will not affect price-to-rent ratios by much, suggesting that amenity changes are not an important driver of price-to-rent ratio changes in the



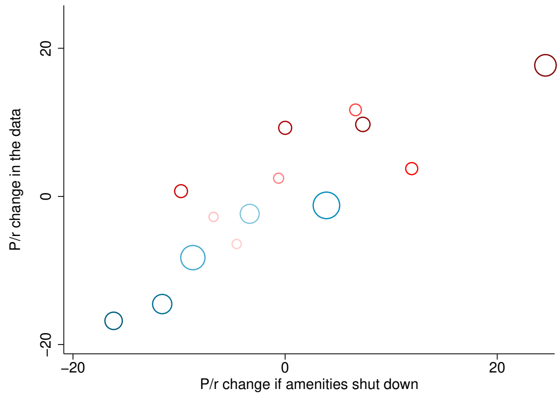
(a) P/r ratio changes, 2015 relative to 2012, assuming no discount rate changes.



(b) P/r ratio changes, 2015 relative to 2012, assuming no amenity changes.



(c) P/r ratio changes, 2019 relative to 2012, assuming no discount rate changes.



(d) P/r ratio changes, 2019 relative to 2012, assuming no discount rate changes.

Figure 7: Price-to-rent ratio changes in the data as well as the implied changes from the model, shutting down amenity changes (upper panel) or shutting down interest rate changes (lower panel).

Notes. To compute price-to-rent ratios implied by the model when amenity changes or interest rate changes are shut down, the equilibrium is recomputed such that all other variables evolve as implied by the model inversion exercise. The change refers to the percentage change in the ratio relative to 2012. The sizes of the circles are proportional to the population in each location.

latter half of the sample period. Instead, if amenity changes are shut down instead of the interest rates, the correlation between price-to-rent changes form the model and in the data is 0.85. This suggests that for the latter half of the sample period, the majority of the price-to-rent variation is loaded on the discount rate changes. These observations are consistent with the parameter estimates depicted in Figure 6. This figure suggests that there were no very dramatic changes in the implied discount rates in the first half of the sample period and that the importance of discount rate changes becomes more important after 2015-2016.

5.3 Interpreting the amenity and discount rate residuals

Since the amenities and discount rates are recovered as residuals, they do not have a structural interpretation. To hypothetise on what these residuals capture, we can examine whether they are associated with regional observables. Figure 8 plots the associations of the change in the discount rates and the change in amenities from 2012 to 2019 with the simultaneous change in population age structure and labor market growth.

The upper left panel shows that the change in the discount rates implied by the model inversion is strongly associated with population aging across locations. The Finnish population is aging quickly, and this affects "Other Finland" relatively more than the large cities with younger populations. The correlation between the change in the model-implied discount rate and and the change in the share of population aged 64 and over is high at 0.73. This suggests that in regions where population is aging faster, discount rates have not been declining as they have elsewhere. In these regions, housing seems to have become less valuable than what it was before, not only because rental growth expectations implied by the model are low, but also for other reasons, captured by the discount rate. One potential such channel is that perhaps in aging locations, new rental tenants also become harder to find, and this is not captured by the model since search frictions are not explicitly modeled. Another potential channel is through regional differences in housing depreciation if, for example, the housing stock in aging locations is more often left deserted because rental revenue no longer covers some fixed costs of operating the rental unit. Finally, it also seems likely that the model is missing out on a part of the rent divergence, since differences across locations in natural population growth are not explicitly modeled. This could be an important extension for future work.

The top and bottom right panels indicate that the implied discount rate and amenity changes are also associted with labor market dynamics, as measured by the change in the number of jobs in service occupations. The model captures labor market dynamics only though mean wages, whereas also job creation has been faster in large cities. If this is then causing some in-migration to larger cities, it will be captured by the parameter describing amenities. Consistent with this hypothesis, the residual amenities are positively associated with the growth of the labor market in each municipality (correlation 0.5).

6 Welfare analysis and counterfactual policy evaluation

In this section, I discuss the welfare consequences of the changes in economic conditions implied by in the model inversion exercise in section 5.1. After that, I analyse the consequences of a counterfactual policy experiment which causes an income shock in the capital region.

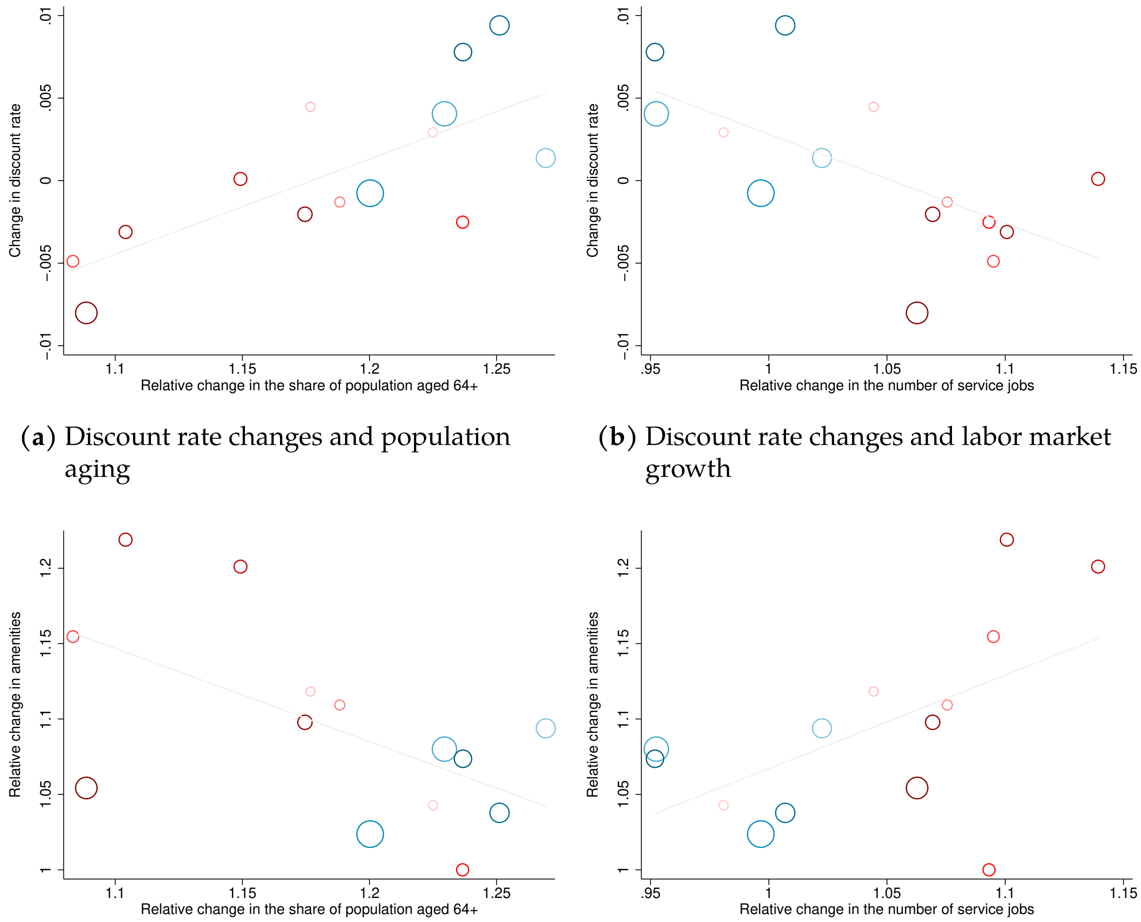
6.1 Welfare implications of Finnish divergence

To adress the welfare implications of the changes in economic fundamentals that are backed out to rationalize Finnish data in section 5.1, I start by considering landlord welfare. Since landlord welfare is simply given by the value of their housing wealth, it suffices to compare the observed prices to the prices that *would have occured* had no changes to economic fundamentals taken place, which is computed using the model. Figure 9 reports these changes, measured as

$$\frac{p'_{l,2019} - p_{l,2019}}{p_{l,2019}}$$

where p refers to the price in 2019 the baseline equilibrium if no changes in economic fundamentals had taken place after 2012, and p' to the price in 2019 in the counterfactual equilibrium in which there were annual changes in fundamentals between 2012-2019. Absent any changes in economic fundamentals, prices would have grown very modestly (by less than 2% over the seven year period) in all regions. However, this is not true for the observed allocation. In the observed allocation, important regional redistribution of wealth has occurred, with price changes ranging from -20% to +20%. The changes in economic fundamentals that have taken place have been very beneficial to homeowners in large cities and harmful for homeowners outside the 9 largest cities.

Next, I consider changes in renter welfare. To see the importance of this exercise, notice that rent levels alone are not informative about renter welfare. First, renter utility is affected by the rent level in their current location, but also by rents in other locations in



(**a**) Discount rate changes and population aging

(**b**) Discount rate changes and labor market growth

(**c**) Amenity changes and population aging

(**d**) Amenity changes and labor market growth

Figure 8: The association of implied amenity and discount rate changes 2012-2019 with the change in age structure and the change in the number of jobs.

Notes. Interest rate changes are measured in levels, as $i_{l,2019} - i_{l,2012}$. Amenity changes are measured in relative changes relative to the reference location d, as $(A_{l,2019}/A_{l,2012})/(A_{d,2019}/A_{d,2012})$. The change in the share of population aged 64 and over is measured as the relative change, $\text{share}_{l,2019}/\text{share}_{l,2012}$. The change in the number of jobs in service sectors in the municipality is measured as the relative change $\text{jobs}_{l,2019}/\text{jobs}_{l,2012}$. Data on the age structure and the number of workplaces in the service occupations by location is from Statistics Finland Municipal Key Figures database. The sizes of the circles are proportional to the population in each location. The light grey line indicates the line of unweighted linear fit.

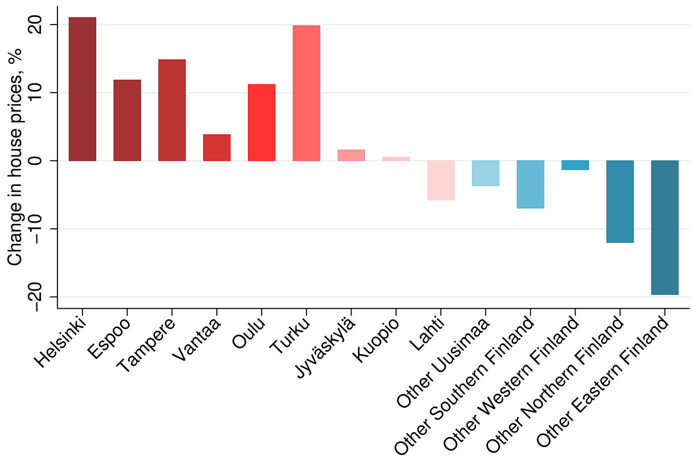


Figure 9: Prices observed in the data relative to prices if no changes in economic fundamentals had taken place after 2012.

Notes. Price changes are measured as $(p'_{l,2019} - p_{l,2019})/p_{l,2019}$, where p refers to the price in the baseline equilibrium if no changes in economic fundamentals had taken place after 2012, and p' to the counterfactual equilibrium in which there were annual changes between 2012-2019.

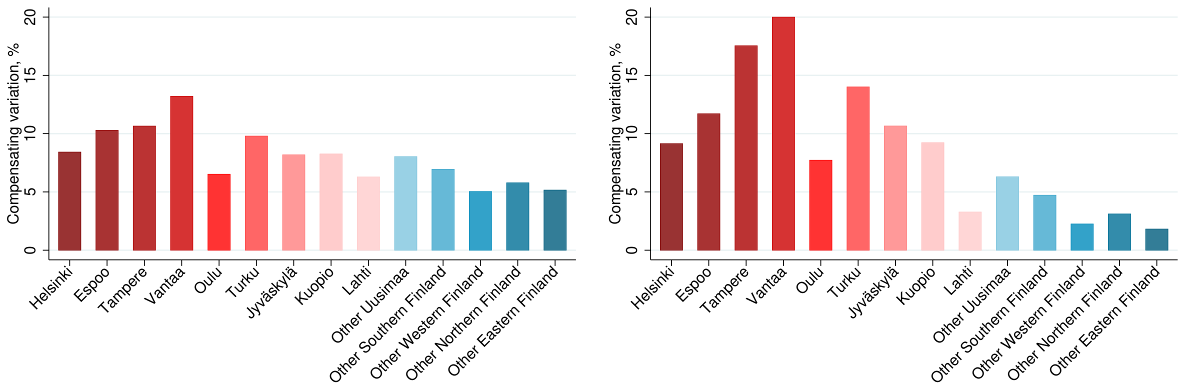
the future (because they might migrate to these locations in the future, which affects their lifetime utility). Second, while renter utility is decreasing in rents, an increase in rents could happen for a reason that has been increasing renter utility or decreasing renter utility. For example, an increase in amenities in a given city or a decrease in land supply in a given city would both push up rents, but the first one would increase renter utility whereas the second one would decrease renter utility. In other words, while the regional price divergence directly suggests that those who own housing in Helsinki have benefitted and those who own housing in other Finland have suffered, rent changes are not directly informative about renter utility across locations, and therefore we can derive additional information from the parameters obtained from the model inversion.

As discussed in section 3, we cannot meaningfully compare renter utilities over time because utility levels are not identified. Figure 10 reports estimates for compensating variation that are obtained after normalizing amenity levels in the reference location in all periods. Thus the numbers can be interpreted as the income variation that would make households indifferent between the world in 2019 and the hypothetical world in 2019 in which no changes to economic fundamentals had taken place, *if* amenities in the reference location would have remained at their 2012 levels also in the counterfactual equilibrium. If, for example, amenities in the reference location would have been declining over time in reality, then we would need to shift downwards our estimate for the compensating variation in every location.

The left panel of Figure 10 suggests that given the chosen normalization, renters have benefited from the changes in 2012-2019 across all locations. This is consistent with how amenity growth was faster elsewhere than in the reference location, as suggested by Figure 6. The compensating variation value of 8 % for Helsinki suggests that households in Helsinki would be indifferent between the world in which changes in economic fundamentals took place between 2012-2019, or an equilibrium in which fundamentals and allocations would have remained as in the initial equilibrium but they received a 8% permanent income increase. There are, however, some differences in the compensating variation across locations. The utility changes are the largest for locations that saw the fastest amenity growth. Overall, the figure suggests that changes in fundamentals 2012-2019 have contributed to important regional welfare differences for renters: Compensating variation across locations varies between 5 and 13 %.

Renters can benefit from location-specific changes through direct channels (if, for example, amenities in their current location increase) or through indirect equilibrium channels. There are two types of indirect channels. First, if there is a positive shock in some location, then there will be out-migration from other locations to that location, which is going to relax housing market congestion and lower rents elsewhere. Second, households in other locations might migrate in the future, so improvements in other locations affect their lifetime utility through the migration option value.

To decompose the welfare increase to the part that is due to declining rents and the part that is due to migration option value, the right panel in Figure 10 reports compensating variation when the option value of migration is shut down as in equation 41. The comparison of the two panels reveals that migration operates as an important buffer which reduces the regional welfare differences stemming from location-specific shocks. When the option value to migrate is shut down. In this scenario, differences across locations are much stronger: compensating variation varies between 2 and 20%. This suggests that migration the option to migrate operates as an important channel to mitigate welfare differences across locations. This also indicates that for households for whom migration is in practice impossible, regional welfare differences can be much larger than what is implied by the baseline estimates.



(**a**) Compensating variation

(**b**) Compensating variation without option to migrate

Figure 10: Compensating variation that would make renting households indifferent between the baseline equilibrium and the new equilibrium, if nothing had happened to amenities in the reference location.

Notes. Compensating variation, as in 39, is measured as the permanent proportional increase in income that would make renting households indifferent between the baseline equilibrium in 2019 (if no changes to fundamentals had taken place after 2012) and the equilibrium in 2019 after changes to economic fundamentals. Compensating variation without option to migrate is measured as in equation 41 by shutting down the effect that migration has on household values. Note that the *levels* of the compensating variation are not identified, but obtained after assuming no amenity changes in the reference location Oulu.

6.2 Counterfactual: Location-specific income changes

As a counterfactual policy, I study the effects that location-specific income shocks would have on the Finnish housing market.

A policy which could lead to location-specific income changes is the proposed undersea tunnel between Helsinki and Tallinn. The proposed tunnel would connect the Finnish and Estonian capitals by train. The tunnel would have a submarine length of approximately 50-80 kilometers, making it slightly longer than the submarine section of the Eurostar tunnel connecting London and Paris. The local authorities of both cities as well as both national govenrnments have been involved in investigating the possibility of the tunnel[17]. Moreover, there have been attempts by private businesses to collect funding for the tunnel[18].

The tunnel would integrate the labor markets of Helsinki and Tallinn. The labor markets already show some level of integration, in particular with a high number of Estonians working in Southern Finland, as average wages are higher on the Finnish side of the gulf. The current mode of transport across the gulf with a ferry takes approximately two hours whereas the proposed train would take 30 minutess or less. On top of connecting the two labor markets, the tunnel would also connect Southern Finland to the European Union via a land connection. As illustrated in Figure 11, the current land connection between Finland and the rest of the Schengen area is through the North of Finland, which is a long route: For example, the distance from Helsinki to Stockholm almost 2 000 kilometers by roads.

In a counterfactual policy analysis, I seek to understand how the tunnel, if a decision on the construction was taken, would affect the housing market in Finland. The dynamic spatial model with housing is particularly suitable for such analysis for two reasons. First, infrastructure improvements usually take a a significant amount of time to complete. Yet, house prices can react already to the announcement of changes, years ahead of the actual improvements (Yiu & Wong, 2005). The model can capture these information effects on house prices even if nothing happens to rents in the short run. Second, the railroad tunnel would be such a significant change in infrastucture that it could potentially have profound effects for example on the allocation of households across Finnish regions. It could, therefore, affect housing markets not only in Helsinki but also in other locations.

[17] See, for example, Helsingin sanomat 16.2.2020 "Tähän tulee tunneli (tai sitten ei)".

[18] A private business "Finest Bay Area Development" announced in 2019 to have secured 15 billion euros in funding for the project. See Talouselämä 8.3.2019 "Peter Vesterbackan Tallinna-tunnelille 15 miljardin rahoitus – Takana kiinalainen sijoitusyhtiö".

Figure 11: Finland and Estonia and the capital cities, Helsinki and Tallinn.

Notes. Finland and Estonia as well as other European Union member countries denoted in green. Shapefile of country borders from Natural Earth.

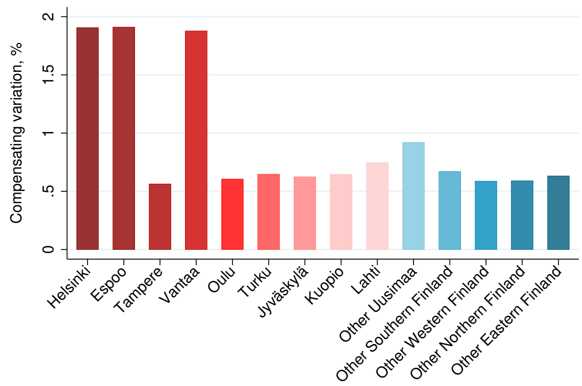
Quantifying the effects of the tunnel on the housing market How the hypothetical tunnel would affect the housing market depends in particular on the effects that it would have on labor markets. However, the effects of labor market integration on labor market outcomes, wages in particular, are theoretically ambiguous. First, integrating two labor markets into one can have positive effects on wages, if the integration leads to significant productivity gains from agglomeration (Ahlfeldt et al., 2015). Second, if the labor markets differ in initial productivity, there can be pressure for the regional wages to converge, lowering wages in the higher-income region (Dustmann et al., 2017). Third, even if there are differences in productivity across locations and an inflow of labor from the lower-income region to the higher-income region, a negative effect on the higher-region wages need not take place (Butikofer et al., 2020)[19]. Moreover, the labor market integration could have welfare effects beyond wages: it could, for example, change the relative prices of nontradable goods (services), or the desirability of different locations through local amenities.

Because of the ambiguity of the effects of a potential labor market integration, I consider two different scenarios. In the first one, the integration of the labor markets has a *positive* impact on the wages in Helsinki and the two neighboring cities, Espoo and Vantaa, labeled as "the capital region". In this scenario, I assume that the wages in the capital region increase permanently by 5 %, and there are no other effects of the tunnel in Finland on economic fundamentals. In the second scenario, the integration has a *negative* effect on wages in the capital region, which decrease permanently by 5%, and no other changes take place.

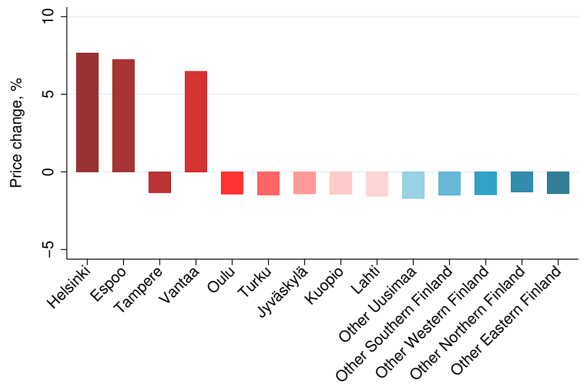
I assume the following timing for the policy: First, no changes in economic fundamentals take place between 2020 and 2023 (so that, for example, there is no global pandemic taking place in 2020). I use data from the equilibrium given the 2019 economic fundamentals from 5 to evaluate how the economy would evolve in the absence of a tunnel. Then, in the beginning of 2024, a favorable decision of the tunnel construction is announced. The tunnel is announced to be completed in 2029. The wage effect takes place immediately after, and is correctly anticipated by everyone.

Effects of a positive income shock in the capital region (scenario 1) What happens to renter welfare and housing wealth across locations if there is a change in one of the economic fundamentals in one of the locations, such as an income change in the capital region? The answer depends on the degree of household mobility as well as on housing supply responsiveness (Moretti, 2010). This is where the dynamic, nonstationary

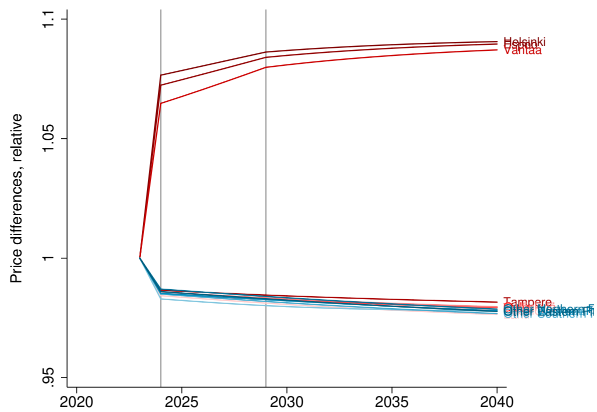
[19] Butikofer et al. (2020) study the opening of the The Øresund Bridge which connects the Danish capital Copenhagen with Sweden's third-largest city Malmö. Despite significant labor supply flows from Malmö to Copenhagen, wages in Copenhagen appear not to be affected by the bridge.



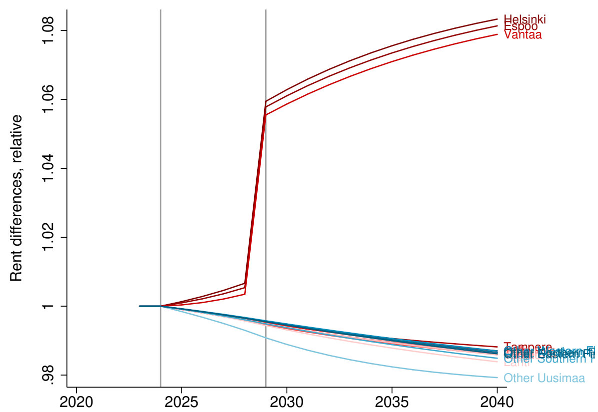
(**a**) Compensating variation (short-run)



(**b**) House prices (short-run)



(**c**) Prices



(**d**) Rents

Figure 12: Effects of a counterfactual policy, positive wage shock in the capital region in 2029 (a tunnel integrating Helsinki and Tallinn labor markets).

Notes. Compensating variation is the consumption-equivalent income increase for year 2024, relative to the equilibrium with no tunnel. Price change is measured in 2024, relative to prices in 2024 in the equilibrium with no tunnel. Prices and rents in the lower panels are reported each year relative to their values in the equilibrium with no tunnel.

model of a housing market is particularly useful. Both household mobility and housing supply changes require time - in the very short-run, both households are immobile and housing stock is fixed, but this is not necessarily the case in the medium-run. Houshold mobility in the model depends on the migration costs and on the migration elasticity. Housing supply responsiveness depends on the land availability and the housing supply elasticity. Computing the transition path of the economy allows us to compare the short-run and long-run changes.

Figure 12 summarizes housing market outcomes following a 5% income increase in the capital region, which takes place in 2029 but becomes public knowledge already in 2024. The upper left panel illustrates changes in renter welfare across locations. Renters benefit in all locations. Those renters who are initially located in the capital region benefit directly through the increasing income. Those renters, who are initially located in other locations, benefit through two channels. First, the wage increase in the capital region causes out-migration from other locations to Helsinki, implying that future rents decrease in other locations. Second, renters in other locations have the possibility of migrating to Helsinki in some future date, and the increase in this option value also increases their utility relative to the baseline. Jointly, these effects significantly mitigate regional welfare differences from local shocks: the regions which were not affected by the policy saw no initial income increases while the capital region saw a 5% income increases, but the compensating variation ranges at a smaller interval from 0.5% to 2%.

The upper right panel describes the changes in house prices. As expected, house prices increase in the capital region. The magnitude of house price increases is considerable: The 5% permanent income increases lead to 6-7% increases in house prices in the affected regions. On the other hand, house prices outside the affected regions *decline*. The lower panels in Figure 12 illustrate how house prices and rents would evolve over time as a response to the shock. House prices react immediately to the new information: The price increases take place immediately following the announcement of the tunnel. Rents, however, do not react at all to the announcement, since migration and construction do not react immediately either (in 2024, households and landlords make location and investment decisions for 2025). Rents start diverging slightly across regions after the announcement of the policy, as some households start migrating to the capital area. A large share of the rents increase takes place simultaneously with the income changes, but not all of it: Rents continue to diverge across regions for decades after the income increase. Thus, in the short- and medium-run, price-to-rent ratios will vary because of the policy.

The positive income shock would thus benefit in the capital region would therefore have asymmetric effects on the utility of renters and the value of housing across locations. Renters benefit in every location. Landlords, however, only benefit in the capital region. Landlords outside the capital region suffer from a decline in housing wealth, even if nothing happened to economic fundamentals in these other locations, because other locations become relatively less attractive than the capital region. As expected, the positive wage shock in the capital region would amplify regional divergence of housing wealth.

Effects of a negative income shock in the capital region (scenario 2) In the alternative scenario, the tunnel has a *negative* effect on wages in the capital region. The welfare effects in this scenario are reported in appendix G.1. These results are qualitatively the mirror image of Figure 12, but not in terms of magnitudes, since the two scenarios imply different incentives for example for construction that we would not expect to be symmetric. For example, when there is a positive shock in the capital region, house prices in other regions decrease by approximately 1.5%. However, when there is a symmetric but negative income shock in the capital region, houses prices elsewhere increase, but significantly less, on average by 0.7%. This is due to these other regions having a relatively responsive housing supply from the higher availability of buildable land.

Alternative counterfactual experiment I appendix G.2, I report the evaluation of another counterfactual public policy that the government *could have implemented* in the beginning of my sample period. For this alternative counterfactual policy, I consider a 5% regional income subsidy to the declining regions, financed via an income tax on residents of large cities. Such a policy redistributes welfare from renters in large cities to renters in deprived locations. The policy would have mitigated regional differences in house prices, but only temporarily: only 5 years after the policy implementation, prices in many locations are back to their initial levels, even if the income subsidy is permanent. This suggests that redistribution of income across locations might not be an effective tool for fighting regional welfare differences, and other policies such as migration subsidies could be more efficient.

7 Conclusions

In this paper, I set up a dynamic spatial equilibrium model of the housing market in order to study the welfare effects of regional divergence of apartment rents and apartment prices. As I am interested in a gradual divergence over time, I need a model for which I can compute the nonstationary transitions to a steady-state equilibrium: applying insights from the trade literature allows me to do so in a tractable way. The model nests many of the previously suggested causes of regional rent-price ratio divergence, and therefore it can be used to assess their relative importance quantitatively.

In my emprical application, I take the model to data in Finland and use the model to understand the causes and consequences of regional divergence. I find that in Finland, there has been a significant regional divergence of welfare across locations both for homeowners and for renters. This is visible from regionally diverging prices, but not visible from regional rents. The key takeaway from the paper is that housing affordability (as measured by rents) and housing wealth (as measured by house prices) do not always go hand-in-hand.

The model can be used to study a number of welfare questions that have previously not been adressed in the literature, but as the framework is relatively stylized, it also calls for a number of possible extensions. One of them is be to incorporate also exogenous changes to location-specific population, stemming for example from location/specific mortality and fertility rates. Fertility and mortality rates are possibly important drivers of regional divergence in countries like Finland, where the population is aging. Another extension that would be relevant for understanding welfare effects of location-specific shocks would be to compare separately renter utility across "incumbent renters", who rent housing in a specific location already before a location-specific shock takes place, and the renters who will move in to a specific location following a shock. If rental contracts have sticky prices then incumbent renters will not have to pay higher prices even if a location-specific shock would take place. Third, it would be interesting to model landlord utility using some other specification than a linear one to understand how changes in house prices can affect their welfare nonlinearly (an alternative would be a CRRA utility function, yielding closed-form expressions for capital accumulation).

An important limitation of the setup is related to uncertainty: the model assumes away any uncertainty related to housing investment, and therefore the discount rates, which presumably reflect at least in part location-specific risk premia, do not have a structural interpretation. While incorporating uncertainty would of course yield more realistic conclusions about the rent and price effects of shocks in a spatial model, current tools in the literature do not yet allow for simple computation of equilibria in dynamic spatial models under uncertaintly, making model inversion exercises with uncertainty infeasible.

The framework also calls for future work on modeling homeownership in a spatial model. A key margin which is missing from the model is the decision between owning a home or renting. Homeownership is important for both understanding *why* rents and prices evolve differently and *who* benefits from these changes. If the housing market were frictionless, then modeling homeownership explicitly or modeling renters and landlords separately like I do should not lead to different conclusions about price and rent divergence, but this is not true in the presence of market frictions. In the real world, changes such as the decline in risk-free interest rates can affect the homeownership versus renting margin, affecting rents and prices, which is not captured in my model. Factors like credit constraints are also beyond the scope of this model. Moreover, modeling homeownership can improve our understanding of the relative importance of the changes. For example, according to my estimates, a homeowner in Eastern Finland has suffered losses from the decline in the value of their homes, but also benefited from increasing amenities. Is the amenity increase important enough to outweigh the wealth loss from house values? I leave this question for future work.

References

Ahlfeldt, G. M., Bald, F., Roth, D., & Seidel, T. (2020). Quality of life in a dynamic spatial model. *Available at SSRN 3751857*.

Ahlfeldt, G. M., Heblich, S., & Seidel, T. (2021). Micro-geographic property price and rent indices.

Ahlfeldt, G. M., Redding, S. J., Sturm, D. M., & Wolf, N. (2015). The economics of density: Evidence from the berlin wall. *Econometrica*, *83*(6), 2127–2189.

Albouy, D. (2008). *Are big cities bad places to live? estimating quality of life across metropolitan areas* (Tech. Rep.). National Bureau of Economic Research.

Amaral, F., Dohmen, M., Kohl, S., & Schularick, M. (2022). The fall in the risk-free rate and rising house price dispersion. *Unpublished manuscript*.

Arcidiacono, P., & Miller, R. A. (2011). Conditional choice probability estimation of dynamic discrete choice models with unobserved heterogeneity. *Econometrica*, *79*(6), 1823–1867.

Balboni, C. A. (2019). *In harm's way? infrastructure investments and the persistence of coastal cities* (Unpublished doctoral dissertation). London School of Economics and Political Science.

Bischoff, O. (2012). Explaining regional variation in equilibrium real estate prices and income. *Journal of Housing Economics*, *21*(1), 1–15.

Bryan, G., & Morten, M. (2019). The aggregate productivity effects of internal migration: Evidence from indonesia. *Journal of Political Economy*, *127*(5), 2229–2268.

Büchler, S., Ehrlich, M. v., & Schöni, O. (2021). The amplifying effect of capitalization rates on housing supply. *Journal of urban economics*, *126*, 103370.

Butikofer, A., Løken, K., & Willén, A. (2020). Building bridges and widening gaps: Wage gains and equity concerns of labor market expansions. *Available at SSRN 3518628*.

Caliendo, L., Dvorkin, M., & Parro, F. (2019). Trade and labor market dynamics: General equilibrium analysis of the china trade shock. *Econometrica*, *87*(3), 741–835.

Campbell, J. Y., & Shiller, R. J. (1988). The dividend-price ratio and expectations of future dividends and discount factors. *The Review of Financial Studies*, *1*(3), 195–228.

Campbell, S. D., Davis, M. A., Gallin, J., & Martin, R. F. (2009). What moves housing markets: A variance decomposition of the rent–price ratio. *Journal of Urban Economics*, *66*(2), 90–102.

Cavalleri, M. C., Cournède, B., & Özsöğüt, E. (2019). How responsive are housing markets in the OECD? National level estimates. (1589).

Chapelle, G., Eyméoud, J.-B., Bruneel-Zupanc, C., & Wasmer, E. (2022). Housing prices propagation: A theory of spatial interactions. *Unpublished manuscript*.

Cun, W., & Pesaran, M. H. (2022). A spatiotemporal equilibrium model of migration and housing interlinkages. *Journal of Housing Economics*, *57*, 101839.

Davis, M. A., Fisher, J. D., & Veracierto, M. (2013). Gross migration, housing and urban population dynamics. *Working Paper No. 2013-19, Federal Reserve Bank of Chicago, Chicago, IL*.

Davis, M. A., Fisher, J. D., & Whited, T. M. (2014). Macroeconomic implications of agglomeration. *Econometrica*, *82*(2), 731–764.

de Haan, J., & Diewert, E. (2013). Hedonic regression methods. Eurostat, Luxembourg.

Diamond, R. (2016). The determinants and welfare implications of us workers' diverging location choices by skill: 1980-2000. *American Economic Review*, *106*(3), 479–524.

Diewert, W. E., Heravi, S., & Silver, M. (2009). Hedonic imputation versus time dummy hedonic indexes. In *Price index concepts and measurement* (pp. 161–196). University of Chicago Press.

Dustmann, C., Schönberg, U., & Stuhler, J. (2017). Labor supply shocks, native wages, and the adjustment of local employment. *The Quarterly Journal of Economics*, *132*(1), 435–483.

Eerola, E., Lyytikäinen, T., & Vanhapelto, T. (2020). Asuntojen hintojen alueellinen eriytyminen suomessa. *Valtion taloudellinen tutkimuskeskus*.

Eerola, E., & Saarimaa, T. (2018). Delivering affordable housing and neighborhood quality: A comparison of place-and tenant-based programs. *Journal of Housing Economics*, *42*, 44–54.

Favilukis, J., Ludvigson, S. C., & Van Nieuwerburgh, S. (2017). The macroeconomic effects of housing wealth, housing finance, and limited risk sharing in general equilibrium. *Journal of Political Economy*, *125*(1), 140–223.

Giannone, E., Li, Q., Paixao, N., & Pang, X. (2020). Unpacking moving. *Unpublished manuscript*.

Glaeser, E. L., Gyourko, J., Morales, E., & Nathanson, C. G. (2014). Housing dynamics: An urban approach. *Journal of Urban Economics*, *81*, 45–56.

Goodman, A. C. (1988). An econometric model of housing price, permanent income, tenure choice, and housing demand. *Journal of urban economics*, *23*(3), 327–353.

Greaney, B. (2022). The distributional effects of uneven regional growth. *Unpublished manuscript*.

Greenwald, D. L., & Guren, A. (2021). Do credit conditions move house prices? *Unpublished manuscript*.

Halket, J., & Vasudev, S. (2014). Saving up or settling down: Home ownership over the life cycle. *Review of Economic Dynamics*, *17*(2), 345–366.

Harding, J. P., Rosenthal, S. S., & Sirmans, C. (2007). Depreciation of housing capital, maintenance, and house price inflation: Estimates from a repeat sales model. *Journal of urban Economics*, *61*(2), 193–217.

Henkel, M., Seidel, T., & Suedekum, J. (2021). Fiscal transfers in the spatial economy. *American Economic Journal: Economic Policy*, *13*(4), 433–68.

Herkenhoff, K. F., Ohanian, L. E., & Prescott, E. C. (2018). Tarnishing the golden and empire states: Land-use restrictions and the us economic slowdown. *Journal of Monetary Economics*, *93*, 89–109.

Hilber, C. A., & Mense, A. (2021). Why have house prices risen so much more than rents in superstar cities?

Hill, R. J., Scholz, M., Shimizu, C., & Steurer, M. (2018). An evaluation of the methods used by european countries to compute their official house price indices. *Economie et Statistique*, *500*(1), 221–238.

Hornbeck, R., & Moretti, E. (2022). Estimating who benefits from productivity growth: local and distant effects of city productivity growth on wages, rents, and inequality. *The Review of Economics and Statistics*, 1–49.

Hotz, V. J., & Miller, R. A. (1993). Conditional choice probabilities and the estimation of dynamic models. *The Review of Economic Studies*, *60*(3), 497–529.

Howard, G., & Liebersohn, J. (2020). Regional divergence and house prices. *Fisher College of Business Working Paper*(2020-03), 004.

Hsieh, C.-T., & Moretti, E. (2019). Housing constraints and spatial misallocation. *American Economic Journal: Macroeconomics*, *11*(2), 1–39.

Kaplan, G., Mitman, K., & Violante, G. L. (2020). The housing boom and bust: Model meets evidence. *Journal of Political Economy*, *128*(9), 3285–3345.

Karlman, M. (2022). Why have house prices diverged? the role of the interest rate. *Unpublished manuscript*.

Kleinman, B., Liu, E., & Redding, S. J. (2021). Dynamic spatial general equilibrium. *Unpublished manuscript*.

Kline, P., & Moretti, E. (2014). People, places, and public policy: Some simple welfare economics of local economic development programs. *Annu. Rev. Econ.*, *6*(1), 629–662.

Levy, A. (2021). Housing policy with home-biased landlords: Evidence from french rental markets.

Miles, D., & Monro, V. (2021). Uk house prices and three decades of decline in the risk-free real interest rate. *Economic Policy*, *36*(108), 627–684.

Molloy, R., Nathanson, C., & Paciorek, A. (2020). Housing supply and affordability: Evidence from rents, housing consumption and household location.

Moretti, E. (2010). Local Labor Markets. In O. Ashenfelter & D. Card (Eds.), *Handbook of Labor Economics* (Vol. 4, p. 1237-1313). Elsevier.

Moretti, E. (2013). Real wage inequality. *American Economic Journal: Applied Economics*, *5*(1), 65–103.

Notowidigdo, M. J. (2020). The incidence of local labor demand shocks. *Journal of Labor Economics*, *38*(3), 687–725.

OECD. (2019). Household spending (accessed on 01 october 2022). Retrieved from https://www.oecd-ilibrary.org/content/data/b5f46047-en doi: https://doi.org/10.1787/b5f46047-en

OECD. (2021). *Brick by brick: Building better housing policies. Chapter 5, figure 5.5.* Organisation for Economic Cooperation and Development.

Retrieved from https://www.oecd-ilibrary.org/sites/08b39994-en/index.html?itemId=/content/component/08b39994-en#figure-d1e430

Oikarinen, E., Peltola, R., & Valtonen, E. (2015). Regional variation in the elasticity of supply of housing, and its determinants: The case of a small sparsely populated country. *Regional Science and Urban Economics*, *50*, 18–30.

Roback, J. (1982). Wages, rents, and the quality of life. *Journal of political Economy*, *90*(6), 1257–1278.

Rosen, S. (1979). Wage-based indexes of urban quality of life. *Current issues in urban economics*, 74–104.

Suhonen, T., & Karhunen, H. (2019). The intergenerational effects of parental higher education: Evidence from changes in university accessibility. *Journal of Public Economics*, *176*, 195–217.

Suzuki, Y. (2021). Local shocks and regional dynamics in an aging economy. *Unpublished manuscript*.

Van Nieuwerburgh, S., & Weill, P.-O. (2010). Why has house price dispersion gone up? *The Review of Economic Studies*, *77*(4), 1567–1606.

Warnes, P. E. (2020). Transport infrastructure improvements and spatial sorting: Evidence from Buenos Aires. *Unpublished manuscript*.

Wilhelmsson, M. (2008). House price depreciation rates and level of maintenance. *Journal of Housing Economics*, *17*(1), 88–101.

Yiu, C. Y., & Wong, S. K. (2005). The effects of expected transport improvements on housing prices. *Urban studies*, *42*(1), 113–125.

Yoon, C. (2017). Estimating a dynamic spatial equilibrium model to evaluate the welfare implications of regional adjustment processes: the decline of the Rust Belt. *International Economic Review*, *58*(2), 473–497.

Zerecero, M. (2021). The birthplace premium. *Unpublished manuscript*.

Appendix

A Descriptive evidence on price and rent divergence

A.1 Constructing hedonic indices

In order to document trends in dwelling prices and dwelling rents, adjusting for potential quality variation, I use a "time-dummy" method, falling within the larger umbrella of methods used to construct quality-adjusted hedonic indices (i.a. de Haan & Diewert (2013), Diewert et al. (2009)). My adaptation of the "time-dummy" method consists of regressing the log of an outcome y_i, either the price or the rent of an apartment i, on apartment-level characteristics x, region-fixed effects and time-by-region fixed effects:

$$\ln y_{it} = \alpha + \beta X_{it} + \gamma_r + \delta_{tr} + \epsilon_{it}$$

where t indexes time and r the region. The time-by-region coefficients for the base year of the index are omitted from the regression and normalized to 0 for each region. This allows interpreting the values for $\exp(\hat{\delta}_{tr}) \quad \forall t \neq t_{base}$ as an index number. For example, comparing the price (or rent) of an apartment of characteristics $\overline{x}$, sold in year z, compared to an apartment of the same characteristics, sold in 2009, if 2009 is the base year of the index:

$$\frac{\hat{y}(\overline{x})_{s,r}}{\hat{y}(\overline{x})_{2009,r}} = \frac{\exp(\hat{\alpha} + \hat{\gamma_r} + \hat{\beta}\overline{x} + \hat{\delta}_{s,r})}{\exp(\hat{\alpha} + \hat{\gamma_r} + \hat{\beta}\overline{x} + \delta_{2009,r})} = \frac{\exp(\hat{\delta}_{s,r})}{\exp(\delta_{2009,r})} = \exp(\hat{\delta}_{s,r})$$

and thus $\exp(\hat{\delta}_{tr})$ can be interpreted as the price or the rent index number. I also construct similar indices where the characteristic vector x and the region fixed effect γ_r are replaced by an apartment-level fixed effect, and I refer to these indices as "resales" or "relisting" indices. Moreover, I interpret the exponential of the bounds of the confidence interval of $\hat{\delta}_{tr}$ as confidence intervals for the regional price indices:

$$\left[\exp(\hat{\delta}_{s,r} - 1.96 \cdot se(\hat{\delta}_{s,r})), \exp(\hat{\delta}_{s,r} + 1.96 \cdot se(\hat{\delta}_{s,r}))\right] \tag{51}$$

This "hedonic time dummy" method has the attractive feature that it is simple to compute and to interpret. Official statistical agencies, who produce annual house price indices, often use different techniques (inputation methods) because they need to add additional time periods to the index without changing the previous estimates (Hill et al. (2018)). However, this is not a concern for me, as I do not need to deal with adding additional years to the data. Official Statistics Finland house price indices are very consistent with my measures.

A.2 Microdata on rents and prices

The microdata I use to document time trends of apartment rents and prices in Finland comes from two separate sources.

Prices To describe dwelling prices, I use transaction price data collected by the Finnish Federation of Real Estate Agency KVKL (*Kiinteistönvälitysalan keskusliitto*). The dataset contains information on all transactions intermediated by the member real estate agencies of the organization, and it covers more than 60% of all apartment transactions taking place in Finland[20]. The advantage of using this dataset as opposed to survey data or aggregate data preprocessed by an official statistical agency comes from the richness of the data: on top of the transaction price, we observe apartment-level characteristics such as the address of the apartment, the type and age of the building, the number of rooms and floor area, the amount of co-operative debt associated with the transaction, etc.

Throughout the analysis, I restrict the transaction sample only to apartments in multi-unit buildings because individual properties (detached homes) are rarely rented, and this could make the composition of the apartment transaction sample different from the rental sample. I exclude all new observations, often sold directly by development companies instead of individuals. I exclude outliers based on floor area, transaction price and building age. I exclude observations where information on the price, sales date or location is missing. In order to include apartment-level fixed effects in the regressions, I construct apartment id's based on the following characteristics of the apartment: The street address, zip code, number of rooms, floor number, and the floor area. Some misclassification is possible if the street address information is not complete and does not contain information on the apartment number within the building.

Column (1) in table A1 summarizes the transaction data.

Rents To describe dwelling rents, I use rents listings data from an online listings platform Vuokraovi.com. This listings service is a platform on which landlords may post announcements of available rental units in exchange for a fee, and it is the second-largest such platform Finland. From this dataset I can access the listings rents which reflect the ask prices by the landlors and might not be perfectly representative of the

[20]For example, in 2015, 28 000 transactions of non-new dwelling apartments (excluding individual property) were recorded by KVKL, and there were 43 000 such trades in total in Finland taking place according to Statistics Finland.

Table A1: Summary statistics for the transaction data.

	Sample	Main specification
	mean	mean
Transaction price	163593.0	150397.0
Operating costs (monthly)	196.0	185.7
Apartment surface	66.0	52.9
Rooms	2.5	2.0
Age	37.0	40.2
Observations	491988	67969

Notes. Transaction price is expressed net of any cooperative debt. Transaction price and maintenance cost are expressed in 2020 euros. "Sample" refers to all apartments in multi-unit buildings (no detached homes), excluding new apartments, between 2009 and 2019. "Main specification" refers to 2-bedroom apartments in blocks of flats where the condition is denoted good, between 2012 and 2019.

realized rents on new contracts, but anecdotal evidence does not suggest important biases arising from pricing mistakes or strategies by landlords. Again, the advantage of using listings data as opposed to survey data comes from the fine granularity of the data: on top of the listing time and the listing rent, I observe the apartment location, size and quality, building type, type of rental contract suggested, and so on.

I restrict the sample to private rental units (excluding public and semi-public rentals, including free-market rentals from both private individual landlords and landlord corporations), as the rents of new rental contracts in these private rental units are unregulated in Finland. I exclude listings for finite-term contracts. Otherwise, I select the sample similar to the transaction data. I restrict the sample to apartments in multi-unit buildings and exclude listings in new developments. I exclude outliers on rent, floor area and building age. I exclude observations with missing values of rent, floor area or location . Again, the apartment-level fixed effects are constructed using the observable characteristics of the apartment: the street address, the zipcode, the floor, the number of rooms and the floor area. Some misclassification is possible due to some of the apartment numbers missing. Moreover, I exclude any observations where an observationally same listing reappears in the data as a new listing in less than 180 days after the previous listing. I do so in order to avoid treating repostings of old listings (if, for example, the original listing was not successful) as new listings.

Column (1) in table A2 summarizes the rent listings data.

Control variables in hedonic indices I run two types of price and rent regressions: *i*) regressions with quality controls and *ii*) resales / relisting regressions.

In the regression used to measure price divergence with quality controls, I regress log price net of cooperative debt on floor area, floor area squared, age, age squared,

Table A2: Summary statistics for the rent listings data

	Sample	Main specification
	mean	mean
Listing rent (monthly)	738.4	742.3
Apartment surface	52.2	51.3
Rooms	2.0	2.0
Age	38.6	36.2
Observations	360214	100780

Notes. "Sample" refers to the set of observations used to compute indices of house prices and rents in appendix ... , which consists of aparments in multi-unit housing in non-new dwellings in 2009-2019. "Main specification" refers to the sample used to compute indices in the main text, which refers to 2-bedroom apartments in blocks of flats in good condition in 2012-2019. Listing rents are measured in 2020 euros. Apartment surface is measured in square meters and building age in years.

number of rooms, floornumber, maintenance charge (building operating costs), a categorical variable for building type, a categorical variable for condition, a categorical variable for the number of rooms, a categorical variable characterizing land lot ownership status as well as zipcode fixed effects. In the respective regression for rents, I control for floor area, floor area squared, age, age squared, floor number, a categorical variable for building type, a categorical variable for condition, a categorical variable for the number of rooms, an indicator for whether the apartment is immediately available, an indicator for the owner type (individual or a company), and zipcode fixed effects. I cluster the standard errors at the zipcode level.

In the relisting/resales regressions, I include apartment-level fixed effects. This implies that I compare repeated observations of the same apartment over time[21]. These indices therefore identify how the price or rent of a given apartment has evolved over time. In these regressions, I cluster the standard errors at the apartment level.

Sample selection for the main analysis In my main empirical application, I measure regional apartment price and rents indices using hedonic regressions with strict sample selection in order to ensure that I measure rents and prices of comparable apartments (instead of comparing, say, family home prices to studio rents). For both the rent and price regressions I only use apartments listed with number of rooms equal to two (typically, a living room and a bedroom), in blocks of flats, in good condition, and excluding new buildings. Prices and rents in the sample after these stricter sample restrictions are summarized in the second columns of tables A1 and A2.

To measure rent and price changes, I use the specification with quality controls, where the list of controls is the same as above.

[21] Up to any misclassification caused by imperfectly observed street address, for example.

I also need to measure rents and prices (as well as operating costs, see below) in levels in the baseline year (this is used for the model inversion exercise). For this purpose, I keep the same sample as for the main hedonic regressions, and compute simple averages of rents, prices and operating costs per square meter by year and by location. Since I do not need the levels for multiple years, using simple averages should not lead to selection issues from changes in the composition of the sample. I cluster the standard errors at the zipcode level.

To measure price-to-rent ratios of two-bedroom apartments in good condition in blocks of flats in figure 7, I use the raw averages for the initial year (2012) and I compute changes relative to the initial year using my measure of hedonic price and rent indices with quality controls. Note that I do not use the rents and prices of the *same* apartmetns for the computation of price-to-rent ratios, as one could do for example to compute price-to-dividend ratios if one had data on stock prices and dividends of specific firms.

The first observed housign transactions are from 2003 and the first observed rents from 2008, which sets bounds on what time periods I can consider in my analysis. In appendix section A.3, I will describe the evolution of rents and prices from 2009 to 2019. However, in the main analysis of the paper, I take 2012 as the base year for the indices to restrict the sample to post-financial crisis recovery, but my results are not sensitive to the selection of the base year. I restrict my sample to the pre-2020 period to avoid any counfounders stemming from the global pandemic.

Operating costs The owners of apartments in blocks of flats must pay monthly fees to building co-operatives to participate in building-level operating costs such as cleaning, heating, et cetera. In the empirical application in the main paper, these fees are deducted from rents to compute the net rents that landlords receive (denoted r^N in the paper). Operating costs are observed in the KVKL HSP transactions dataset, and therefore I treat operating costs, denoted by ξ_l in the paper, as observed. They are measured as the mean operating costs per m^2 in the same sample that is used to measure prices, excluding any repayment of housing cooperative debt (Finnish: *hoitovastike*). Typically these operating costs cover only costs related to the direct operations of the building. However, in some instances, the housing cooperative can decide to use this revenue also to conduct building maintenance. This should be taken into account in our choice of the depreciation rate (Harding et al. (2007)).

Figure A1 illustrates the evolution of the operating costs in the KVKL transaction sample relative to the rents as measured in the vuokraovi.com sample, using the same sample selection and regional classification as the main analysis in the paper. Operating cost shares are relatively high - in the order of magnitude of 20-30% of the rent.

Costs are higher outside the largest cities, probably because in cities, apartment building cooperatives receive more rental revenue for example by letting out commercial spaces in ground floors. The estimaets used to measure ξ_l are simply the averages from 2012-2019 by location. The share is the lowest in Helsinki (0.21), and the highest in Other Southern Finland (0.31).

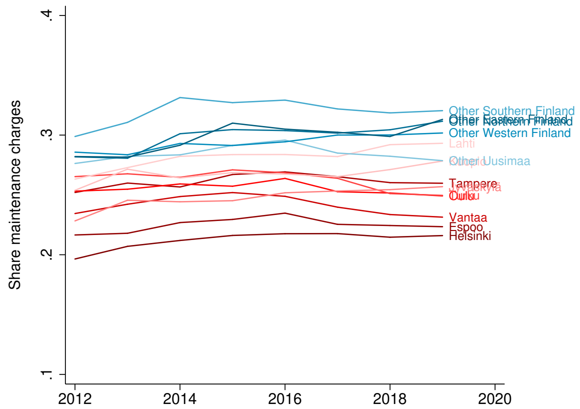


Figure A1: Share building charges of rents

Notes. The average operating costs per m^2 relative to average rents per m^2, by location. Both are computed for 2-room apartments in blocks of flats in good condition. Operating costs are measured using the KVKL data and listing rents using the vuokraovi.com data.

Data references:

Prices. Kiinteistönvälitysalan Keskusliitto ry, KVKL Hintaseurantapalvelu. Data accessed via VATT Institute for Economic Research.

Listing rents. vuokraovi.com. Data accessed via VATT Institute for Economic Research.

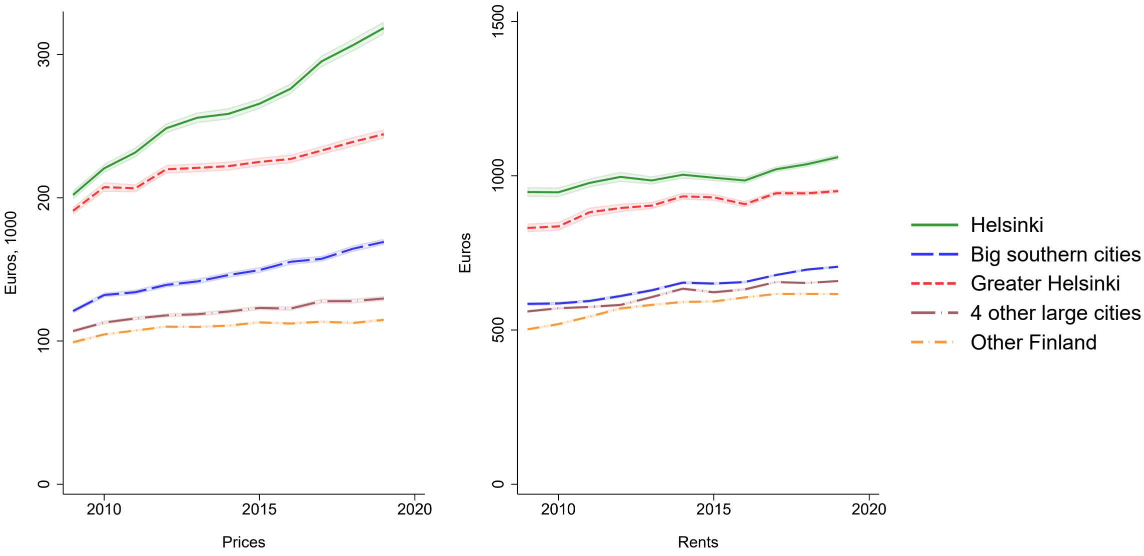
A.3 Regional divergence of house prices and rents in Finland

In this section, I describe the evolution of house prices and rents across Finnish regions in the decade following the financial crisis using transactions data from KVKL and rental listings data from `vuokraovi.com`. The data sources and sample selection are summarized in Appendix A.2.

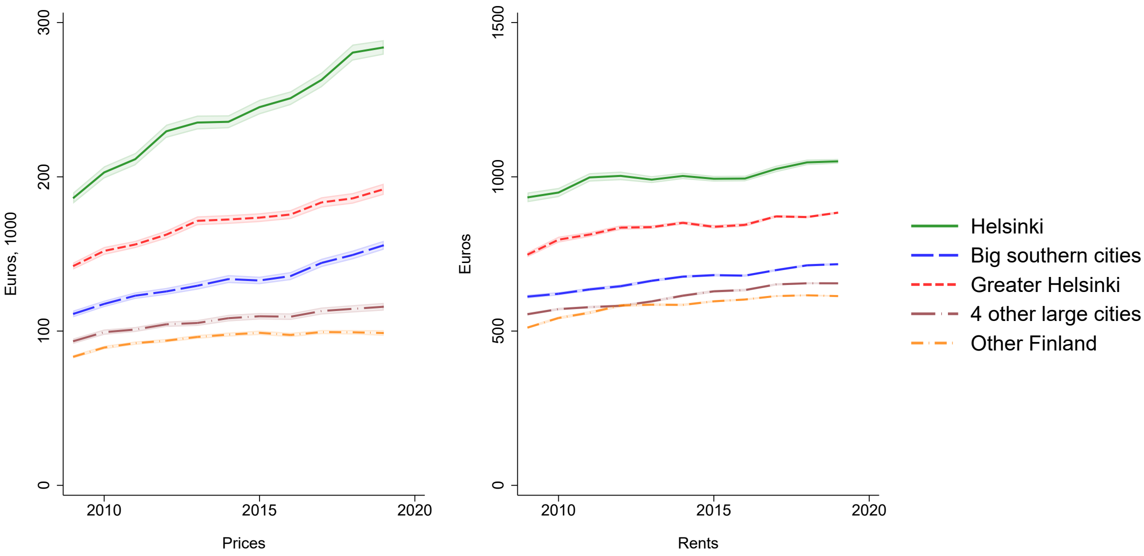
The regional classification is selected to mirror that in the main text: the capital Helsinki is displayed separately, and "Other Greater Helsinki" refers to the neighboring cities Espoo and Vantaa. Tampere and Turku are the two other large Southern metropolitan areas, and the "Other 4 large cities" refers to the remaining cities with a population above 100 000. Finally, "Other Finland" pools together the remaining regions (Other Northern, Eastern, Western, Southern, and other Uusimaa).

To illustrate the differences in levels and in changes in prices and rents, the top panel in Figure A2 plots the average nominal prices and rents of apartment transactions in multi-unit housing across Finnish regions. These raw averages do not include any controls. The average nominal transaction price in Helsinki has increased from $\approx$ 200 000 euros in 2009 to more than $\approx$ 300 000 euros in 2019. Simultaneously, price increases in other regions are much more modest - for example, transaction prices in "Other large cities" are up from $\approx$ 107 000 euros to $\approx$ 130 000 euros. Average listing rents in Helsinki have not increased by a quantity comparable to the price increase - to the contrary, average rental growth has been slightly faster in other regions than Helsinki.

The bottom panel of the figure reports the same averages in a sample selected as in my main specification: 2-bedroom apartments in good condition in blocks of flats. This figure mirrors the top panel. While in 2009, the price of a 2-bedroom aparment in Helsinki was approximately 2 times higher than it was in middle-sized cities, in 2019 the prices were almost three-fold. On the other hand, rental differences have, if anything, mitigated through the time period.



(a) All apartments in multi-unit housing



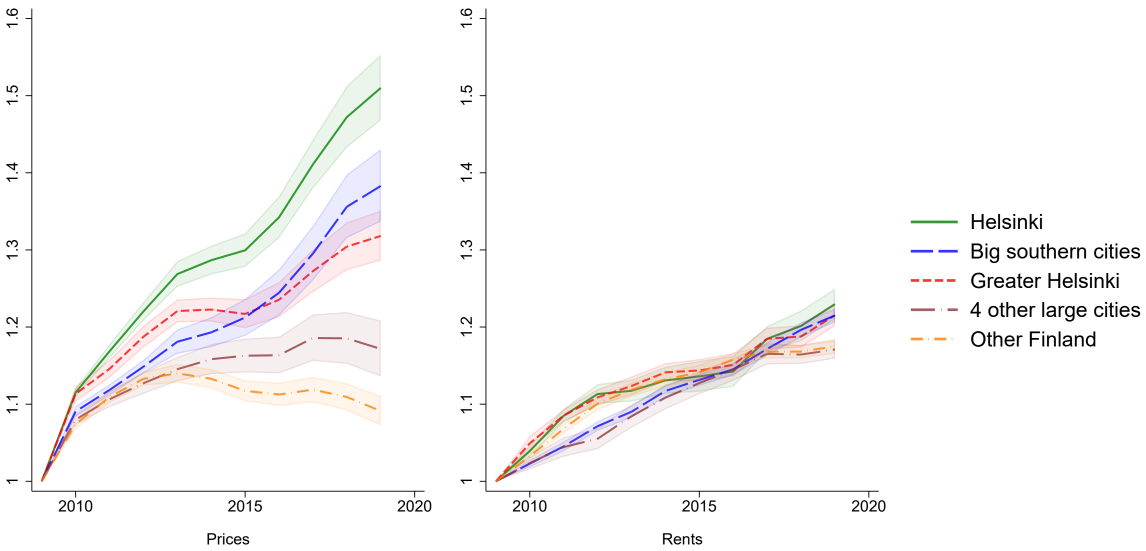
(b) 2-bedroom apartments in blocks of flats in good condition

Figure A2: Average prices and rents across Finnish geographic areas

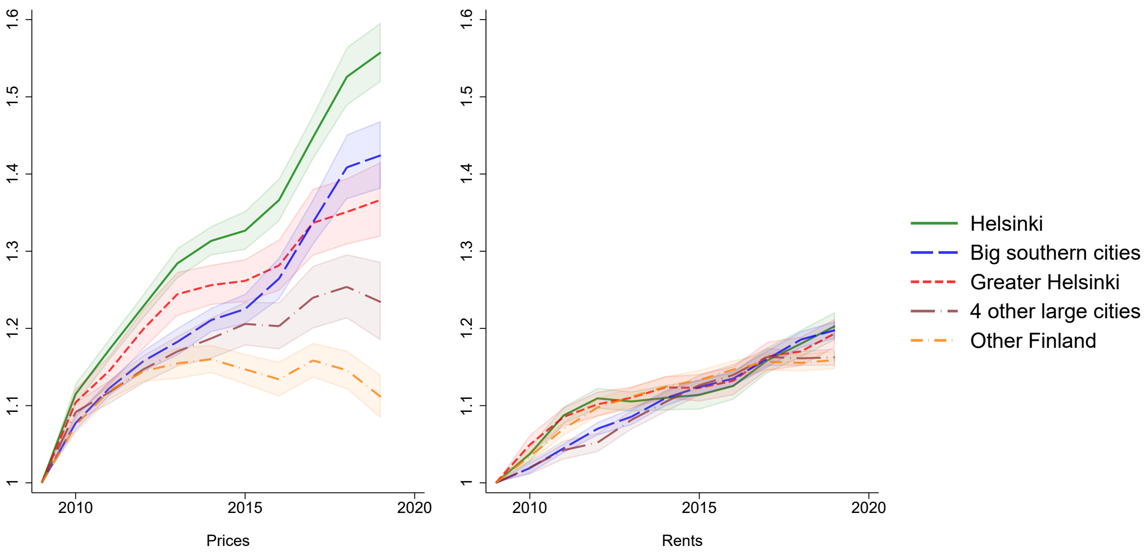
Notes: Figure plots the mean and the 95% confidence interval of the mean. Prices and rents are nominal. Other Greater Helsinki refers to the two cities neighboring Helsinki: Espoo and Vantaa. Big Southern cities refers to Turku and Tampere which are the other large metropolitan areas in the country. Next 4 largest cities refers to Oulu, Jyväskylä, Kuopio and Lahti. Other Finland contains all other locations.

In order to address concerns about changes in sample composition, Figure A3 plots point estimates and confidence intervals from hedonic price and rent indices with quality controls. The construction of these indices is described in Appendix A.1. The top panel shows price and rent indices for all apartments in the sample and the bottom panel for the sample used in the main specification. All regressions control for among others floor area, floor area squared, age, age squared, and include zipcode fixed effects. For a full list of controls, see appendix A.2.

The coefficients are largely in line with the raw means in Figure A2. Nominal prices in Helsinki are up by $\approx 50\%$, and, while prices in medium-sized cities are up by only ≈ 15 % and they stagnate in particular in the latter half of the sample period. Yet, the evolution of rents has been nearly identical in all regions. The indices for the total sample (in the top panel) and in the main specification sample (in the bottom panel) are very comparable.



(**a**) All apartments in multi-unit housing

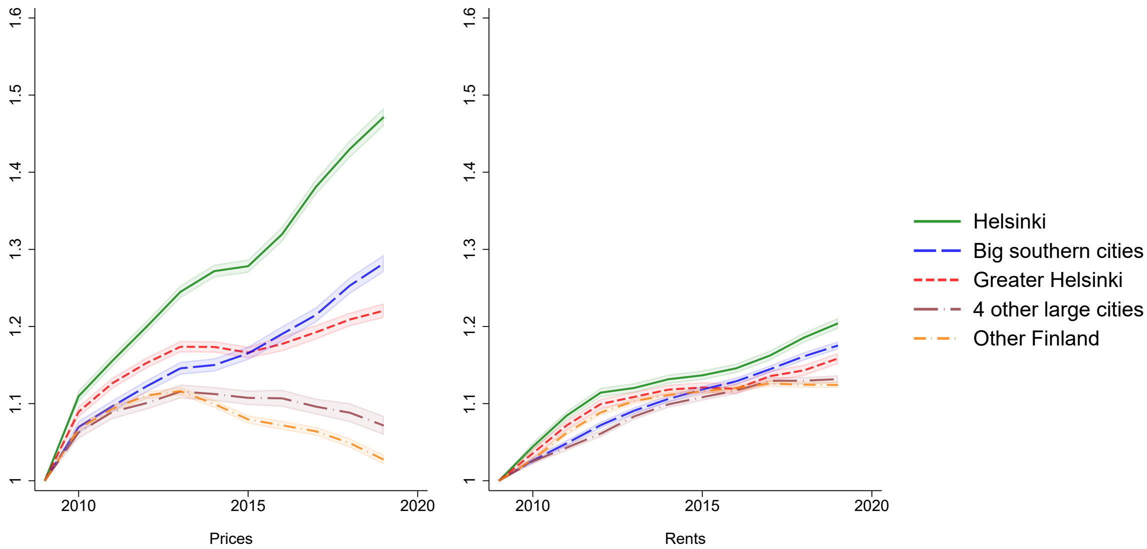


(**b**) 2-bedroom apartments in blocks of flats in good condition

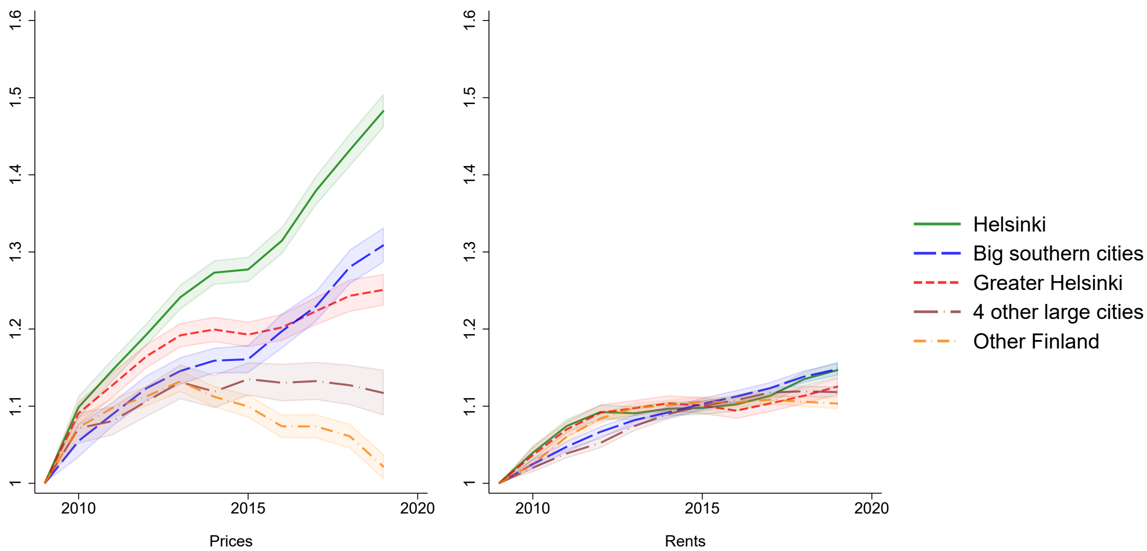
Figure A3: Hedonic price and rent indices across Finnish geographic areas

Notes: Figure plots coefficients from hedonic regressions and 95% confidence intervals, obtained via a hedonic regression with quality controls. Both regressions control among others for apartment floor area, age and zipcode fixed effects. For a full list of controls, see appendix section A.2. Standard errors are clustered at the zipcode level. Prices and rents are nominal. Other Greater Helsinki refers to the two cities neighboring Helsinki: Espoo and Vantaa. Big Southern cities refers to Turku and Tampere which are the other large metropolitan areas in the country. Next 4 largest cities refers to Oulu, Jyväskylä, Kuopio and Lahti. Other Finland contains all other locations.

Moreover, Figure A4 plots coefficients from regressions in which which apartment-level controls are replaced by apartment-level fixed effects. These regressions control for unobserved time-invariant apartment characteristics that could drive the regional divergence by using repeated observations of the same apartment. Hence they can be interpreted as resales or relisting indices. However, the divergence pattern in prices persists, and the estimates are significantly more precisely estimated.



(**a**) All apartments in multi-unit housing



(**b**) 2-bedroom apartments in blocks of flats in good condition

Figure A4: Resale price / relisting rent indices indices across Finnish geographic areas

Notes: Figures plot coefficients and 95% confidence intervals from hedonic regressions with apartment level fixed effects. For the construction of the apartment-id's, see appendix section A.2. Standard errors are clustered at the apartment level. Prices and rents are nominal. Other Greater Helsinki refers to the two cities neighboring Helsinki: Espoo and Vantaa. Big Southern cities refers to Turku and Tampere which are the other large metropolitan areas in the country. Next 4 largest cities refers to Oulu, Jyväskylä, Kuopio and Lahti. Other Finland contains all other locations.

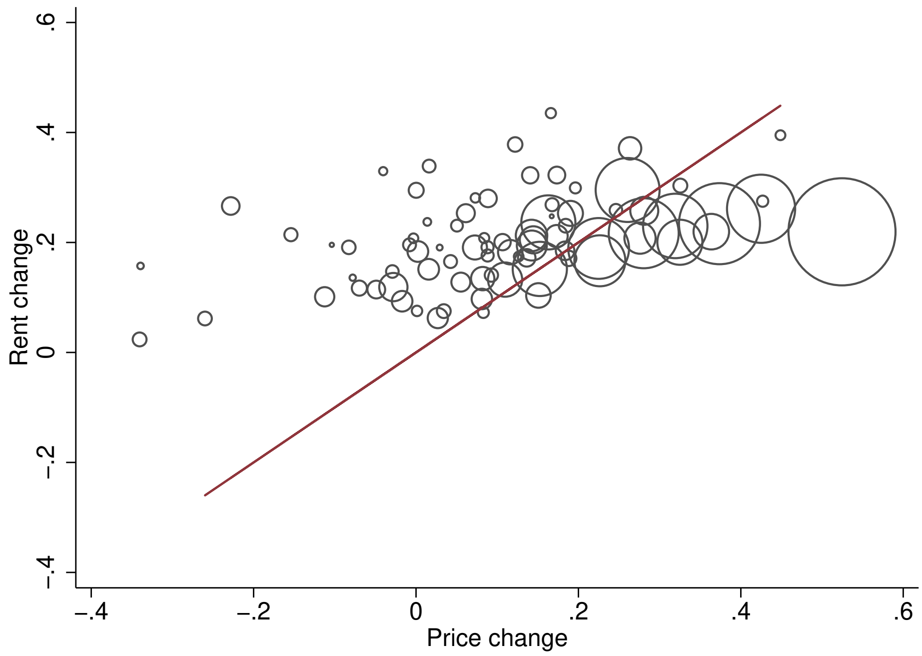


Figure A5: Mean prices and rents per m^2 in 2019 relative to 2009 by municipality.

Notes. Nominal mean prices and rents per m^2 in 2019 relative to respective means in 2009. Size of circle reflects the number of observations No quality controls. Municipalities with less than 10 rent or price observations per year in either 2009 or 2019 are excluded, leading to a total of 80 municipalities. Red line depicts the 45 degree line.

Finally, to document price and rent changes at a less aggregated level, Figure A5 plots raw means or rents and prices per m^2 in 2019 relative to 2009 by *municipality*. The size of the circle reflects the number of observations in the transaction sample. This figure suggests that *i*) the variance of price changes is much larger than the variance of rents changes *ii*) price growth is stronger in larger cities, *iii*) in most large cities, price changes have been faster than rent changes (dots are to the right of the 45 degree line) *iv*) in smaller cities, price changes are smaller than rent changes (dots are to the left of the 45 degree line). This figure confirms the finding from the hedonic indices: there is no straight-forward mapping from rental growth to price growth.

A.4 Alternative rent and price data

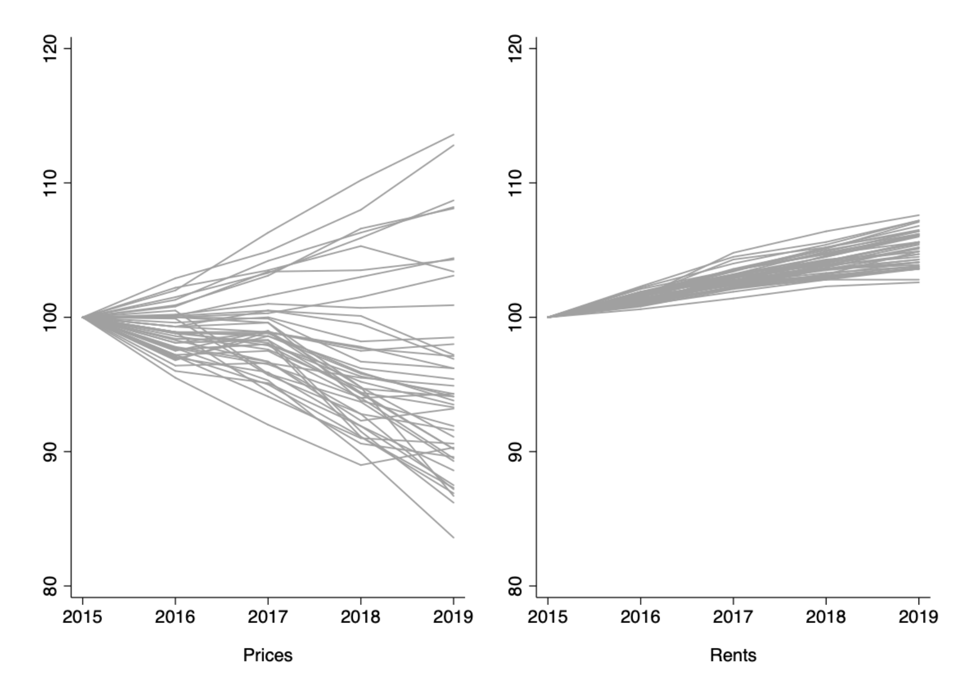


Figure A6: Statistics Finland official price and rent indices, 2015=100, 2015-2019, by regions.

Notes. Statistics Finland dwelling price index for apartments in non-new dwellings. Statistics Finland rents index for private market rental units. Regions displayed in the figures are: Espoo-Kauniainen, Etelä-Karjala, Etelä-Pohjanmaa, Etelä-Savo, Helsinki, Hyvinkää, Hämeenlinna, Joensuu, Jyväskylä, Järvenpää, Kainuu, Kajaani, Kanta-Häme, Kerava, Keski-Pohjanmaa, Keski-Suomi, Kokkola, Kotka, Kouvola, Kuopio, Kymenlaakso, Lahti, Lappeenranta, Lappi, Mikkeli, Oulu, Pirkanmaa, Pohjanmaa, Pohjois-Karjala, Pohjois-Pohjanmaa, Pohjois-Savo, Pori, Porvoo, Päijät-Häme, Rauma, Riihimäki, Rovaniemi, Satakunta, Seinäjoki, Tampere, Turku, Uusimaa, Vaasa, Vantaa, Varsinais-Suomi.

In order to validate my measure of rent and price divergence, I compare it to official Statistics Finland house price and rent indices, which are available from 2015 onwards. Figure A6 suggests that the Statistics Finland estimates suggest a very similar divergence story as my estimates[22]. Further evidence about regional divergence of *prices* is provided in Eerola et al. 2020.

[22]Statistics Finland price index is for old dwellings but the rents index could include new dwellings too.

References:

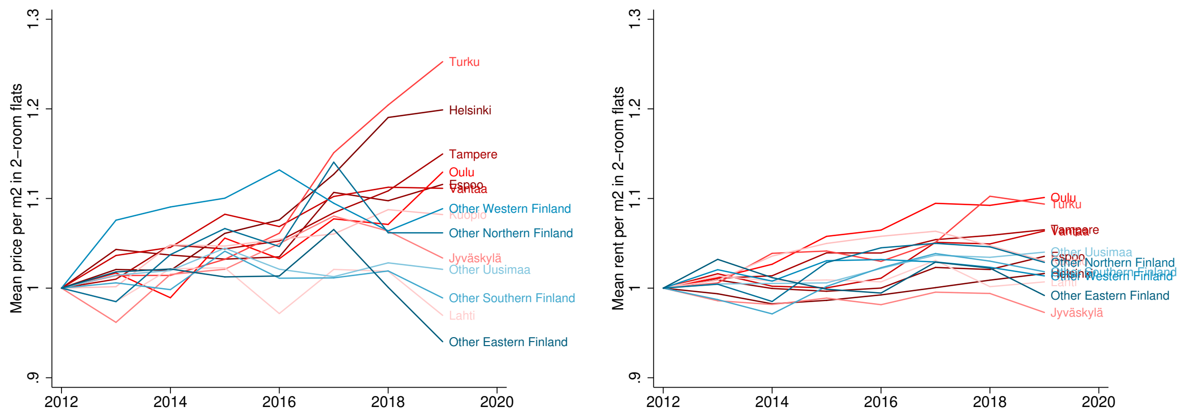
Official Statistics of Finland (OSF): Prices of dwellings in housing companies [e-publication]. ISSN=2323-8801. August 2021. Helsinki: Statistics Finland [referred: 6.10.2022]. Access method:
http://www.stat.fi/til/ashi/2021/08/ashi_2021_08_2021-09-29_tie_001_en.html
112l – Vanhojen osakeasuntojen hintaindeksi (2015=100) ja kauppojen lukumäärät, vuositasolla, 2015-2021.

Official Statistics of Finland (OSF): Rents of dwellings [e-publication]. ISSN=1798-1018. 2nd quarter 2021. Helsinki: Statistics Finland [referred: 6.10.2022]. Access method:
http://www.stat.fi/til/asvu/2021/02/asvu_2021_02_2021-08-05_tie_001_en.html
11x5 – Vuokraindeksi (2015=100) ja keskineliövuokrat, vuositiedot, 2015-2021

A.5 Indices in the main text

The price and rent indices reported in the main text in Figure 4 are comparable to the point estimates in the bottom panel of A3, with the following differences: *ii*) the regional classification in the main text consists of 14 regions (as opposed to 5 in the appendix), *ii*) the sample period used for the computation of the indices in the main text is limited to 2012-2019 and *iii*) the indices are real, not nominal (prices and rents are expressed in 2020 euros). However, the sample selection and the quality controls are otherwise similar.

For completeness, I report in Figure A7 the raw means or rents and prices per m^2, relative to the respective raw means in 2012, using the same sample, same regional classification and same (real) outcomes as Figure 4 in the main text. While Figure A7 is significantly noisier than 4, divergence patterns seem similar.



(**a**) Prices (**b**) Rents

Figure A7: Mean prices and rents per m^2 relative to 2012, by location.

Notes. Mean prices and rents per m^2 relative to 2012 in the same sample as what is used in the main text. No quality controls.

The figure in the introduction (Figure 1) is the same as the top panel in Figure A3.

B Other data sources

In this section of the appendix, I provide details on the data that are used in the empirical exercise. Data on house prices and rents are separately discussed in appendix A.2.

Geography In order to implement my empirical exercise in practice, I need to split Finland into geographically distinct regions. I provide here a list of the mapping of the "maakunta" regions to my 14-region classification of Finland. All data is obtained at the municipality level using the 2021 municipality classifications.

Eastern and Nothern Finland NUTS 2 region is split separately to Eastern (Pohjois-Karjala, Pohjois-Savo, Etelä-Savo) and Northern (Kainuu, Keski-Pohjanmaa, Pohjois-Pohjanmaa, Lappi). Other regions are mapped as: Uusimaa: Uusimaa. Western Finland: Keski-Suomi, Etelä-Pohjanmaa, Pohjanmaa, Satakunta, Pirkanmaa. Southern Finland: Etelä-Suomi, Varsinais-Suomi, Kanta-Häme, Päijät-Häme, Kymenlaakso, Etelä-Karjala. Åland islands will be excluded throughout the analysis.

The shapefile (map) of Finnish municipalities is obtained via Statistics Finland. The shapefile of country borders is obtained from Natural Earth.

References:

Municipality area boundaries, Statistics Finland. The material was downloaded from Statistics Finland's interface service on 4.8.2022 with the licence CC BY 4.0. https://tilastokeskuskartta.swgis.fi/#

Country borders. Natural Earth. https://www.naturalearthdata.com/

Regional population and migration To measure population at the municipality level, I use Statistics Finland Vital Statistics. To measure migration from one region to another, I use Statistics Finland matrices on intermunicipal migration. These data are publicly available on the Statistics Finland website. Migration probabilities are measured by relating the bilateral population flows from location i to location j during year t to the population in location i in year $t-1$. The probability of remaining in a location is calculated as $1 - \mathbb{P}(\text{migrate})$.

References:

Official Statistics of Finland: 11a1 – Intermunicipal migration by area of arrival and departure, 1990-2020 [referred 6.5.2022].

Official Statistics of Finland: 12au – Vital statistics and population by area, 1990-2020 [referred 6.5.2022].

New construction The data used on municipality-level information on the amount of new housing units constructed and their average sizes is administrative data from the Finnish Population Information System, publicly available via the Liiteri service.

Reference:

Väestötietojärjestelmä/Digi- ja väestötietovirasto: Uusien asuntojen määrä ja Uusien asuntojen keskikoko. Accessed 29.7.2022 via Liiteri, Syke:n Elinympäristön tietopalvelu.

Initial housing stock and housing consumption For my empirical exercise, I also need to measure the initial housing stock in each location. I recover the initial housing stock in each location by computing the total housing consumption in each location by multiplying the information on individual housing consumption in each location by the number of individuals in each location. The information on housing consumption by individual municipality is obtained from Statistics Finland household-dwelling statistics (computed from administrative data).

Reference:

Official Statistics of Finland: 115a – Household-dwelling units and housing population by housing density, 1990-2021, referred 24.6.2022

Robustness: Regional incomes I will use data on regional incomes in a robustness exercise in which I compare the model-consistent wages to the actual mean wages in each location. The information on municipality-level incomes is obtained via the Statistics Finland Income Distribution statistics (in part administrative, in part survey data).

Reference:

Official Statistics of Finland: 118w – Number, income and income structure of household-dwelling units by municipality, 1995-2020, referred 7.6.2022.

Consumer price indices Rents, prices and incomes are deflated to 2020 euros using the CPI published by Statistics Finland.

Reference:

Official Statistics of Finland (OSF): Consumer price index [e-publication]. ISSN=1799-0254. Helsinki: Statistics Finland, referred: 29.9.2022. Access method: http://www.stat.fi/til/khi/index_en.html

Data for housing consumption share Housing and nonhousing consumption shares are chosen based on the Statistics Finland Household consumption expenditure survey. The share of housing and energy of all household consumption is 0.3057 in 2016.

Reference:

Official Statistics of Finland (OSF): Households' consumption [e-publication]. ISSN=2323-3028. Helsinki: Statistics Finland [referred: 29.9.2022]. Access method: http://www.stat.fi/til/ktutk/index_en.html

Data for institutional context The section "Institutional context" in the main text contains references to Statistics Finland.

Reference:

Statistics Finland, 115y – Household-dwelling units by tenure status, type of building, number of persons, 2020, referred 21.9.2022.

Data for age structure and number of workplaces by municipality The section 5.3 uses data on the age structure by location and on the number of jobs in different municipalities, obtained from Statistics Finland Muicipal Key Figures.

Reference:

Statistics Finland, Municipal Key Figures 1987-2021, referred 26.10.2022.

Data for the number of housing transactions Appendix section A.2 contains a reference to number of transactions recorded by Statistics Finland.

Reference:

Official Statistics of Finland (OSF): Prices of dwellings in housing companies [e-publication]. ISSN=2323-8801. Helsinki: Statistics Finland [referred: 5.10.2022]. Access method: http://www.stat.fi/til/ashi/index.html

Data for the number of Airbnb listings The introduction contains a reference to the number of short-term rental listings in some European cities. These figures are obtained via website "airdna.co", with written permission for use obtained on 28.10.2022.

References:

https://www.airdna.co/vacation-rental-data/app/dk/default/copenhagen/overview
https://www.airdna.co/vacation-rental-data/app/dk/default/helsinki/overview
https://www.airdna.co/vacation-rental-data/app/dk/default/lisbon/overview
Referred 28.10.2022.

C Numerical implementation of model inversion

In this section of the appendix, I discus matters related to the practical implementation of the inversion procedure.

C.1 Uniqueness

To implement the model inversion in practice for some period t, we can start by inverting the static market clearing equations, yiedling estimates for $w_{l,t}$ and $\overline{L}_{l,t}$ for all locations. However, we still need to simultanously estimate interest rates in all locations and amenities in all locations, resulting in $2L-1$ parameters to be recovered, denoted by $\Omega_t = \{\{A_{l,t}, i_{l,t}\}_{l=1}^{L-1}, i_{L,t}\}$, where amenities in one location are normalized. In the model, interest rates and amenities in all locations affect prices and rents in all locations and all future time periods. Therefore, to recover Ω_t, we need to solve simultaneously for a system of $2L-1$ equations

$$V_{l,t+1} - V_{k,t+1} = \sum_{s=1}^{\infty} \beta^s \log\Big(\frac{A_l}{A_k}\frac{w_l}{w_k}(\frac{r_{k,t+s}(\Omega_t)}{r_{l,t+s}(\Omega_t)})^{(1-\phi)}(\frac{\mu_{t+s}^{k,k}(\Omega_t)}{\mu_{t+s}^{l,l}(\Omega_t)})^{\eta}\Big) \qquad \forall l \neq k$$

$$p_{l,t} = \sum_{s=0}(\frac{1}{1+i_{l,t}})^s(1-\delta)^s r_{l,t+s}^N(\Omega_t) \qquad \forall l$$

Propositions 3.4 and 3.5 suggest that we can uniquely identify one of the unknowns in Ω_t at a time, if the other elements of Ω_t are known. It is of course possible that this uniqueness does not carry on to the case of the system of equations, since in the model, interest rates and amenities in all locations affect prices and rents in all locations. We cannot therefore rule out the possibility of multiple solutions to the system of equations, which would imply that there are multiple vectors of interest rates and amenities that allow rationalizing the data. However, this seems to be of little practical concern, as independent of starting values I always only recover a single solution.

C.2 Measuring values

While helpful, the assumption about the symmetry of migration costs is of course a strong one, since in this model the migration costs are not related only to geographic distance but also factors like culture. This symmetry assumption will also have direct implications for the empirical exercise. Since in equation 47 we are measuring differences in values between locations k and d using the migration flows from location l to k and l to d, the choices of l and d matter, if in reality the migration costs are not symmetric. In my empirical exercise, I will fix a single reference location d, and measure value differences as the average over all l, as given by

$$V_{k,t+1}(\Theta_t) - V_{d,t+1}(\Theta_t) = \frac{1}{L}\sum_{l=1}^{L}\left[\frac{\eta}{\beta}\Big(\ln(\mu_t^{l,k}) - \ln(\mu_t^{l,d})\Big) + \frac{1}{\beta}\Big(\tau^{l,k} - \tau^{l,d}\Big)\right] \tag{52}$$

This allows me to estimate the value differences with less noise than if I was only using migration flows from a single location.

C.3 Numerical algorithm

In this section, I propose a fast iterative algorithm for the simultaneous implementation of the model inversion.

In order to implement the inversion, we also need an algorithm to solve the model. The model can be solved by modifying the algorithm suggested by Caliendo et al. (2019) to correspond to the sets of equations in propositions 2.1 and 2.2.

Model solution algorithm for the baseline equilibrium To solve the baseline equilibrium, take as inputs the anticipated sequences of changes to the fundamentals, the anticipated sequence of discount rates, and the intial equilibrium allocation. The numerical solution starts by taking a convergent sequence of guesses for $\{\dot{u}_{l,s}^{(0)}\}_{s=1}^{T}\ \forall l$, which is similar to Caliendo et al. (2019) (we could also take a convergent sequence of guesses for changes in household allocations, like in Ahlfeldt et al. (2020)). Moreover, we also take a convergent sequence of guesses for changes in the housing stock: $\{\dot{Q}_{l,s}^{(0)}\}_{s=1}^{T}\ \forall l$.

Given the guess $\{\dot{u}_{l,s}^{(0)}\}_{s=1}^{T}\ \forall l$ and the initial migration probabilities, solve for a path of migration probabilities using equation 21. This gives guess sequences of household allocations $\{N_{l,s}^{(0)}\}_{s=1}^{T}\ \forall l$ using 23. Given the guess $\{\dot{Q}_{l,s}^{(0)}\}_{s=1}^{T}\ \forall l$, obtain directly guess housing stock sequences $\{H_{l,s}^{(0)}\}_{s=1}^{T}\ \forall l$ using 26. Using $\{N_{l,s}^{(0)}\}_{s=1}^{T}\ \forall l$ and $\{H_{l,s}^{(0)}\}_{s=1}^{T}\ \forall l$, compute market-clearing rents using 20 to get $\{r_{l,s}^{(0)}\}_{s=1}^{T}\ \forall l$. Solve backwards for the sequences of house prices from equation 28 using guess interest rates and $\{r_{l,s}^{(0)}\}_{s=1}^{T}\ \forall l$.

Solve backwards for an updated guess for $\dot{u}$ using 22, and denote the updated guess with $\{\dot{u}_{l,s}^{(1)}\}_{s=1}^{T}\ \forall l$. Solve forwards for an updated guess for $\dot{Q}$ using 24, and denote the updated guess with $\{\dot{Q}_{l,s}^{(1)}\}_{s=1}^{T}\ \forall l$.

Compute the distances between $\{\dot{u}_{l,s}^{(0)}\}_{s=1}^{T}\ \forall l$ and $\{\dot{u}_{l,s}^{(1)}\}_{s=1}^{T}\ \forall l$ as well as $\{\dot{Q}_{l,s}^{(0)}\}_{s=1}^{T}\ \forall l$ and $\{\dot{Q}_{l,s}^{(1)}\}_{s=1}^{T}\ \forall l$. Take them as new initial conditions and repeat until convergence.

Model solution algorithm for counterfactual equilibria The terminology used here is that in period z, there is a change in the anticipated sequences of economic fundamentals, and the equilibrium under these new beliefs is called a "counterfactual" equilibrium. We can solve for the counterfactual equilibrium whether the period-$z-1$ there was a change in the anticipated sequences of economic fundamentals or not. Call the sequential equilibrium under the period-$z-1$ anticipated sequences the "previous equilibrium".

To solve a counterfactual equilibrium from period z onwards, take as inputs the previous equilibrium from period z onwards, the anticipated convergent sequences of changes to economic fundamentals in the previous equilibrium, denoted by single primes, $\{\dot{A}'_{l,z+s}, \dot{w}'_{l,z+s}, \dot{\bar{L}}'_{l,z+s}\}_{s=0}^{s=T} \forall l$, the *new* anticipated convergent sequence of changes to economic fundamentals, denoted by double primes, $\{\dot{A}''_{l,z+s}, \dot{w}''_{l,z+s}, \dot{\bar{L}}''_{l,z+s}\}_{s=0}^{s=T} \forall l$, the sequences of discount rates in the previous equilibrium, denoted by $\{i'_{l,s}\}_{s=z}^{s=T}$ $\forall l$, and the sequences of discount rates in the *new* equilibrium, denoted by $\{i''_{l,s}\}_{s=z}^{s=T}$ $\forall l$.

Take a convergent guess sequence for $\{\hat{\dot{u}}^{(0)}_{l,s}\}_{s=z}^{T}$ $\forall l$, where the notation is that

$$\hat{\dot{u}}_{l,s} = \frac{\dot{u}''_{l,s}}{\dot{u}'_{l,s}} = \frac{u''_{l,s}/u''_{l,s-1}}{u'_{l,s}/u'_{l,s-1}}$$

Take also a convergent sequence of guesses for counterfactual changes in the housing stock: $\{\dot{Q}''^{(0)}_{l,s}\}_{s=z}^{T}$ $\forall l$.

Use the guess $\hat{\dot{u}}$ with equation 29 for periods $t > z$ and the corresponding equation in appendix E.2 for period $t = z$ to get migration probabilities consistent with the guess. This then gives household allocations $\{N''^{(0)}_{l,s}\}_{s=1}^{T}$ $\forall l$ using equation 31. Given the guess $\dot{Q}''$, get $\{H''^{(0)}_{l,s}\}_{s=1}^{T}$ $\forall l$ from equation 34. Using H'' and N'', solve for rents using 20. Solve backwards for the sequence of house prices from 36.

Solve backwards for an updated guess for $\hat{\dot{u}}$ using 30, and denote the updated guess with $\{\hat{\dot{u}}^{(1)}_{l,s}\}_{s=1}^{T}$ $\forall l$. Solve forwards for an updated guess for $\dot{Q}$ using 32, and denote the updated guess with $\{\dot{Q}''^{(1)}_{l,s}\}_{s=1}^{T}$ $\forall l$. Compare with the initial guesses, update and repeat until convergence.

Model inversion algorithm To invert the model, I take advantage of the fact that for a given vector of fundamentals, the solution of the sequential equilibrium is deterministic and fast to compute, and the solution implies values for today's prices and values. Thus, given a guess for the unknown fundamentals, we can solve the model, giving us values and prices consistent with the guess, which we can then compare to the data, and update our guesses for

The algorithm is similar to Ahlfeldt et al. (2020) in the sense that the sequential equilibrium needs to be solved for each guess of the unknown fundamentals. As opposed to Ahlfeldt et al. (2020), I take advantage of closed-form rules that can be used to update guesses for unknonwn fundamentals. This makes the model inversion for my application very fast to implement[23].

To implement the inversion suggested in section ... , assuming that we want to recover the news that arrive in period t, I propose the following algorithm:

1. Take a $L \times 1$ guess vector for interest rates and $(L-1) \times 1$ guess vector for location-sepcific amenities relative to a reference location k. Denote initial guesses as $\Omega^{(0)}$. Take as given migration costs, obtained as in 46. Take as given the solution to the sequential equilibrium under period-$t-1$ beliefs and the period-$t-1$ beliefs on economic fundamentals, Θ_{t-1}.

2. Take the market clearing equations for period t, and invert them to recover wages and land supplies in each location, $w_l \forall l$ and $\overline{L}_l \forall l$.

3. Compute the sequential counterfactual equilibrium under $w_l, \overline{L}_l \; \forall l$ and guess $\Omega^{(0)}$ using proposition 2.2.

4. Compute model-implied values and prices from 37 and 12.

5. Compute differences between model-implied values and p/r ratios and empirical counterparts.

6. If difference $>$ tolerance, update the guesses using

$$\frac{1}{1-\beta} \log \left[\frac{A^{(1)}_{l,t}}{A_{k,t}}\right] = \underbrace{V_{l,t+1} - V_{k,t+1} - \frac{1}{1-\beta} \log \frac{w_{l,t}}{w_{k,t}}}_{\text{observed}}$$
$$+ \underbrace{(1-\phi) \sum_{s=0}^{\infty} \beta^s \log \frac{r_{l,t+1+s}(\Omega^{(0)})}{r_{k,t+1+s}(\Omega^{(0)})} + \eta \sum_{s=0}^{\infty} \beta^s \log \frac{\mu^{l,l}_{t+1+s}(\Omega^{(0)})}{\mu^{k,k}_{t+1+s}(\Omega^{(0)})}}_{\text{implied by previous guess}}$$

for the amenities and

$$D^{(1)}_{l,t} = 1 - \underbrace{\frac{r^N_{l,t}}{p_{l,t}}}_{\text{observed}} - \underbrace{\frac{r^N_{l,t}}{p_{l,t}}}_{\text{observed}} \left[1 - D^{(0)}_{l,t}\right] \sum_{s=0}^{\infty} \left(D^{(0)}_{l,t}\right)^s \underbrace{g_{l,t}(\Omega^{(0)})}_{\text{implied by previous guess}}$$

[23] In my application, I do not need to restrict agents' choice sets, whereas Ahlfeldt et al. (2020) need to restrict the number fo moves that agents can consider for computational reasons. However, their model is more complex in terms of the number of locations and the number of household groups, so it is not possible to directly compare the efficiency of the two approaches.

for the discount rates, where the notation is

$$r_{l,t+s} = (1 + g_{l,t+s}) r_{l,t}$$
$$D_{l,t} = \frac{1-\delta}{1+i_{l,t}}$$

and the updated guesses are denoted as $x^{(1)}$.

7. Repeat steps 2-6 until convergence.

The equations used to update guesses To arrive to the first expression used in updating, start from 37

$$\begin{aligned} V_{l,t+1} - V_{k,t+1} =& \sum_{s=0}^{\infty} \beta^s \log \left(\frac{A_{l,t}}{A_{k,t}} \frac{w_{l,t}}{w_{k,t}} (\frac{r_{k,t+1+s}}{r_{l,t+1+s}})^{(1-\phi)} (\frac{\mu^{k,k}_{t+1+s}}{\mu^{l,l}_{t+1+s}})^{\eta} \right) \\ =& \frac{1}{1-\beta} \log \frac{A_{l,t}}{A_{k,t}} + \frac{1}{1-\beta} \log \frac{w_{l,t}}{w_{k,t}} - (1-\phi) \sum_{s=0}^{\infty} \beta^s \log \frac{r_{l,t+1+s}}{r_{k,t+1+s}} - \eta \sum_{s=0}^{\infty} \beta^s \log \frac{\mu^{l,l}_{t+1+s}}{\mu^{k,k}_{t+1+s}} \end{aligned}$$

Reorganising:

$$\frac{1}{1-\beta} \log \frac{A_{l,t}}{A_{k,t}} = V_{l,t+1} - V_{k,t+1} - \frac{1}{1-\beta} \log \frac{w_{l,t}}{w_{k,t}} + (1-\phi) \sum_{s=0}^{\infty} \beta^s \log \frac{r_{l,t+1+s}}{r_{k,t+1+s}} - \eta \sum_{s=0}^{\infty} \beta^s \log \frac{\mu^{l,l}_{t+1+s}}{\mu^{k,k}_{t+1+s}}$$

And this needs to hold when r and μ are functions of the unknown fundamental vector Ω.

For the second expression, start from equation 13, where $p_{l,t}$ is the observed house price:

$$p_{l,t} = \sum_{s=0}^{\infty} (\frac{1-\delta}{1+i_{l,t}})^s r^N_{l,t+s}$$

Divide both sides by $r^N_{l,t}$ and denote the discountor $\frac{1-\delta}{1+i_{l,t}} = D_{l,t}$ and the rental growth relative to period t $r^N_{l,t+s}/r^N_{l,t} = 1 + g_{l,s}$

$$\frac{p_{l,t}}{r^N_{l,t}} = \sum_{s=0}^{\infty} D^s_{l,t}(1+g_{l,s}) = \sum_{s=0}^{\infty} D^s_{l,t} 1 + \sum_{s=0}^{\infty} D^s_{l,t} g_{l,s} = \frac{1}{1-D_{l,t}} + \sum_{s=0}^{\infty} D^s_{l,t} g_{l,s}$$

since $D \leq 1$. Thus,

$$\begin{aligned}
\left[1 - D_{l,t}\right] \frac{p_{l,t}}{r^N_{l,t}} &= 1 + \left[1 - D_{l,t}\right] \sum_{s=0}^{\infty} D^s_{l,t} g_{l,s} \\
\left[1 - D_{l,t}\right] &= \frac{r^N_{l,t}}{p_{l,t}} + \frac{r^N_{l,t}}{p_{l,t}} \left[1 - D_{l,t}\right] \sum_{s=0}^{\infty} D^s_{l,t} g_{l,s} \\
D_{l,t} &= 1 - \frac{r^N_{l,t}}{p_{l,t}} - \frac{r^N_{l,t}}{p_{l,t}} \left[1 - D_{l,t}\right] \sum_{s=0}^{\infty} D^s_{l,t} g_{l,s}
\end{aligned}$$

Which again needs to hold when $g_{l,t}$ is a function of Ω. Then, using the definition of D,

$$D_{l,t} = \frac{1 - \delta}{1 + i_{l,t}} \quad \Rightarrow \quad i_{l,t} = \frac{1 - \delta}{D_{l,t}} - 1$$

D Model derivations

D.1 Households' sub-period problem

Within periods, households solve

$$\max_{c,h} \; \log A + \phi \log(c) + (1-\phi)\log(h)$$
$$\text{s. t. } w = c + rh$$

implying Lagrangian

$$\phi \log(c) + (1-\phi)\log(h) + \lambda(w - c - rh)$$

and the FOC's

$$\phi \frac{1}{c} = \lambda$$
$$(1-\phi)\frac{1}{h} = r\lambda$$

Combining

$$\phi \frac{1}{c} = (1-\phi)\frac{1}{rh} \Rightarrow \; c = \frac{\phi}{1-\phi} rh$$

Combining with the budget constraint $c = w - rh$

$$\frac{\phi}{1-\phi} rh = w - rh$$
$$\left[1 + \frac{\phi}{1-\phi}\right] rh = w$$
$$\frac{1}{1-\phi} rh = w$$
$$h = (1-\phi)\frac{w}{r}$$

And

$$c = \frac{\phi}{1-\phi} rh = \frac{\phi}{1-\phi} r(1-\phi)\frac{w}{r} = \phi w$$

Substituting into the utility function

$$\max_{c,h} \log A + \phi \log(c) + (1-\phi)\log(h) \tag{53}$$

$$= \log A + \phi \log(\phi w) + (1-\phi)\log((1-\phi)\frac{w}{r}) \tag{54}$$

$$= \log A + \phi\Big[\log\phi + \log w\Big] + (1-\phi)\Big[\log(1-\phi) + \log w - \log r\Big] \tag{55}$$

$$= \phi\log\phi + (1-\phi)\log\phi + \log A + \phi\log w + (1-\phi)\log w - (1-\phi)\log r \tag{56}$$

$$= \underbrace{\log(\phi)}_{\tilde{\phi}} + \log A + \phi\log w + (1-\phi)\log w - (1-\phi)\log r \tag{57}$$

$$= \tilde{\phi} + \log A + \log w - (1-\phi)\log r \tag{58}$$

D.2 Household value functions and migration probabilities

Households' idiosyncratic preference shocks ϵ follow a type 1 extreme value distribution with location parameter $-\gamma$ and scale parameter 1, where γ is the Euler's constant, so that the cumulative distribution function is given by $F(\epsilon) = \exp\big(-\exp(-\epsilon-\gamma)\big)$, the mean is zero and the variance is $\pi^2/6$.

To see that if in

$$v_{l,t}(\epsilon_{i,t}) = u(w_{l,t}, r_{l,t}, A_{l,t}) + \max_{d\in 1,\ldots,L}\left[\beta\mathbb{E}_\epsilon\big(v_{d,t+1}\big) - \tau^{l,d} + \eta\epsilon^d_{it}\right] \tag{59}$$

then

$$V_{l,t} = \mathbb{E}_\epsilon(v_{l,t}) = u(w_{l,t}, r_{l,t}, A_{l,t}) + \eta\Big[\log\sum_{k=1}^{L}\big(\exp(\beta V_{k,t+1} - \tau^{l,k})\big)^{1/\eta}\Big] \tag{60}$$

and

$$\mu_t^{k,d} = \frac{\big(\exp(\beta V_{d,t+1} - \tau^{k,d})\big)^{1/\eta}}{\sum_{l=1}^{L}\big(\exp(\beta V_{l,t+1} - \tau^{k,l})\big)^{1/\eta}} \tag{61}$$

See Caliendo et al. (2019).

D.3 Developer's problem

In a given location, for a given market price for new housing p^Q, the representative (competitive) development firm solves

$$\max_n p^Q Q - wn - p^L \overline{L}$$
$$\text{s.t. } Q = n^\gamma \overline{L}^{1-\gamma}$$

where p^L is the price of land, p^Q the price of new housing (paid by the landowner sector) and w the cost of labor n. Q is the amount of housing built. Housing is produced with a CRS technology using land and labor with a technology parameter γ.

The FOC writes

$$\gamma n^{\gamma-1} p^Q \overline{L}^{1-\gamma} = w$$

And the optimal labor usage

$$n = \Big(\frac{w}{\gamma p^Q}\Big)^{\frac{1}{\gamma-1}} \overline{L}$$

Substituting this into the production function, the supply of construction then depends on the house price, wages and land supply as follows

$$Q = \Big(\frac{\gamma p^Q}{w}\Big)^{\frac{\gamma}{1-\gamma}} \overline{L}$$

meaning that the elasticity of housing construction w.r.t. house prices is $\frac{\gamma}{1-\gamma}$, or conversely

$$p^Q = \frac{w}{\gamma}\Big(\frac{Q}{\overline{L}}\Big)^{\frac{1-\gamma}{\gamma}} \tag{62}$$

And from the zero profits condition, the land price must satisfy

$$p^Q Q - wn - p^L L = 0$$
$$p^L = (p^Q)^{\frac{1}{1-\gamma}} \big(\frac{w}{\gamma}\big)^{\frac{\gamma}{1-\gamma}} - w\big(\frac{w}{\gamma p^Q}\big)^{\frac{1}{\gamma-1}}$$

D.4 Landlord demand curve for buildings

Landlords solve

$$\max_{q_{l,t}\geq 0} \sum_{s=0}^{\infty} \rho_{l,t,t+s}\pi_{l,t+s} \tag{63}$$

$$\pi_{l,t} = r^N_{l,t} h_{l,t} - p^Q_{l,t} q_{l,t} \tag{64}$$

$$h_{l,t+1} = (1-\delta)h_{l,t} + q_{l,t} \tag{65}$$

Rewriting this in a recursive form, the value of being a landlord in location l in time t who currently has a housing stock of size h is

$$\begin{aligned} V(h_{l,t}) &= \max_{q_{l,t}} \; \pi_{l,t} + \rho_{l,t,t+1} V\big(h_{l,t+1}\big) \\ &= \max_{q_{l,t}} \; r^N_{l,t} h_{l,t} - p^Q_{l,t} q_{l,t} + \rho_{l,t,t+1} V\big(h_{l,t+1}\big) \\ h_{l,t+1} &= (1-\delta)h_{l,t} + q_{l,t} \end{aligned}$$

The first-order condition w.r.t. $q_{l,t}$ writes:

$$\begin{aligned} p^Q_{l,t} &= \rho_{l,t,t+1} V'\big(h_{l,t+1}\big) \frac{\partial h_{l,t+1}}{\partial q_{l,t}} \\ &= \rho_{l,t,t+1} V'\big(h_{l,t+1}\big) \end{aligned}$$

The envelope condition writes:

$$\begin{aligned} V'(h_{l,t}) &= r^N_{l,t} + \rho_{l,t,t+1} V'\big(h_{l,t+1}\big) \frac{\partial h_{l,t+1}}{\partial h_{l,t}} \\ &= r^N_{l,t} + \rho_{l,t,t+1}(1-\delta) V'\big(h_{l,t+1}\big) \end{aligned}$$

We notice that the optimality conditions do not depend on q. Guess and verify that the value of capital is linear in the amount of capital $V(h_{l,t}) = P_{l,t}h_{l,t}$ (like in Suzuki

(2021)). Substitute to the envelope condition:

$$\begin{aligned}
V'(h_{l,t}) &= r^N_{l,t} + \rho_{l,t,t+1}(1-\delta)V'\big(h_{l,t+1}\big) \\
P_{l,t} &= r^N_{l,t} + \rho_{l,t,t+1}(1-\delta)P_{l,t+1} \\
P_{l,t}h_{l,t} &= r^N_{l,t}h_{l,t} + \rho_{l,t,t+1}(1-\delta)P_{l,t+1}h_{l,t} \\
&= \pi_{l,t} + p^Q_{l,t}q_{l,t} + \rho_{l,t,t+1}(1-\delta)P_{l,t+1}h_{l,t} \\
&= \pi_{l,t} + p^Q_{l,t}q_{l,t} + \rho_{l,t,t+1}P_{l,t+1}(h_{l,t+1} - q_{l,t}) \\
&= \pi_{l,t} + p^Q_{l,t}q_{l,t} - \rho_{l,t,t+1}P_{l,t+1}q_{l,t} + \rho_{l,t,t+1}P_{l,t+1}h_{l,t+1} \\
&= \pi_{l,t} + \Big[p^Q_{l,t} - \rho_{l,t,t+1}P_{l,t+1}\Big]q_{l,t} + \rho_{l,t,t+1}P_{l,t+1}h_{l,t+1}
\end{aligned}$$

and $p^Q_{l,t} - \rho_{l,t,t+1}P_{l,t+1} = 0$ under the guess from the FOC. Hence,

$$P_{l,t}h_{l,t} = \pi_{l,t} + \rho_{l,t,t+1}P_{l,t+1}h_{l,t+1}$$

Then, we can verify that we recover the correct value functions, since the discount factor satisfies $\rho_{l,z,s} = \rho_{l,z,s-1} \cdot \rho_{l,s-1,s}$:

$$P_{l,t}h_{l,t} = \sum_{s=0}^{\infty} \rho_{l,t,t+s}\pi_{l,t+s} = V(h_{l,t})$$

confirming the guess. So indeed the value of housing capital is linear in the amount of capital $V(h_{l,t}) = P_{l,t}h_{l,t}$.

Rewriting the optimality conditions under $V(h_{l,t}) = P_{l,t}h_{l,t}$:

$$\begin{aligned}
p^Q_{l,t} &= \rho_{l,t,t+1}P_{l,t+1} \\
P_{l,t} &= r^N_{l,t} + \rho_{l,t,t+1}(1-\delta)P_{l,t+1}
\end{aligned}$$

Combine to get

$$P_{l,t} = r^N_{l,t} + (1-\delta)p^Q_{l,t}$$

Substitute back to FOC:

$$\begin{aligned}
p^Q_{l,t} &= \rho_{l,t,t+1}P_{l,t+1} && (66) \\
&= \rho_{l,t,t+1}r^N_{l,t+1} + \rho_{l,t,t+1}(1-\delta)p^Q_{l,t+1} && (67) \\
&= \sum_{s=1} \rho_{l,t,t+s}(1-\delta)^{s-1}r^N_{l,t+s} && (68)
\end{aligned}$$

And this gives p^Q, the price that landlords are willing to pay for *new* housing capital, which yields the first rent in the following period $t+1$. Note that q does not appear in the optimality conditions because landlords are indifferent between purchasing any quantity of housing in period t, as long as the equilibrium price p^Q is given by the net present value of future rents: The demand curve for housing structures is a flat line.

In equilibrium, landlords are indifferent between selling or purchasing more capital for a price that equals the unit value of capital $P_{l,t}$, the NPV of rents:

$$\begin{aligned} P_{l,t} &= r^N_{l,t} + (1-\delta)p^Q_{l,t} \\ &= r^N_{l,t} + (1-\delta)\sum_{s=1} \rho_{l,t,t+s}(1-\delta)^{s-1} r^N_{l,t+s} \\ &= \sum_{s=0} \rho_{l,t,t+s}(1-\delta)^s r^N_{l,t+s} \end{aligned}$$

Finally, there is a simple link between the price of *new* housing capital p^Q and the price of existing housing capital P:

$$\begin{aligned} p^Q_{l,t} &= \rho_{l,t,t+1} P_{l,t+1} \\ &= \rho_{l,t,t+1} \Big[\sum_{s=0} \rho_{l,t+1,t+1+s}(1-\delta)^s r^N_{l,t+1+s} \Big] \end{aligned}$$

which tells us that a landlord purchasing a unit of new capital in period t implies will get a unit of capital, with value P, in the next period.

D.5 Rental market clearing

Each household's housing consumption in location l in time t is $h_{l,t} = (1-\phi)\frac{w_{l,t}}{r_{l,t}}$ (from household FOC's). Equating housing demand $N_{l,t} \times h_{l,t}$ with housing supply $H_{l,t}$,

$$N_{l,t} \cdot (1-\phi)\frac{w_{l,t}}{r_{l,t}} = H_{l,t} \tag{69}$$

$$r_{l,t} = \frac{(1-\phi)N_{l,t}}{H_{l,t}} \tag{70}$$

D.6 Construction market clearing

In a given location, for a sequence of anticipated rents $\{r_{l,t+s}\}_{s=0}^{\infty}$ and a sequence of discount factors $\{\rho_{l,t,t+s}\}_{s=0}^{\infty}$, landlord demand curve for new construction is given by equation 66

$$p^Q_{t,l} = \sum_{s=1} \rho_{l,t,t+s}(1-\delta)^{s-1}\, r^N_{l,t+s} \tag{71}$$

Inverse housing supply is given by equation 62:

$$p^Q_{l,t} = \frac{w_{l,t}}{\gamma}\Big(\frac{Q_{l,t}}{\overline{L}_{l,t}}\Big)^{\frac{1-\gamma}{\gamma}} \tag{72}$$

Equating the inverse supply curve with the (flat) inverse demand, we find that the construction supply in location l in time t is equal to

$$Q_{l,t} = \Big(\frac{\gamma p^Q_{l,t}}{w_{l,t}}\Big)^{\frac{\gamma}{1-\gamma}}\overline{L}_{l,t}$$

where p^Q is given by 66.

E Proofs of propositions for model solution

This section contains the proofs of the so-called "dynamic hat algebra" propositions (model solution in time differences).

E.1 Proof of proposition 2.1

We show that given initial observed allocations, sequences of discount rates and sequences of anticipate changes to A, w and $\overline{L}$ in all locations, the sequential equilibrium is the solution to the set equations in proposition 2.1. The proof of proposition 2.1 follows closely the steps in Caliendo et al. (2019).

E.1.1 Static market clearing

The rental market clearing equation 20 implies tha conditional on N and H, the rent is given by

$$r_{l,t}(N_{l,t}, H_{l,t}) = \frac{(1-\phi)w_{l,t}N_{l,t}}{H_{l,t}} \tag{73}$$

Taking this condition for t and $t+1$, we get

$$\frac{r_{l,t+1}(N_{l,t+1}, H_{l,t+1})}{r_{l,t}(N_{l,t}, H_{l,t})} = \frac{w_{l,t+1}}{w_{l,t}}\frac{N_{l,t+1}}{N_{l,t}}\frac{1}{\frac{H_{l,t+1}}{H_{l,t}}} \tag{74}$$

Using the dot notation,

$$\dot{r}_{l,t+1} = \dot{w}_{l,t+1}\dot{N}_{l,t+1}\frac{1}{\dot{H}_{l,t+1}} \tag{75}$$

which is equation 27. Now, as in Caliendo et al. (2019), starting with $r_{l,t}$ and given changes $\dot{w}$, $\dot{L}$, $\dot{H}$, we can solve for $r_{l,t+1}$ even without knowing w in levels.

Similarly, taking the construction market clearing equation 19 in time differences, conditional on a price,

$$Q_{l,t} = (w_{l,t}^{-\frac{\gamma}{1-\gamma}})(\gamma\, p^Q_{l,t})^{\frac{\gamma}{1-\gamma}}\overline{L}_{l,t} \tag{76}$$

$$\frac{Q_{l,t+1}}{Q_{l,t}} = \frac{w_{l,t+1}}{w_{l,t}}^{-\frac{\gamma}{1-\gamma}}\frac{p^Q_{l,t+1}}{p^Q_{l,t}}^{\frac{\gamma}{1-\gamma}}\frac{\overline{L}_{l,t+1}}{\overline{L}_{l,t}} \tag{77}$$

$$\dot{Q}_{l,t+1} = \left(\frac{\dot{p}^Q_{l,t+1}}{\dot{w}_{l,t+1}}\right)^{\frac{\gamma}{1-\gamma}}\dot{\overline{L}}_{l,t+1} \tag{78}$$

This gives us equation 32. Thus, if $\dot{p}^{Q}_{l,t+1}$, $\dot{w}_{l,t+1}$ and $\dot{\overline{L}}_{l,t+1}$ are known, and $Q_{l,t}$ observed, we can recover $Q_{l,t+1}$ without knowing w or $\overline{L}$ in levels.

E.1.2 Migration decisions

Next we show that equation 6

$$\mu_t^{k,d} = \frac{\exp(\beta V_{d,t+1} - \tau^{k,d})^{1/\eta}}{\sum_{l=1}^{L} \exp(\beta V_{l,t+1} - \tau^{k,l})^{1/\eta}}$$

implies equation 21

$$\mu_{t+1}^{k,d} = \frac{\mu_t^{k,d} \cdot \dot{u}_{d,t+2}^{\beta/\eta}}{\sum_{l=1}^{L} \mu_t^{k,l} \cdot \dot{u}_{l,t+2}^{\beta/\eta}}$$

following the steps in Caliendo et al. (2019) (appendix 3, proof of proposition 2).

Start from equation 6 and take the time difference between periods t and $t-1$:

$$\frac{\mu_t^{k,d}}{\mu_{t-1}^{k,d}} = \frac{\frac{\exp\left(\beta V_{d,t+1} - \tau^{k,d}\right)^{1/\eta}}{\sum_{l=1}^{L} \exp\left(\beta V_{l,t+1} - \tau^{k,l}\right)^{1/\eta}}}{\frac{\exp\left(\beta V_{d,t} - \tau^{k,d}\right)^{1/\eta}}{\sum_{l=1}^{L} \exp\left(\beta V_{l,t} - \tau^{k,l}\right)^{1/\eta}}} = \frac{\frac{\exp\left(\beta V_{d,t+1} - \tau^{k,d}\right)^{1/\eta}}{\exp\left(\beta V_{d,t} - \tau^{k,d}\right)^{1/\eta}}}{\frac{\sum_{l=1}^{L} \exp\left(\beta V_{l,t+1} - \tau^{k,l}\right)^{1/\eta}}{\sum_{l=1}^{L} \exp\left(\beta V_{l,t} - \tau^{k,l}\right)^{1/\eta}}} \tag{79}$$

Rewrite the numerator as:

$$\begin{aligned}
\frac{\exp\left(\beta V_{d,t+1} - \tau^{k,d}\right)^{1/\eta}}{\exp\left(\beta V_{d,t} - \tau^{k,d}\right)^{1/\eta}} &= \frac{\exp\left(\frac{1}{\eta}\beta V_{d,t+1} - \frac{1}{\eta}\tau^{k,d}\right)}{\exp\left(\frac{1}{\eta}\beta V_{d,t} - \frac{1}{\eta}\tau^{k,d}\right)} \\
&= \exp\left[\left(\frac{1}{\eta}\beta V_{d,t+1} - \frac{1}{\eta}\tau^{k,d}\right) - \left(\frac{1}{\eta}\beta V_{d,t} - \frac{1}{\eta}\tau^{k,d}\right)\right] \\
&= \exp\left[\frac{\beta}{\eta} V_{d,t+1} - \frac{\beta}{\eta} V_{d,t}\right] = \left[\exp\left(V_{d,t+1} - V_{d,t}\right)\right]^{\frac{\beta}{\eta}}
\end{aligned}$$

And the denominator as:

$$\frac{\sum_{l=1}^{L} \exp\left(\beta V_{l,t+1} - \tau^{k,l}\right)^{1/\eta}}{\sum_{l=1}^{L} \exp\left(\beta V_{l,t} - \tau^{k,l}\right)^{1/\eta}} = \sum_{l=1}^{L} \mu_{t-1}^{k,l} \exp\left(V_{l,t+1} - V_{l,t}\right)^{\beta/\eta}$$

where we used equation 6 in $t-1$ (for the intermediate steps, see below for the derivation of the values).

Combining the numerator and the denominator,

$$\frac{\mu_t^{k,d}}{\mu_{t-1}^{k,d}} = \frac{\exp\left[V_{d,t+1} - V_{d,t}\right]^{\frac{\beta}{\eta}}}{\sum_{l=1}^{L} \mu_{t-1}^{k,l} \exp\left[V_{l,t+1} - V_{l,t}\right]^{\beta/\eta}} \tag{80}$$

Multiplying by $\mu_{t-1}^{k,d}$

$$\mu_t^{k,d} = \frac{\mu_{t-1}^{k,d} \exp\left[V_{d,t+1} - V_{d,t}\right]^{\frac{\beta}{\eta}}}{\sum_{l=1}^{L} \mu_{t-1}^{k,l} \exp\left[V_{l,t+1} - V_{l,t}\right]^{\beta/\eta}} \tag{81}$$

and using the definition of $\dot{u}$:

$$\mu_t^{k,d} = \frac{\mu_{t-1}^{k,d}\ \dot{u}_{d,t+1}^{\beta/\eta}}{\sum_{l=1}^{L} \mu_{t-1}^{k,l}\ \dot{u}_{l,t+1}^{\beta/\eta}} \tag{82}$$

which gives us equation 21 as desired.

In particular, setting $t = 0$, we get the first period changes in the baseline sequential equilibrium:

$$\mu_0^{k,d} = \frac{\mu_{-1}^{k,d}\ \dot{u}_{d,1}^{\beta/\eta}}{\sum_{l=1}^{L} \mu_{-1}^{k,l}\ \dot{u}_{l,1}^{\beta/\eta}} \tag{83}$$

and for any period $t > 0$ we can use 82 so with data on μ_{-1} and a known sequence $\{\dot{u}_s\}_{s=1}^{\infty}$, it is possible to generate the full sequence of μ's.

E.1.3 Indirect utility

Starting from the household indirect utility,

$$\begin{aligned} u(w_{l,t}, r_{l,t}, A_{l,t}) &= \log A_{l,t} + \log(w_{l,t}) - (1-\phi)\log r_{l,t} + \tilde{\phi} \\ &= \log A_{l,t} + \log(w_{l,t}) - (1-\phi)\log r_{l,t} + \log(\phi) \\ &= \log(\phi A_{l,t} w_{l,t} r_{l,t}^{-(1-\phi)}) \end{aligned}$$

Take $\exp$ to arrive to

$$\exp\left(u(w_{l,t}, r_{l,t}, A_{l,t})\right) = \phi A_{l,t} w_{l,t} r_{l,t}^{-(1-\phi)}$$

and denoting $\omega_{l,t} = \exp(u(A_{l,t}, w_{l,t}, r_{l,t}))$ and taking time differences, we get

$$\dot{\omega}_{l,t+1} = \dot{A}_{l,t+1}\dot{w}_{l,t+1}\dot{r}_{l,t+1}^{-(1-\phi)}$$

This is used as an input to equation 22.

E.1.4 Values

Next we show that equation 5

$$V_{l,t} = \mathbb{E}_{\epsilon}(v_{l,t}) = u(w_{l,t}, r_{l,t}, A_{l,t}) + \eta\big[\log\sum_{k=1}^{L}\exp(\beta V_{k,t+1} - \tau^{l,k})^{1/\eta}\big]$$

implies that

$$\dot{u}_{l,t+1} = \dot{\omega}_{l,t+1}\Big(\sum_{l=1}^{L}\mu_t^{k,l}\cdot\dot{u}_{l,t+2}^{\beta/\eta}\Big)^{\eta}$$

Again, the steps follow closely Caliendo et al. (2019).

Start from taking time differences $V_{l,t} - V_{l,t-1}$ for any period $t > 0$.

$$\begin{aligned}
V_{l,t} - V_{l,t-1} &= u(w_{l,t}, r_{l,t}, A_{l,t}) - u(w_{l,t-1}, r_{l,t-1}, A_{l,t-1}) \\
&\quad + \ \eta\big[\log\sum_{k=1}^{L}\exp(\beta V_{k,t+1} - \tau^{l,k})^{1/\eta}\big] - \eta\big[\log\sum_{k=1}^{L}\exp(\beta V_{k,t} - \tau^{l,k})^{1/\eta}\big] \\
&= u(w_{l,t}, r_{l,t}, A_{l,t}) - u(w_{l,t-1}, r_{l,t-1}, A_{l,t-1}) \\
&\quad + \ \eta\log\Big[\frac{\sum_{k=1}^{L}\exp(\beta V_{k,t+1} - \tau^{l,k})^{1/\eta}}{\sum_{k=1}^{L}\exp(\beta V_{k,t} - \tau^{l,k})^{1/\eta}}\Big]
\end{aligned}$$

Multiply and divide each term in the numerator by $\exp(\beta V_{k,t}-\tau^{l,k})^{1/\eta}$ to write

$$\eta\log\left[\frac{\sum_{k=1}^{L}\exp(\beta V_{k,t+1}-\tau^{l,k})^{1/\eta}}{\sum_{k=1}^{L}\exp(\beta V_{k,t}-\tau^{l,k})^{1/\eta}}\right]=\eta\log\left[\frac{\sum_{k=1}^{L}\exp(\beta V_{k,t+1}-\tau^{l,k})^{1/\eta}\frac{\exp(\beta V_{k,t}-\tau^{l,k})^{1/\eta}}{\exp(\beta V_{k,t}-\tau^{l,k})^{1/\eta}}}{\sum_{k=1}^{L}\exp(\beta V_{k,t}-\tau^{l,k})^{1/\eta}}\right]$$

$$=\eta\log\left[\frac{\sum_{k=1}^{L}\frac{\exp(\beta V_{k,t+1}-\tau^{l,k})^{1/\eta}}{\exp(\beta V_{k,t}-\tau^{l,k})^{1/\eta}}\frac{\exp(\beta V_{k,t}-\tau^{l,k})^{1/\eta}}{1}}{\sum_{k=1}^{L}\exp(\beta V_{k,t}-\tau^{l,k})^{1/\eta}}\right]=\eta\log\left[\frac{\sum_{k=1}^{L}\frac{\exp(V_{k,t+1}-V_{k,t})^{\beta/\eta}}{1}\frac{\exp(\beta V_{k,t}-\tau^{l,k})^{1/\eta}}{1}}{\sum_{k=1}^{L}\exp(\beta V_{k,t}-\tau^{l,k})^{1/\eta}}\right]$$

$$=\eta\log\left[\frac{\sum_{k=1}^{L}\exp(V_{k,t+1}-V_{k,t})^{\beta/\eta}\ \exp(\beta V_{k,t}-\tau^{l,k})^{1/\eta}}{\sum_{d=1}^{L}\exp(\beta V_{d,t}-\tau^{l,d})^{1/\eta}}\right]$$

$$=\eta\log\left[\sum_{k=1}^{L}\left[\exp(V_{k,t+1}-V_{k,t})^{\beta/\eta}\ \frac{\exp(\beta V_{k,t}-\tau^{l,k})^{1/\eta}}{\sum_{d=1}^{L}\exp(\beta V_{d,t}-\tau^{l,d})^{1/\eta}}\right]\right]$$

$$=\eta\log\left[\sum_{k=1}^{L}\left[\exp(V_{k,t+1}-V_{k,t})^{\beta/\eta}\ \mu_{t-1}^{l,k}\right]\right]$$

using $\mu_{t-1}^{l,k}=\frac{\exp(\beta V_{k,t}-\tau^{l,k})^{1/\eta}}{\sum_{d=1}^{L}\exp(\beta V_{d,t}-\tau^{l,d})^{1/\eta}}$. Therefore

$$V_{l,t}-V_{l,t-1}=u(w_{l,t},r_{l,t},A_{l,t})-u(w_{l,t-1},r_{l,t-1},A_{l,t-1})\ +\ \eta\log\left[\sum_{k=1}^{L}\left[\mu_{t-1}^{l,k}\ \exp(V_{k,t+1}-V_{k,t})^{\beta/\eta}\right]\right]$$

Taking $\exp$ on both sides

$$\begin{aligned}\exp(V_{l,t}-V_{l,t-1})=&\frac{\exp(u(w_{l,t},r_{l,t},A_{l,t}))}{\exp(u(w_{l,t-1},r_{l,t-1},A_{l,t-1}))}*\exp\left(\log\left[\sum_{k=1}^{L}\left[\mu_{t-1}^{l,k}\ \exp(V_{k,t+1}-V_{k,t})^{\beta/\eta}\right]\right]^{\eta}\right)\\=&\frac{\exp(\log(\phi A_{l,t}w_{l,t}r_{l,t}^{-(1-\phi)})))}{\exp(\log(\phi A_{l,t-1}w_{l,t-1}r_{l,t-1}^{-(1-\phi)}))}*\left[\sum_{k=1}^{L}\left[\mu_{t-1}^{l,k}\ \exp(V_{k,t+1}-V_{k,t})^{\beta/\eta}\right]\right]^{\eta}\\=&\frac{A_{l,t}w_{l,t}r_{l,t}^{-(1-\phi)}}{A_{l,t-1}w_{l,t-1}r_{l,t-1}^{-(1-\phi)}}*\left[\sum_{k=1}^{L}\left[\mu_{t-1}^{l,k}\ \exp(V_{k,t+1}-V_{k,t})^{\beta/\eta}\right]\right]^{\eta}\end{aligned}$$

Finally, using the notations: $\dot{u}_{l,t}=\exp(V_{l,t}-V_{l,t-1})$ and $\dot{\omega}_t=\dot{A}_{l,t}\dot{w}_{l,t}\dot{r}_{l,t}^{-(1-\phi)}$:

$$\dot{u}_{l,t}=\dot{\omega}_t\cdot\left(\sum_{k=1}^{L}\left[\mu_{t-1}^{l,k}\ \dot{u}_{l,t+1}^{\beta/\eta}\right]\right)^{\eta}$$

as desired. This holds for all periods $t \geq 1$.

E.1.5 Laws of motion

Laws of motion for N and H are the same equations as in the equilibrium definition.

E.1.6 Prices

The price equation in the text, equation 12,

$$p_{l,t}^{Q} = \sum_{s=1}^{\infty} \rho_{l,t,t+s}(1-\delta)^{s-1} r_{l,t+s}^{N} \tag{84}$$

is equivalent to equation 28 in proposition 2.1,

$$p_{l,t}^{Q} = \rho_{l,t,t+1} r_{l,t+1}^{N} + \rho_{l,t,t+1} p_{l,t+1}^{Q}$$

by inputing $p_{l,t+1}^{Q}$.

E.2 Proof of proposition 2.2

This section contains the proof of propositon 2.2. Again, the proof follows Caliendo et al. (2019), but I highlight that we can solve the counterfactual equilibrium using proposition 2.2 even if there is new information arriving in period subsequent periods. Consider news that arrive in period z. The previous period, period $z-1$, could have been one where agents had, or had not, received new information, so that in $z-1$, they might (or might not) have the same beliefs about the sequences of economic fundamentals as in $z-2$.

Notation. Let us start by defining some notation. In order to highlight that we can solve for a counterfactual equilibrium in z, whether $z-1$ was a period in which news arrived or not, denote the equilibrium under period $z-1$ anticipated sequences with primes $'$, and the equilibrium under the new information that was received in period z with double primes $''$. By setting $''$ to $'$ and $'$ to no primes, we recover the equations in the original proposition. Denote differences between $\dot{x}''$ and $\dot{x}'$ by double-hats $\hat{\hat{x}}$. For example,

$$\dot{\mu}''_{t+1} = \frac{\mu''_{t+1}}{\mu''_t}$$

$$\dot{\mu}'_{t+1} = \frac{\mu'_{t+1}}{\mu'_t}$$

$$\hat{\hat{\mu}}_{t+1} = \frac{\dot{\mu}''_{t+1}}{\dot{\mu}'_{t+1}}$$

Refer to equilibrium under single primes as the "previous" equilibrium.

E.2.1 Market clearing equations

In the counterfactual economy, it needs to be that the rental market clears given the counterfactual values, denoted by primes:

$$r''_{l,t} = \frac{(1-\phi)w''_{l,t}N''_{l,t}}{H''_{l,t}} \tag{85}$$

Taking the time difference,

$$\dot{r}''_{l,t} = \frac{\dot{w}''_{l,t}\dot{N}''_{l,t}}{\dot{H}''_{l,t}} \tag{86}$$

which is equation 35. Similarly, the construction market needs to clear given the counterfactual values

$$Q''_{l,t} = ({w''}_{l,t}^{-\frac{\gamma}{1-\gamma}})\big(\gamma\, {p^{Q''}}_{l,t}\big)^{\frac{\gamma}{1-\gamma}} \overline{L}''_t \tag{87}$$

Equation 32 is this equation in time differences.

Note that these equations also hold for whichever period $t \geq z$: To get the values in the period in which news arrive, $\dot{x}''_z$, use that before the news the values were the same as in the "single-prime" equilibrium $x''_{z-1} = x'_{z-1}$ for all variables.

E.2.2 Migration and household values

Migration in t greater than z We start by showing that in periods $t > z$, it is true that

$$\mu''^{d,l}_{t} = \frac{\mu''^{d,k}_{t-1}\ \dot{\mu'}^{d,k}_{t}\ \hat{\dot{u}}^{\beta/\eta}_{k,t+1}}{\sum_{k=1}^{L} \mu''^{d,l}_{t-1}\ \dot{\mu'}^{d,l}_{t}\ \hat{\dot{u}}^{\beta/\eta}_{l,t+1}}$$

Start by noticing that equation 6

$$\mu^{k,d}_t = \frac{\exp(\beta V_{d,t+1} - \tau^{k,d}\big)^{1/\eta}}{\sum_{l=1}^{L} \exp(\beta V_{l,t+1} - \tau^{k,l})^{1/\eta}} \tag{88}$$

can be written as

$$\mu^{k,d}_t = \frac{\mu^{k,d}_{t-1}\ \dot{u}^{\beta/\eta}_{d,t+1}}{\sum_{l=1}^{L} \mu^{k,l}_{t-1}\ \dot{u}^{\beta/\eta}_{l,t+1}} \tag{89}$$

as shown in appendix E.1.2.

Notice that also in a counterfactual equilibrium, equation 6 is also true, meaning that the counterfactual migration in period t is driven by the counterfactual values in $t + 1$, (irrespective of what has happened before t), so under some counterfactual sequences $'$ (or $''$) we can also write

$$\mu'^{k,d}_{t} = \frac{\exp(\beta V'_{d,t+1} - \tau^{k,d}\big)^{1/\eta}}{\sum_{l=1}^{L} \exp(\beta V'_{l,t+1} - \tau^{k,l})^{1/\eta}} \tag{90}$$

Then, notice that if the news arrive in period z, then for any periods strictly *after* the arrival of the news $t > z$, we can replicate the derivation in E.1.2 to rewrite equation 90 as

$$\mu'^{d,l}_{t} = \frac{\mu'^{d,l}_{t-1}\ \dot{u}'^{\beta/\eta}_{l,t+1}}{\sum_{k=1}^{L} \mu'^{d,k}_{t-1}\ \dot{u}'^{\beta/\eta}_{k,t+1}} \tag{91}$$

However we cannot use this for period $t = z$, because we do not know μ'_{z-1} (the counterfactual migration in a period before the news were known). Therefore, for period $t = z$, see below.

Take 91 under $'$ and $''$ for any $t > z$

$$\mu'^{d,l}_{t} = \frac{\mu'^{d,l}_{t-1}\ \dot{u}'^{\beta/\eta}_{l,t+1}}{\sum_{k=1}^{L} \mu'^{d,k}_{t-1}\ \dot{u}'^{\beta/\eta}_{k,t+1}} \tag{92}$$

move the time indices one period forward, so that for any $t \geq z$, under $'$ or $''$

$$\mu'^{d,l}_{t+1} = \frac{\mu'^{d,l}_{t}\ \dot{u}'^{\beta/\eta}_{l,t}}{\sum_{k=1}^{L} \mu'^{d,k}_{t}\ \dot{u}'^{\beta/\eta}_{k,t}} \tag{93}$$

so 93 is true for $t \geq z$. Comparing the sequence $''$, which becomes known in z to a sequence $'$ which became known *before* z, we can use 93 under $'$ and $''$ for any $t \geq z$ to get migration strictly *after* the news. Divide to get

$$\frac{\mu''^{d,l}_{t+1}}{\mu'^{d,l}_{t+1}} = \frac{\dfrac{\mu''^{d,l}_{t}\ \dot{u}''^{\beta/\eta}_{l,t+2}}{\sum_{k=1}^{L} \mu''^{d,k}_{t}\ \dot{u}''^{\beta/\eta}_{k,t+2}}}{\dfrac{\mu'^{d,l}_{t}\ \dot{u}'^{\beta/\eta}_{l,t+2}}{\sum_{k=1}^{L} \mu'^{d,k}_{t}\ \dot{u}'^{\beta/\eta}_{k,t+2}}} \tag{94}$$

$$= \frac{\dfrac{\mu''^{d,l}_{t}\ \dot{u}''^{\beta/\eta}_{l,t+2}}{\mu'^{d,l}_{t}\ \dot{u}'^{\beta/\eta}_{l,t+2}}}{\dfrac{\sum_{k=1}^{L} \mu''^{d,k}_{t}\ \dot{u}''^{\beta/\eta}_{k,t+2}}{\sum_{k=1}^{L} \mu'^{d,k}_{t}\ \dot{u}'^{\beta/\eta}_{k,t+2}}} \tag{95}$$

Multiplying both sides with ${\mu'}_{t+1}^{d,l}$

$$
{\mu''}_{t+1}^{d,l} = \frac{{\mu''}_t^{d,l} \frac{{\mu'}_{t+1}^{d,l} \; \dot{u''}_{l,t+2}^{\beta/\eta}}{{\mu'}_t^{d,l} \; \dot{u'}_{l,t+2}^{\beta/\eta}}}{\frac{\sum_{k=1}^{L} {\mu''}_t^{d,k} \; \dot{u''}_{k,t+2}^{\beta/\eta}}{\sum_{k=1}^{L} {\mu'}_t^{d,k} \; \dot{u'}_{k,t+2}^{\beta/\eta}}}
$$

Replacing $\dot{\mu'}_{t+1}^{d,l}$ and $\hat{\hat{u}}_{l,t+2}$

$$
{\mu''}_{t+1}^{d,l} = \frac{{\mu''}_t^{d,l} \dot{\mu'}_{t+1}^{d,l} \hat{\hat{u}}_{l,t+2}^{\beta/\eta}}{\frac{\sum_{k=1}^{L} {\mu''}_t^{d,k} \; \dot{u''}_{k,t+2}^{\beta/\eta}}{\sum_{k=1}^{L} {\mu'}_t^{d,k} \; \dot{u'}_{k,t+2}^{\beta/\eta}}}
$$

Looking at the denominator:

$$
\frac{\sum_{k=1}^{L} {\mu''}_t^{d,k} \; \dot{u''}_{k,t+2}^{\beta/\eta}}{\sum_{k=1}^{L} {\mu'}_t^{d,k} \; \dot{u'}_{k,t+2}^{\beta/\eta}} = \sum_{h=1}^{L} \frac{{\mu''}_t^{d,h} \; \dot{u''}_{h,t+2}^{\beta/\eta}}{\sum_{k=1}^{L} {\mu'}_t^{d,k} \; \dot{u'}_{k,t+2}^{\beta/\eta}}
$$

Divide and multiply each term with ${\mu'}_t^{d,h} \; \dot{u'}_{h,t+2}^{\beta/\eta}$

$$
\begin{aligned}
&= \sum_{h=1}^{L} \frac{{\mu'}_t^{d,h} \; \dot{u'}_{h,t+2}^{\beta/\eta} \; {\mu''}_t^{d,h} \; \dot{u''}_{h,t+2}^{\beta/\eta}}{{\mu'}_t^{d,h} \; \dot{u'}_{h,t+2}^{\beta/\eta} \; \sum_{k=1}^{L} {\mu'}_t^{d,k} \; \dot{u'}_{k,t+2}^{\beta/\eta}} = \sum_{h=1}^{L} \left[\frac{{\mu''}_t^{d,h} \; \dot{u''}_{h,t+2}^{\beta/\eta} \;\; {\mu'}_t^{d,h} \; \dot{u'}_{h,t+2}^{\beta/\eta}}{{\mu'}_t^{d,h} \; \dot{u'}_{h,t+2}^{\beta/\eta} \;\; \sum_{k=1}^{L} {\mu'}_t^{d,k} \; \dot{u'}_{k,t+2}^{\beta/\eta}} \right] \\
&= \sum_{h=1}^{L} \left[\frac{{\mu''}_t^{d,h} \; \dot{u''}_{h,t+2}^{\beta/\eta} \; {\mu'}_{t+1}^{d,h}}{{\mu'}_t^{d,h} \; \dot{u'}_{h,t+2}^{\beta/\eta}} \right] = \sum_{h=1}^{L} \left[{\mu''}_t^{d,h} \; \hat{\hat{u}}_{h,t+2}^{\beta/\eta} \; \dot{\mu'}_{t+1}^{d,h} \right]
\end{aligned}
$$

since ${\mu'}_{t+1}^{d,h} = \frac{{\mu'}_t^{d,h} \; \dot{u'}_{h,t+2}^{\beta/\eta}}{\sum_{k=1}^{L} {\mu'}_t^{d,k} \; \dot{u'}_{k,t+2}^{\beta/\eta}}$ and replacing for the dot and hat notation.

Substituting the denominator back

$$
{\mu''}_{t+1}^{d,l} = \frac{{\mu''}_t^{d,l} \; \dot{\mu'}_{t+1}^{d,l} \; \hat{\hat{u}}_{l,t+2}^{\beta/\eta}}{\sum_{h=1}^{L} \left[{\mu''}_t^{d,h} \; \dot{\mu'}_{t+1}^{d,h} \; \hat{\hat{u}}_{h,t+2}^{\beta/\eta} \right]}
$$

we find the expression we wanted.

Household values in t greater than z Next we show that in periods $t > z$, it is true that

$$\hat{\hat{u}}_{l,t} = \hat{\hat{\omega}}_{l,t} \cdot \Big[\sum_{k=1}^{L} \mu''^{l,k}_{t-1} \dot{\mu}^{l,k}_{t} \ \hat{\hat{u}}^{\beta/\eta}_{k,t+1}\Big]^{\eta}$$

In section E.1 we showed that equation 5

$$V_{l,t} = \mathbb{E}_{\epsilon}(v_{l,t}) = u(w_{l,t}, r_{l,t}, A_{l,t}) + \eta\big[\log \sum_{k=1}^{L} \exp(\beta V_{k,t+1} - \tau^{l,k})^{1/\eta}\big]$$

implies that

$$\dot{u}_{l,t} = \dot{\omega}_{l,t}\Big(\sum_{l=1}^{L} \mu^{k,l}_{t-1} \cdot \dot{u}^{\beta/\eta}_{l,t+1}\Big)^{\eta}$$

for periods $t > 0$ in the baseline equilibrium. We can repeat similar steps for counterfactual equilibrium, denoted by $'$ (or $''$), but only for periods $t > z$ where z is the period in which the news arrive, since again, we do not have $\mu'^{k,l}_{z-1}$.

Take periods $t \geq z$ (notice change of timing), where

$$\dot{u''}_{l,t+1} = \dot{\omega''}_{l,t+1} \ \cdot \Big[\sum_{k=1}^{L} \ \mu''^{l,k}_{t} \ \dot{u''}^{\beta/\eta}_{k,t+2}\Big]^{\eta}$$

and divide by

$$\dot{u'}_{l,t+1} = \dot{\omega'}_{l,t+1} \ \cdot \Big[\sum_{k=1}^{L} \ \mu'^{l,k}_{t} \ \dot{u'}^{\beta/\eta}_{k,t+2}\Big]^{\eta}$$

to get

$$\frac{\dot{u''}_{l,t+1}}{\dot{u'}_{l,t+1}} = \frac{\dot{\omega''}_{l,t+1}}{\dot{\omega'}_{l,t+1}} \cdot \frac{\Big[\sum_{k=1}^{L} \ \mu''^{l,k}_{t} \ \dot{u''}^{\beta/\eta}_{k,t+2}\Big]^{\eta}}{\Big[\sum_{k=1}^{L} \ \mu'^{l,k}_{t} \ \dot{u'}^{\beta/\eta}_{k,t+2}\Big]^{\eta}}$$

$$\hat{\hat{u}}_{l,t+1} = \hat{\hat{\omega}}_{l,t+1} \frac{\Big[\sum_{k=1}^{L} \ \mu''^{l,k}_{t} \ \dot{u''}^{\beta/\eta}_{k,t+2}\Big]^{\eta}}{\Big[\sum_{k=1}^{L} \ \mu'^{l,k}_{t} \ \dot{u'}^{\beta/\eta}_{k,t+2}\Big]^{\eta}} = \hat{\hat{\omega}}_{l,t+1} \ \Big[\sum_{k=1}^{L} \frac{\mu''^{l,k}_{t} \ \dot{u''}^{\beta/\eta}_{k,t+2}}{\sum_{d=1}^{L} \ \mu'^{l,d}_{t} \ \dot{u'}^{\beta/\eta}_{d,t+2}}\Big]^{\eta}$$

Multiply and divide right-hand side by ${\mu'}^{l,k}_{t}\ \dot{u'}^{\beta/\eta}_{k,t+2}$

$$\hat{\hat{u}}_{l,t+1} = \hat{\hat{\omega}}_{l,t+1}\left[\sum_{k=1}^{L} \frac{{\mu'}^{l,k}_{t}\ \dot{u'}^{\beta/\eta}_{k,t+2}\ {\mu''}^{l,k}_{t}\ \dot{u''}^{\beta/\eta}_{k,t+2}}{{\mu'}^{l,k}_{t}\ \dot{u'}^{\beta/\eta}_{k,t+2}\ \sum\limits_{d=1}^{L}\ {\mu'}^{l,d}_{t}\ \dot{u'}^{\beta/\eta}_{d,t+2}}\right]^{\eta}$$

Use ${\mu'}^{l,k}_{t+1} = \frac{{\mu'}^{l,k}_{t}\ \dot{u'}^{\beta/\eta}_{k,t+2}}{\sum\limits_{d=1}^{L} {\mu'}^{l,d}_{t}\ \dot{u'}^{\beta/\eta}_{d,t+2}}$ and denote $\hat{\hat{u}}_{k,t+2}$ and $\dot{\mu'}^{l,k}_{t+1}$ to write

$$\begin{aligned}
\hat{\hat{u}}_{l,t+1} &= \hat{\hat{\omega}}_{l,t+1}\left[\sum_{k=1}^{L} \frac{{\mu''}^{l,k}_{t}\ \dot{u''}^{\beta/\eta}_{k,t+2}\quad {\mu'}^{l,k}_{t}\ \dot{u'}^{\beta/\eta}_{k,t+2}}{{\mu'}^{l,k}_{t}\ \dot{u'}^{\beta/\eta}_{k,t+2}\quad \sum\limits_{d=1}^{L}\ {\mu'}^{l,d}_{t}\ \dot{u'}^{\beta/\eta}_{d,t+2}}\right]^{\eta} \\
&= \hat{\hat{\omega}}_{l,t+1}\left[\sum_{k=1}^{L} \frac{{\mu''}^{l,k}_{t}\quad \hat{\hat{u}}^{\beta/\eta}_{k,t+2}\quad {\mu'}^{l,k}_{t+1}}{{\mu'}^{l,k}_{t}}\right]^{\eta} \\
&= \hat{\hat{\omega}}_{l,t+1}\left[\sum_{k=1}^{L} {\mu''}^{l,k}_{t}\ \frac{{\mu'}^{l,k}_{t+1}}{{\mu'}^{l,k}_{t}}\quad \hat{\hat{u}}^{\beta/\eta}_{k,t+2}\right]^{\eta} \\
&= \hat{\hat{\omega}}_{l,t+1}\left[\sum_{k=1}^{L} {\mu''}^{l,k}_{t}\ \dot{\mu'}^{l,k}_{t+1}\ \hat{\hat{u}}^{\beta/\eta}_{k,t+2}\right]^{\eta}
\end{aligned}$$

as desired, for all periods $t \geq z$. For $\hat{\hat{u}}_{l,z}$, see below.

Household values in t = z Next, consider the period in which the news $''$ arrive. As highlighted above, we cannot use equations

$$\hat{\hat{u}}_{l,t} = \hat{\hat{\omega}}_{l,t}\left[\sum_{k=1}^{L} {\mu''}^{l,k}_{t-1}\ \dot{\mu'}^{l,k}_{t}\ \hat{\hat{u}}^{\beta/\eta}_{k,t+1}\right]^{\eta}$$

$${\mu''}^{d,l}_{t} = \frac{{\mu''}^{d,l}_{t-1}\ \dot{\mu'}^{d,l}_{t}\ \hat{\hat{u}}^{\beta/\eta}_{l,t+1}}{\sum\limits_{h=1}^{L}\left[{\mu''}^{d,h}_{t-1}\ \dot{\mu'}^{d,h}_{t}\ \hat{\hat{u}}^{\beta/\eta}_{h,t+1}\right]}$$

for the period in which the news arrive, z, since ${\mu''}_{z-1}$ are not known (we do not know what would ahve been the migration behavior under the new information *before* the new information).

I derive the equilibrium conditions for $t = z$ similar to Caliendo et al. (2019) highlighting that: *i*) We do not need the migration probabilities in $z - 1$ to compute the equilibrium *ii*) it does not matter if the belief in $z - 1$ was the same as the belief in $z - 2$ or not. In other words, there can (or can not) be news that arrive in subsequent periods, and we can still compute the equilibrium using the strategy from Caliendo et al. (2019).

The new information that arrives in period z cannot affect any equilibriu outcomes before z. Therefore, in all periods before the news, $z-1$ and earlier, we denote (with a slight abuse of notation) that

$$\mu''_{z-1} = \mu'_{z-1}$$

since the news which did not arrive yet could not affect the equilibrium yet, and similarly for all other variables. With a similar logic, we also denote that

$$\hat{\hat{u}}_{l,z-1} = \frac{\dot{u}''_{l,z-1}}{\dot{u}'_{l,z-1}} = \frac{\exp(V''_{l,z-1} - V''_{l,z-2})}{\exp(V'_{l,z-1} - V'_{l,z-2})} = 1$$

which says that the information which arrives in period z does not affect the values $z-1$ and $z-2$, and similarly for all other periods $t < z$.

Like in Caliendo et al. (2019), we want to solve for values $\mu'^{d,k}_{z}$ and $\dot{u}'_{l,z}$, the period in which the news arrive, to be then able to apply equation 29 from there onwards.

We show that

$$\mu''^{d,k}_{z} = \frac{\theta^{d,k}_{z}\ \hat{\hat{u}}^{\beta/\eta}_{l,z+1}}{\sum_{k=1}^{L} \theta^{d,k}_{z}\ \hat{\hat{u}}^{\beta/\eta}_{k,z+1}}$$

and

$$\hat{\hat{u}}_{l,z} = \hat{\hat{\omega}}_{l,z}\Big(\sum_{k=1}^{L} \theta^{d,k}_{z}\ \hat{\hat{u}}^{\beta/\eta}_{k,z+1}\Big)^{\eta}$$

where

$$\theta^{d,k}_{z} = \mu'^{d,k}_{z}\ (\hat{\hat{u}}_{k,z})^{\beta/\eta}$$

so θ depends on the migration that would have taken place in z absent the new information, as well as the change in utility given the new information.

Start from equation 5 for period $z-1$, in the *previous* equilibrium (*before* the news in z arrived, so the single primes)

$$V'_{l,z-1} = u(w'_{l,z-1}, r'_{l,z-1}, A'_{l,z-1}) + \eta\big[\log \sum_{k=1}^{L} \exp(\beta V'_{k,z} - \tau^{l,k})^{1/\eta}\big]$$

Take exp and denote $\exp(V') = u'$

$$u'_{l,z-1} = A'_{l,z-1} \cdot w'_{l,z-1} \cdot {r'}_{l,z-1}^{-(1-\phi)} \;\cdot\; \big[\sum_{k=1}^{L} \exp(\beta V'_{k,z} - \tau^{l,k})^{1/\eta}\big]^{\eta}$$
$$= A'_{l,z-1} \cdot w'_{l,z-1} \cdot {r'}_{l,z-1}^{-(1-\phi)} \;\cdot\; \big[\sum_{k=1}^{L} {u'}_{k,z}^{\beta/\eta} \cdot \exp(\tau^{l,k})^{-1/\eta})\big]^{\eta}$$

Divide and multiply right-hand side by $({u''}_{k,z})^{\beta/\eta}$

$$u'_{l,z-1} = A'_{l,z-1} \cdot w'_{l,z-1} \cdot {r'}_{l,z-1}^{-(1-\phi)} \;\cdot\; \big[\sum_{k=1}^{L} {u'}_{k,z}^{\beta/\eta} \, \frac{({u''}_{k,z})^{\beta/\eta}}{({u''}_{k,z})^{\beta/\eta}} \;\cdot \exp(\tau^{l,k})^{-1/\eta})\big]^{\eta}$$
$$= A'_{l,z-1} \cdot w'_{l,z-1} \cdot {r'}_{l,z-1}^{-(1-\phi)} \;\cdot\; \big[\sum_{k=1}^{L} (\frac{{u'}_{k,z}}{{u''}_{k,z}})^{\beta/\eta} \, ({u''}_{k,z})^{\beta/\eta} \;\cdot \exp(\tau^{l,k})^{-1/\eta})\big]^{\eta}$$

And denote $(\frac{{u'}_{k,z}}{{u''}_{k,z}})^{\beta/\eta} = \phi_{k,z}$

$$u'_{l,z-1} = A'_{l,z-1} \cdot w'_{l,z-1} \cdot {r'}_{l,z-1}^{-(1-\phi)} \;\cdot\; \big[\sum_{k=1}^{L} \;\phi_{k,z} \, ({u''}_{k,z})^{\beta/\eta} \;\cdot \exp(\tau^{l,k})^{-1/\eta})\big]^{\eta}$$

Next, take equation 5 for period z, in the equilibrium *after* the news in z arrived, (so the double primes)

$$V''_{l,z} = u(w''_{l,z}, r''_{l,z}, A''_{l,z}) + \eta\big[\log \sum_{k=1}^{L} \exp(\beta V''_{k,z+1} - \tau^{l,k})^{1/\eta}\big]$$
$$u''_{l,z} = A''_{l,z} \cdot w''_{l,z} \cdot {r''}_{l,z}^{-(1-\phi)} \;\cdot\; \big[\sum_{k=1}^{L} {u''}_{k,z+1}^{\beta/\eta} \cdot \exp(\tau^{l,k})^{-1/\eta}\big]^{\eta}$$

And divide $u''_{l,z}$ by $u'_{l,z-1}$ to get

$$\frac{u''_{l,z}}{u'_{l,z-1}} = \frac{A''_{l,z} \cdot w''_{l,z} \cdot {r''}_{l,z}^{-(1-\phi)}}{A'_{l,z-1} \cdot w'_{l,z-1} \cdot {r'}_{l,z-1}^{-(1-\phi)}} \left(\frac{\sum_{k=1}^{L} {u''}_{k,z+1}^{\beta/\eta} \cdot \exp(\tau^{l,k})^{-1/\eta}}{\sum_{k=1}^{L} \;\phi_{k,z} \, ({u''}_{k,z})^{\beta/\eta} \;\cdot \exp(\tau^{l,k})^{-1/\eta}} \right)^{\eta}$$
$$= \frac{\omega''_{l,z}}{\omega'_{l,z-1}} \left(\frac{\sum_{k=1}^{L} {u''}_{k,z+1}^{\beta/\eta} \cdot \exp(\tau^{l,k})^{-1/\eta})}{\sum_{k=1}^{L} \;\phi_{k,z} \, ({u''}_{k,z})^{\beta/\eta} \;\cdot \exp(\tau^{l,k})^{-1/\eta}} \right)^{\eta}$$

Next, take equation 6 in period $z-1$, *before the news* (so single primes) :

$$\mu'^{k,d}_{z-1} = \frac{\exp\left(\beta V'_{d,z} - \tau^{k,d}\right)^{1/\eta}}{\sum\limits_{l=1}^{L} \exp(\beta V'_{l,z} - \tau^{k,l})^{1/\eta}} = \frac{(u'_{d,z})^{\beta/\eta} \exp(\tau^{k,d})^{-1/\eta}}{\sum\limits_{l=1}^{L} \exp(u'_{l,z})^{\beta/\eta} \exp(\tau^{k,l})^{-1/\eta}}$$

Divide and multiply with $(u''_{d,z})^{\beta/\eta}$

$$\begin{aligned}
\mu'^{k,d}_{z-1} &= \frac{(u'_{d,z})^{\beta/\eta} \exp(\tau^{k,d})^{-1/\eta}}{\sum\limits_{l=1}^{L} \exp(u'_{l,z})^{\beta/\eta} \exp(\tau^{k,l})^{-1/\eta}} \\
&= \frac{(\frac{u'_{d,z}}{u''_{d,z}})^{\beta/\eta} u''^{\beta/\eta}_{d,z} \exp(\tau^{k,d})^{-1/\eta}}{\sum\limits_{l=1}^{L} (\frac{u'_{l,z}}{u''_{l,z}})^{\beta/\eta} u''^{\beta/\eta}_{l,z} \exp(\tau^{k,l})^{-1/\eta}} \\
&= \frac{\phi_{d,z} u''^{\beta/\eta}_{d,z} \exp(\tau^{k,d})^{-1/\eta}}{\sum\limits_{l=1}^{L} \phi_{l,z} u''^{\beta/\eta}_{l,z} \exp(\tau^{k,l})^{-1/\eta}}
\end{aligned}$$

So we can write

$$\begin{aligned}
\frac{u''_{l,z}}{u'_{l,z-1}} &= \frac{\omega''_{l,z}}{\omega'_{l,z-1}} \left(\frac{\sum\limits_{k=1}^{L} u''^{\beta/\eta}_{k,z+1} \cdot \exp(\tau^{l,k})^{-1/\eta}}{\sum\limits_{k=1}^{L} \phi_{k,z} \ (u''_{k,z})^{\beta/\eta} \ \cdot \exp(\tau^{l,k})^{-1/\eta})} \right)^{\eta} \\
&= \frac{\omega''_{l,z}}{\omega'_{l,z-1}} \left(\sum_{k=1}^{L} \frac{u''^{\beta/\eta}_{k,z+1} \cdot \exp(\tau^{l,k})^{-1/\eta} \frac{\phi_{k,z}(u''_{k,z})^{\beta/\eta}}{\phi_{k,z}(u''_{k,z})^{\beta/\eta}}}{\sum\limits_{h=1}^{L} \phi_{h,z} \ (u''_{h,z})^{\beta/\eta} \ \cdot \exp(\tau^{l,h})^{-1/\eta}} \right)^{\eta} \\
&= \frac{\omega''_{l,z}}{\omega'_{l,z-1}} \left(\sum_{k=1}^{L} \frac{\phi_{k,z}(u''_{k,z})^{\beta/\eta} \exp(\tau^{l,k})^{-1/\eta}}{\sum\limits_{h=1}^{L} \phi_{h,z} \ (u''_{h,z})^{\beta/\eta} \ \cdot \exp(\tau^{l,h})^{-1/\eta}} \frac{u''^{\beta/\eta}_{k,z+1}}{\phi_{k,z}(u''_{k,z})^{\beta/\eta}} \right)^{\eta}
\end{aligned}$$

where we can substitute for $\mu'^{l,k}_{z-1}$

$$\frac{u''_{l,z}}{u'_{l,z-1}} = \frac{\omega''_{l,z}}{\omega'_{l,z-1}} \left(\sum_{k=1}^{L} \mu'^{l,k}_{z-1} \frac{u''^{\beta/\eta}_{k,z+1}}{\phi_{k,z}(u''_{k,z})^{\beta/\eta}} \right)^{\eta}$$

Now, use the notation $u''_{l,z-1} = u'_{l,z-1}$ and $\omega''_{l,z-1} = \omega'_{l,z-1}$ (from the timing assumption)

$$\frac{u''_{l,z}}{u''_{l,z-1}} = \frac{\omega''_{l,z}}{\omega''_{l,z-1}} \left(\sum_{k=1}^{L} \frac{{\mu'}^{l,k}_{z-1}}{\phi_{k,z}} \left(\frac{{u''}_{k,z+1}}{{u''}_{k,z}}\right)^{\beta/\eta} \right)^{\eta}$$

And use the dot notation to get

$$\dot{u''}_{l,z} = \dot{\omega''}_{l,z} \left(\sum_{k=1}^{L} \frac{{\mu'}^{l,k}_{z-1}}{\phi_{k,z}} (\dot{u''}_{k,z+1})^{\beta/\eta} \right)^{\eta} \tag{96}$$

Next, take equation 84 under the "single-prime" equilibrium for period z, in other words, what would have been the change in values if no news had arrived in z:

$$\dot{u'}_{l,z} = \dot{\omega'}_{z} \cdot \left(\sum_{k=1}^{L} \left[{\mu'}^{l,k}_{z-1} \ \dot{u'}^{\beta/\eta}_{l,z+1} \right] \right)^{\eta}$$

and divide 96 by this to get

$$\begin{aligned} \frac{\dot{u''}_{l,z}}{\dot{u'}_{l,z}} &= \frac{\dot{\omega''}_{l,z}}{\dot{\omega'}_{z}} \left(\frac{\sum_{k=1}^{L} \frac{{\mu'}^{l,k}_{z-1}}{\phi_{k,z}} (\dot{u''}_{k,z+1})^{\beta/\eta}}{\sum_{k=1}^{L} \left[{\mu'}^{l,k}_{z-1} \ \dot{u'}^{\beta/\eta}_{l,z+1} \right]} \right)^{\eta} \\ &= \frac{\dot{\omega''}_{l,z}}{\dot{\omega'}_{z}} \left(\sum_{k=1}^{L} \frac{\frac{{\mu'}^{l,k}_{z-1}}{\phi_{k,z}} (\dot{u''}_{k,z+1})^{\beta/\eta}}{\sum_{h=1}^{L} \left[{\mu'}^{l,h}_{z-1} \ \dot{u'}^{\beta/\eta}_{h,z+1} \right]} \right)^{\eta} \end{aligned}$$

Replace for the hat notation

$$\hat{\hat{u}}_{l,z} = \hat{\hat{\omega}}_{l,z} \left(\sum_{k=1}^{L} \frac{{\mu'}^{l,k}_{z-1}}{\phi_{k,z}} \frac{(\dot{u''}_{k,z+1})^{\beta/\eta}}{\sum_{h=1}^{L} \left[{\mu'}^{l,h}_{z-1} \ \dot{u'}^{\beta/\eta}_{h,z+1} \right]} \right)^{\eta}$$

Multiply and divide by $(\dot{u'}_{k,z+1})^{\beta/\eta}$

$$\hat{\hat{u}}_{l,z} = \hat{\hat{\omega}}_{l,z} \left(\sum_{k=1}^{L} \frac{(\dot{u''}_{k,z+1})^{\beta/\eta} / (\dot{u'}_{k,z+1})^{\beta/\eta}}{\phi_{k,z}} \frac{{\mu'}^{l,k}_{z-1} (\dot{u'}_{k,z+1})^{\beta/\eta}}{\sum_{h=1}^{L} \left[{\mu'}^{l,h}_{z-1} \ \dot{u'}^{\beta/\eta}_{h,z+1} \right]} \right)^{\eta}$$

And substitute in the migration in period z, had there been no news in z,

$$\mu'^{l,k}_{z} = \frac{\mu'^{l,k}_{z-1}(\dot{u}'_{k,z+1})^{\beta/\eta}}{\sum_{h=1}^{L}\left[\mu'^{l,h}_{z-1}\ \dot{u}'^{\beta/\eta}_{h,z+1}\right]}$$

from 91 to get

$$\begin{aligned}\hat{\hat{u}}_{l,z} &= \hat{\hat{\omega}}_{l,z}\left(\sum_{k=1}^{L}\frac{(\dot{u}''_{k,z+1})^{\beta/\eta}/(\dot{u}'_{k,z+1})^{\beta/\eta}}{\phi_{k,z}}\mu'^{l,k}_{z}\right)^{\eta} \\ &= \hat{\hat{\omega}}_{l,z}\left(\sum_{k=1}^{L}(\hat{\hat{u}}_{k,z+1})^{\beta/\eta}\frac{\mu'^{l,k}_{z}}{\phi_{k,z}}\right)^{\eta}\end{aligned}$$

Notice that $(\frac{u'_{k,z}}{u''_{k,z}})^{\beta/\eta} = \phi_{k,z}$ so that

$$\frac{\mu'^{l,k}_{z}}{\phi_{k,z}} = \mu'^{l,k}_{z}\cdot(\frac{u''_{k,z}}{u'_{k,z}})^{\beta/\eta}$$

Finally, notice that

$$\hat{\hat{u}}_{k,z} = \frac{\dot{u}''_{k,z}}{\dot{u}'_{k,z}} = \frac{u''_{k,z}/u''_{k,z-1}}{u'_{k,z}/u'_{k,z-1}} = \frac{u''_{k,z}}{u'_{k,z}}\cdot\frac{u''_{k,z-1}}{u'_{k,z-1}} = \frac{u''_{k,z}}{u'_{k,z}}$$

since by the timing assumption, news in z do not affect the values in $z-1$. Therefore,

$$\frac{\mu'^{l,k}_{z}}{\phi_{k,z}} = \mu'^{l,k}_{z}\cdot(\hat{\hat{u}}_{k,z})^{\beta/\eta}$$

Then, recalling the definition of θ: $\theta^{d,k}_{z} = \mu'^{d,k}_{z}\ (\hat{\hat{u}}_{k,z})^{\beta/\eta}$ we have that

$$\hat{\hat{u}}_{l,z} = \hat{\hat{\omega}}_{l,z}\left(\sum_{k=1}^{L}\theta^{d,k}_{z}(\hat{\hat{u}}_{k,z+1})^{\beta/\eta}\right)^{\eta}$$

as desired.

Migration in t = z Start from equation 6 in the equilibrium before the news in z, (so single primes), for period z:

$$\mu'^{k,d}_{z} = \frac{\exp(\beta V'_{d,z+1} - \tau^{k,d})^{1/\eta}}{\sum_{l=1}^{L}\exp(\beta V'_{l,z+1} - \tau^{k,l})^{1/\eta}} = \frac{(u'_{d,z+1})^{\beta/\eta}\exp(\tau^{k,d})^{-1/\eta}}{\sum_{l=1}^{L}(u'_{l,z+1})^{\beta/\eta}\exp(\tau^{k,l})^{-1/\eta}}$$

Take the same equation for the same period but under the new information (note that this holds also in z)

$$\mu''^{k,d}_{z} = \frac{\exp(\beta V''_{d,z+1} - \tau^{k,d})^{1/\eta}}{\sum_{l=1}^{L} \exp(\beta V''_{l,z+1} - \tau^{k,l})^{1/\eta}} = \frac{(u''_{d,z+1})^{\beta/\eta} \exp(\tau^{k,d})^{-1/\eta}}{\sum_{l=1}^{L} (u''_{l,z+1})^{\beta/\eta} \exp(\tau^{k,l})^{-1/\eta}}$$

Divide one by the other to get

$$\frac{\mu''^{k,d}_{z}}{\mu'^{k,d}_{z}} = \frac{\frac{(u''_{d,z+1})^{\beta/\eta} \exp(\tau^{k,d})^{-1/\eta}}{\sum_{l=1}^{L} (u''_{l,z+1})^{\beta/\eta} \exp(\tau^{k,l})^{-1/\eta}}}{\frac{(u'_{d,z+1})^{\beta/\eta} \exp(\tau^{k,d})^{-1/\eta}}{\sum_{l=1}^{L} (u'_{l,z+1})^{\beta/\eta} \exp(\tau^{k,l})^{-1/\eta}}} = \frac{\frac{(u''_{d,z+1})^{\beta/\eta} \exp(\tau^{k,d})^{-1/\eta}}{(u'_{d,z+1})^{\beta/\eta} \exp(\tau^{k,d})^{-1/\eta}}}{\frac{\sum_{l=1}^{L} (u''_{l,z+1})^{\beta/\eta} \exp(\tau^{k,l})^{-1/\eta}}{\sum_{l=1}^{L} (u'_{l,z+1})^{\beta/\eta} \exp(\tau^{k,l})^{-1/\eta}}}$$

$$= \frac{(u''_{d,z+1}/u'_{d,z+1})^{\beta/\eta}}{\sum_{l=1}^{L} \frac{(u''_{l,z+1})^{\beta/\eta} \exp(\tau^{k,l})^{-1/\eta}}{\sum_{h=1}^{L} (u'_{h,z+1})^{\beta/\eta} \exp(\tau^{k,h})^{-1/\eta}}} = \frac{(u''_{d,z+1}/u'_{d,z+1})^{\beta/\eta}}{\sum_{l=1}^{L} \frac{\frac{(u'_{l,z+1})^{\beta/\eta}}{(u'_{l,z+1})^{\beta/\eta}} (u''_{l,z+1})^{\beta/\eta} \exp(\tau^{k,l})^{-1/\eta}}{\sum_{h=1}^{L} (u'_{h,z+1})^{\beta/\eta} \exp(\tau^{k,h})^{-1/\eta}}}$$

$$= \frac{(u''_{d,z+1}/u'_{d,z+1})^{\beta/\eta}}{\sum_{l=1}^{L} \frac{\frac{(u''_{l,z+1})^{\beta/\eta}}{(u'_{l,z+1})^{\beta/\eta}} \mu'^{k,l}_{z}}{1}} = \frac{(u''_{d,z+1}/u'_{d,z+1})^{\beta/\eta}}{\sum_{l=1}^{L} (u''_{l,z+1}/u'_{l,z+1})^{\beta/\eta} \mu'^{k,l}_{z}}$$

Multiply and divide denominator by $(u''_{l,z}/u'_{l,z})^{\beta/\eta}$ to get

$$\frac{\mu''^{k,d}_{z}}{\mu'^{k,d}_{z}} = \frac{(u''_{d,z+1}/u'_{d,z+1})^{\beta/\eta}}{\sum_{l=1}^{L} (u''_{l,z+1}/u'_{l,z+1})^{\beta/\eta} \mu'^{k,l}_{z} \frac{(u''_{l,z}/u'_{l,z})^{\beta/\eta}}{(u''_{l,z}/u'_{l,z})^{\beta/\eta}}}$$

$$= \frac{(u''_{d,z+1}/u'_{d,z+1})^{\beta/\eta}}{\sum_{l=1}^{L} \mu'^{k,l}_{z} (u''_{l,z}/u'_{l,z})^{\beta/\eta} \frac{(u''_{l,z+1}/u'_{l,z+1})^{\beta/\eta}}{(u''_{l,z}/u'_{l,z})^{\beta/\eta}}}$$

$$= \frac{(u''_{d,z+1}/u'_{d,z+1})^{\beta/\eta}}{\sum_{l=1}^{L} \mu'^{k,l}_{z} (u''_{l,z}/u'_{l,z})^{\beta/\eta} \frac{(u''_{l,z+1}/u''_{l,z})^{\beta/\eta}}{(u'_{l,z+1}/u'_{l,z})^{\beta/\eta}}}$$

$$= \frac{(u''_{d,z+1}/u'_{d,z+1})^{\beta/\eta}}{\sum_{l=1}^{L} \mu'^{k,l}_{z} (u''_{l,z}/u'_{l,z})^{\beta/\eta} (\hat{\hat{u}}_{l,z})^{\beta/\eta}}$$

And substitute in the numerator $\hat{\hat{u}}_{d,z} \cdot {u''}_{d,z}/{u'}_{d,z} = {u''}_{d,z+1}/{u'}_{d,z+1}$

$$\frac{{\mu''}^{k,d}_{z}}{{\mu'}^{k,d}_{z}} = \frac{(\hat{\hat{u}}_{d,z} \cdot {u''}_{d,z}/{u'}_{d,z})^{\beta/\eta}}{\sum_{l=1}^{L} {\mu'}^{k,l}_{z}({u''}_{l,z}/{u'}_{l,z})^{\beta/\eta}(\hat{\hat{u}}_{l,z})^{\beta/\eta}}$$

Then, recalling the definition of θ: $\theta^{k,d}_{z} = {\mu'}^{k,d}_{z}\ (\hat{\hat{u}}_{d,z})^{\beta/\eta}$

$${\mu''}^{k,d}_{z} = \frac{{\mu'}^{k,d}_{z}(\hat{\hat{u}}_{d,z} \cdot {u''}_{d,z}/{u'}_{d,z})^{\beta/\eta}}{\sum_{l=1}^{L} {\mu'}^{k,l}_{z}({u''}_{l,z}/{u'}_{l,z})^{\beta/\eta}(\hat{\hat{u}}_{l,z})^{\beta/\eta}} = \frac{\theta^{k,d}_{z}({u''}_{d,z}/{u'}_{d,z})^{\beta/\eta}}{\sum_{l=1}^{L} \theta^{k,l}_{z}({u''}_{l,z}/{u'}_{l,z})^{\beta/\eta}}$$

And using once again that

$$\hat{\hat{u}}_{k,z} = \frac{{u''}_{k,z}}{{u'}_{k,z}}$$

we get

$${\mu''}^{k,d}_{z} = \frac{\theta^{k,d}_{z}(\hat{\hat{u}}_{d,z})^{\beta/\eta}}{\sum_{l=1}^{L} \theta^{k,l}_{z}(\hat{\hat{u}}_{l,z})^{\beta/\eta}}$$

which is what we wanted.

E.2.3 Prices

The price equation in the text, equation 12, also needs to hold in counterfactual equilibria,

$${p^{Q}}''_{l,t} = \sum_{s=1}^{\infty} \rho''_{l,t,t+s}(1-\delta)^{s-1}{r^{N}}''_{l,t+s} \tag{97}$$

And equation 36 is this one in a recursive form.

F Proofs of propositions for model inversion

F.1 Proof of proposition 3.1

This section contains the proof of proposition 3.1.

Assuming that migration costs are symmetric ($\tau^{l,k} = \tau^{k,l}$) and assuming $\tau^{l,l}$=0 $\forall t$ (own-migration cost is zero), we show that

$$\tau^{l,k} = \frac{1}{2} \log \Big[\Big(\frac{\mu_t^{k,d} \mu_t^{d,k}}{\mu_t^{k,k} \mu_t^{d,d}} \Big)^{-\eta} \Big]$$

Start from the probability to migrate from k to d, relative to the probability of staying in k

$$\frac{\mu_t^{k,d}}{\mu_t^{k,k}} = \frac{\frac{\exp\left(\beta V_{d,t+1} - \tau^{k,d}\right)^{1/\eta}}{\sum_{l=1}^{L} \exp(\beta V_{l,t+1} - \tau^{k,l})^{1/\eta}}}{\frac{\exp\left(\beta V_{k,t+1} - \tau^{k,k}\right)^{1/\eta}}{\sum_{l=1}^{L} \exp(\beta V_{l,t+1} - \tau^{k,l})^{1/\eta}}} = \frac{\exp\left(\beta V_{d,t+1} - \tau^{k,d}\right)^{1/\eta}}{\exp\left(\beta V_{k,t+1} - \tau^{k,k}\right)^{1/\eta}}$$

and the same equation for migration from d to k

$$\frac{\mu_t^{d,k}}{\mu_t^{d,d}} = \frac{\exp\left(\beta V_{k,t+1} - \tau^{d,k}\right)^{1/\eta}}{\exp\left(\beta V_{d,t+1} - \tau^{d,d}\right)^{1/\eta}}$$

Multiply to get

$$\begin{aligned}
\frac{\mu_t^{k,d}}{\mu_t^{k,k}} \frac{\mu_t^{d,k}}{\mu_t^{d,d}} &= \frac{\exp\left(\beta V_{d,t+1} - \tau^{k,d}\right)^{1/\eta}}{\exp\left(\beta V_{k,t+1} - \tau^{k,k}\right)^{1/\eta}} \frac{\exp\left(\beta V_{k,t+1} - \tau^{d,k}\right)^{1/\eta}}{\exp\left(\beta V_{d,t+1} - \tau^{d,d}\right)^{1/\eta}} \\
\Big(\frac{\mu_t^{k,d}}{\mu_t^{k,k}} \frac{\mu_t^{d,k}}{\mu_t^{d,d}} \Big)^{\eta} &= \frac{\exp\left(\beta V_{d,t+1} - \tau^{k,d}\right)}{\exp\left(\beta V_{k,t+1} - \tau^{k,k}\right)} \frac{\exp\left(\beta V_{k,t+1} - \tau^{d,k}\right)}{\exp\left(\beta V_{d,t+1} - \tau^{d,d}\right)} \\
&= \exp\Bigg(\Bigg[\beta V_{d,t+1} - \tau^{k,d}\Big) + \Big(\beta V_{k,t+1} - \tau^{d,k}\Big) \\
&\quad - \Big(\beta V_{k,t+1} - \tau^{k,k}\Big) - \Big(\beta V_{d,t+1} - \tau^{d,d}\Big) \Bigg] \\
&= \exp\Big[-\tau^{k,d} - \tau^{d,k} + \tau^{k,k} + \tau^{d,d} \Big]
\end{aligned}$$

Then, taking logs on both sides

$$\log\Big[\Big(\frac{\mu_t^{k,d}}{\mu_t^{k,k}}\frac{\mu_t^{d,k}}{\mu_t^{d,d}}\Big)^{\eta}\Big] = \Big[-\tau^{k,d}-\tau^{d,k}+\tau^{k,k}+\tau^{d,d}\Big]$$

$$\log\Big[\Big(\frac{\mu_t^{k,d}}{\mu_t^{k,k}}\frac{\mu_t^{d,k}}{\mu_t^{d,d}}\Big)^{-\eta}\Big] = \Big[\tau^{k,d}+\tau^{d,k}-\tau^{k,k}-\tau^{d,d}\Big]$$

$$= 2\tau^{k,d}$$

Where the last line uses the 2 assumptions on the migration costs (symmetry and own migration cost 0). Then we can write the migration cost as a function of the observed migration probabilities

$$\tau^{k,d} = \frac{1}{2}\log\Big[\Big(\frac{\mu_t^{k,d}}{\mu_t^{k,k}}\frac{\mu_t^{d,k}}{\mu_t^{d,d}}\Big)^{-\eta}\Big]$$

F.2 Proof of proposition 3.2

This section contains the proof of proposition 3.2.

From the model, we know that migration probabilities are given by

$$\mu_t^{k,d} = \frac{\exp(\beta V_{d,t+1}-\tau^{k,d})^{1/\eta}}{\sum_{l=1}^{L}\exp(\beta V_{l,t+1}-\tau_t^{k,l})^{1/\eta}}$$

We can rewrite this as

$$\mu_t^{k,d} = \frac{\exp[\frac{1}{\eta}(\beta V_{d,t+1}-\tau^{k,d})]}{\sum_{l=1}^{L}\exp[\frac{1}{\eta}(\beta V_{l,t+1}-\tau_t^{k,l})]}$$

$$\ln(\mu_t^{k,l}) = \frac{1}{\eta}(\beta V_{d,t+1}-\tau^{k,d}) - \ln\Big[\sum_{l=1}^{L}\exp[\frac{1}{\eta}(\beta V_{l,t+1}-\tau_t^{k,l})]\Big]$$

$$\ln(\mu_t^{l,k}) - \ln(\mu_t^{l,d}) = \frac{1}{\eta}(\beta V_{k,t+1}-\tau^{l,k}) - \frac{1}{\eta}(\beta V_{d,t+1}-\tau^{l,d})$$

$$V_{k,t+1} - V_{d,t+1} = \frac{\eta}{\beta}\Bigg[\ln(\mu_t^{l,k}) - \ln(\mu_t^{l,d})\Bigg] + \frac{1}{\beta}\Big[\tau^{l,k}-\tau^{l,d}\Big]$$

And the migration probabilities on the right-hand side are observed, and migration costs recovered in a previous step. Hence differences in the ex ante-values across locations l and d, $V_{k,t+1} - V_{d,t+1}$, can be recovered.

F.3 Proof of proposition 3.3

This section contains the proof of proposition 3.3 (recovering amenities if there are no news).

Starting from the expression for the migration probability:

$$\mu_t^{l,l} = \frac{\exp(\beta \mathbb{E}_t V_{l,t+1})^{1/\eta}}{\sum_{k=1}^{L} \exp(\beta \mathbb{E}_t V_{k,t+1} - \tau^{l,k})^{1/\eta}}$$

$$\Rightarrow \quad \sum_{k=1}^{L} \exp(\beta \mathbb{E}_t V_{k,t+1} - \tau^{l,k})^{1/\eta} = \frac{1}{\mu_t^{l,l}} * \exp(\beta \mathbb{E}_t V_{l,t+1})^{1/\eta}$$

We can write the ex-ante values:

$$\begin{aligned} V_{l,t} &= \mathbb{E}_\epsilon(v_{l,t}) = u(w_{l,t}, r_{l,t}, A_{l,t}) + \eta \big[\log \sum_{k=1}^{L} \exp(\beta \mathbb{E}_t V_{k,t+1} - \tau^{l,k})^{1/\eta} \big] \\ &= u(w_{l,t}, r_{l,t}, A_{l,t}) + \eta * \log \big[\frac{1}{\mu_t^{l,l}} * \exp(\beta \mathbb{E}_t V_{l,t+1})^{1/\eta} \big] \\ &= u(w_{l,t}, r_{l,t}, A_{l,t}) + \eta * \log \frac{1}{\mu_t^{l,l}} + \beta \mathbb{E}_t V_{l,t+1} \end{aligned}$$

If there are no unanticipated changes in the beginning of period t then $\mathbb{E}_t V_{k,t} - \mathbb{E}_t V_{d,t} = V_{k,t} - V_{d,t}$. Using migration probabilities in t and $t-1$, we can recover $V_{k,t} - V_{d,t}$ and $\mathbb{E}_t V_{k,t+1} - \mathbb{E}_t V_{d,t+1}$. From here can recover $u_{k,t} - u_{d,t}$ as

$$\begin{aligned} V_{k,t} - V_{d,t} &= u_{k,t} - u_{d,t} + \eta \big[\log \frac{1}{\mu_t^{k,k}} - \log \frac{1}{\mu_t^{d,d}} \big] + \beta \big[\mathbb{E}_t V_{k,t+1} - \mathbb{E}_t V_{d,t+1} \big] \\ u_{k,t} - u_{d,t} &= V_{k,t} - V_{d,t} - \eta \big[\log \frac{1}{\mu_t^{k,k}} - \log \frac{1}{\mu_t^{d,d}} \big] - \beta \big[V_{k,t+1} - V_{d,t+1} \big] \end{aligned}$$

Then, using the expression for the indirect utility in ..., from $u_{k,t} - u_{d,t}$ we can recover $\log A_{k,t} - \log A_{d,t}$ as

$$\begin{aligned} &\log A_{k,t} - \log A_{d,t} = u_{k,t} - u_{d,t} - \big[\log(w_{k,t}) - \log(w_{d,t}) \big] \; + \; (1-\phi) \big[\log r_{k,t} - r_{d,t} \big] \\ &\frac{A_{k,t}}{A_{d,t}} = \exp\Big(u_{k,t} - u_{d,t} - \big[\log(w_{k,t}) - \log(w_{d,t}) \big] \; + \; (1-\phi) \big[\log r_{k,t} - r_{d,t} \big] \Big) \end{aligned}$$

F.4 Identifying amenities in spatial models

This section highlights why amenity levels in models of discrete location choice are hard to interpret, and not always comparable across models. Suppose that the utility of giving in some location l is given by the location-specific amenities A_l, wages w_l and rents r_l. Amenities operate as location-specific utility shifters. This suggests that they can only be identified up to some normalisations. As is usual for discrete choice models, we need two types of normalizations: A normalization on the *location* and on the *scale* of utility.

F.4.1 Static model with no idiosyncratic preference shocks

Start by considering a model with homogenous agents and no migration costs. From spatial argitrage, utility, given by function u, has to be equalized across locations l:

$$u(A_l, w_l, r_l) = \overline{u} \qquad \forall\ l$$

where $\overline{u}$ is a reservation utility, constant across locations. If incomes and rents are observed ($\hat{w}_l$, $\hat{r}_l$), and the utility function is monotonely increasing in amenities, then there exists a mapping g s.t.

$$A_l = g(\hat{w}_l, \hat{r}_l, \overline{u})$$

For example, if the utility specification would be Cobb-Douglas (as in equation 3 in the main text), then

$$\begin{aligned}
&u(A_l, w_l, r_l) = \log(A_l w_l r_l^{-(1-\phi)}) = \log(A_l) + \log(w_l) - (1-\phi)\log(r_l) \\
&\log(A_l w_l r_l^{-(1-\phi)}) = \overline{u} \\
&A_l = \frac{\exp(\overline{u})}{w_l r_l^{-(1-\phi)}} = \exp(\overline{u})\frac{r_l^{1-\phi}}{w_l}
\end{aligned}$$

Thus, it is possible to obtain a location-specific amenity estimate by inverting the utility equation. The information that is used to identify amenities are the rents, the wages, and the reservation utility.

As usual, utility levels are not identified, so instead of identifying the reservation utility $\overline{u}$, it is set exogenously. Thus it operating as a "location normalization" for the utility function. Equivalently, one can normalize the level of amenities in a specific location. Setting the utility function as Cobb-Douglas implies a scale normalization for household utility: Utils are measured in log wages.

Interpreting the amenity estimates The implied amenity estimates are scale-invariant to the definition of a spatial units: If location l is redefined to consist of sublocations l_1 and l_2 s.t. wages and rents in l as well as in l_1 and l_2 are the same, then also the amenity estimates for the sublocations are the same as the original amenity estimate.

F.4.2 Static model with idiosyncratic preference shocks

Consider, on the other hand, a discrete-choice model like in Diamond (2016), where location choices are also affected by idiosyncratic preference shocks. Suppose utility of household i in location l is given by

$$u(A_l, w_l, r_l, \epsilon_{il}) = \delta(A_l, w_l, r_l) + \sigma\epsilon_{il}$$

where $\delta(A_l, w_l, r_l)$ is the part of utility that is common to all household and ϵ is an idiosyncratic location-specific preference shock which follows a type-1 extreme value distribution. From the distributional assumption it follows that the share of people choosing location l is given by

$$P(\text{choose location 1}) = s_l = \frac{\big(\exp\delta(A_l, w_l, r_l)\big)^{1/\sigma}}{\sum_{k=1}^{L}\big(\exp\delta(A_k, w_k, r_k)\big)^{1/\sigma}}$$

Denote some location d as a reference location.

$$\begin{aligned}\log(s_l) - \log(s_d) = \log\big(\frac{s_l}{s_d}\big) &= \log\Big(\frac{\big(\exp\delta(A_l, w_l, r_l)\big)^{1/\sigma}}{\big(\exp\delta(A_d, w_d, r_d)\big)^{1/\sigma}}\Big)\\ &= \frac{1}{\sigma}\Big[\delta(A_l, w_l, r_l) - \delta(A_d, w_d, r_d)\Big]\end{aligned}$$

Thus the mean utility of location l over the reference location is given by the number of people choosing location l relative to the number of people choosing the reference location d:

$$\delta(A_l, w_l, r_l) - \delta(A_d, w_d, r_d) = \sigma\big(\log(s_l) - \log(s_d)\big) = \sigma\big(\log\big(\frac{s_l}{s_d}\big)\big) = \sigma\log\Big(\frac{\frac{N_l}{N_{total}}}{\frac{N_d}{N_{total}}}\Big) = \sigma\log\Big(\frac{N_l}{N_d}\Big)$$

Now, assuming that w, r and N are observed, σ is known and the utility function δ is known and monotone in amenities A, then δ can be inverted to recover A as

$$A_l = k(\hat{w}_l, \hat{w}_d, \hat{r}_l, \hat{r}_d, \hat{N}_l, \hat{N}_d, \sigma, A_d)$$

where the level of amenities in the reference location, A_d, is normalized. Again, using a Cobb-Douglas utility function as an example,

$$
\begin{aligned}
&\delta(A_l, w_l, r_l) = \log(A_l w_l r_l^{-(1-\phi)}) \\
&\delta(A_l, w_l, r_l) - \delta(A_d, w_d, r_d) = \log\Big[\frac{A_l}{A_d} \cdot \frac{w_l}{w_d} \cdot (\frac{r_l}{r_d})^{-(1-\phi)}\Big] = \log\big((\frac{N_l}{N_d})^{\sigma}\big) \\
&A_l = (\frac{N_l}{N_d})^{\sigma} \cdot \frac{w_d}{w_l} \cdot (\frac{r_d}{r_l})^{-(1-\phi)} \cdot A_d
\end{aligned}
$$

As highlighted by Diamond (2016), if out of two otherwise similar locations, one has higher population, then it is interpreted to have a higher amenities, too.

A location normalization on utility is obtained by normalizing the level of amenities in the reference location (for example, setting $A_d = 1$). Equivalently, it is possible to normalize the mean utility δ in some location d. The scale normalization, again, comes from the functional form assumption on the utility function.

Interpreting the amenity estimates With idiosyncratic preference shocks, the measure of amenities is **not** neutral to location definitions. If location l is redefined to consist of to two equal-sized sublocations l_1 and l_2, then the amenity estimates are affected even if wages and rents in l as well as in l_1 and l_2 are the same:

$$
\begin{aligned}
&A_{l_1} = (\frac{N_{l_1}}{N_d})^{\sigma} \cdot \frac{w_d}{w_{l1}} \cdot (\frac{r_d}{r_{l1}})^{-(1-\phi)} \cdot A_d = (\frac{1}{2})^{\sigma} A_l \\
&A_{l_2} = (\frac{1}{2})^{\sigma} A_l
\end{aligned}
$$

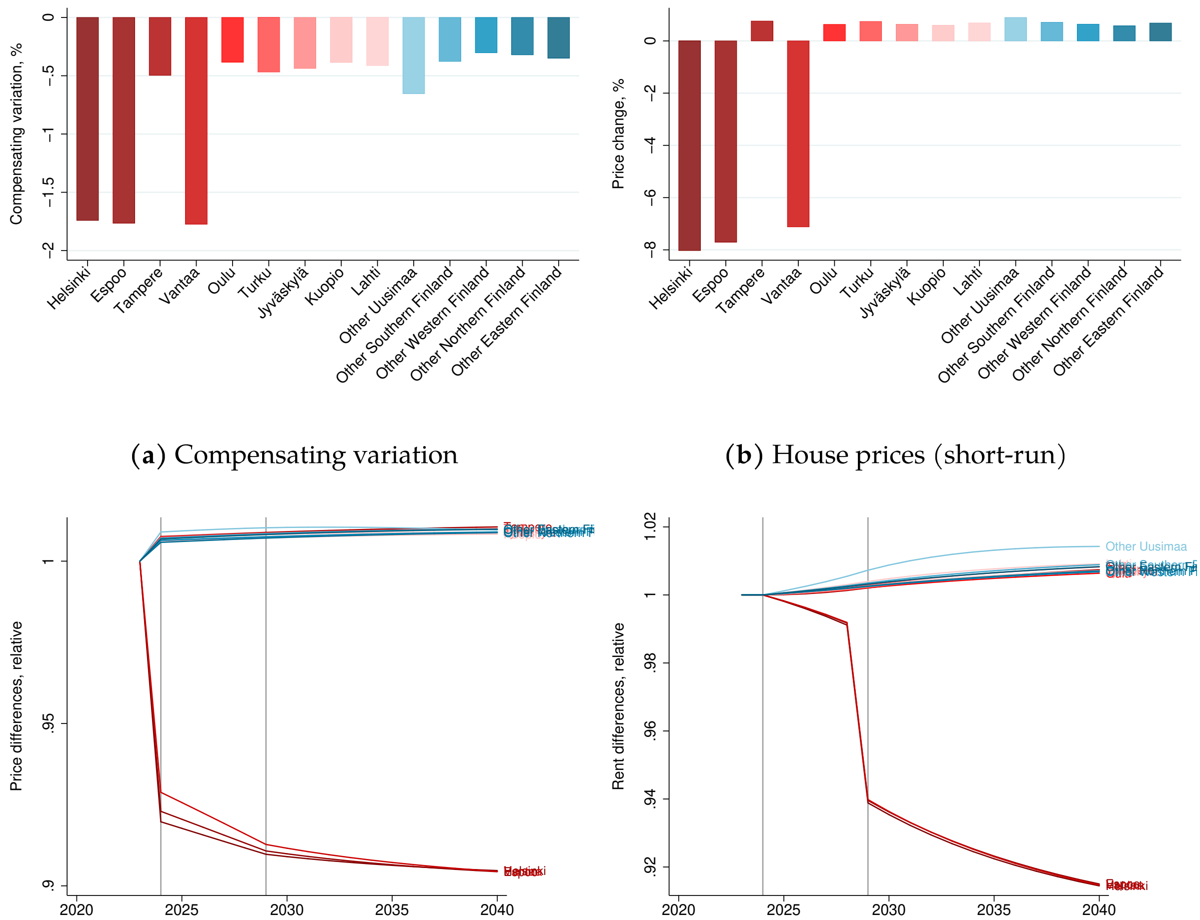
In other words, amenities recovered for location l are *not* a convex combination of the amenities recovered for locations l_1 and l_2. This suggests that amenity estimates obtained from this inversion are sensitive to how locations are defined.

Defining city size Because of the ambiguity related to amenity levels, I treat amenity levels as uninformative and focus on amenity changes over time, holding constant the definitions of geographic units. If one wanted to interpret amenities in levels, possible ways to overcome this challenge would be to *i*) make sure locations are defined in an economically meaningful ways, *ii*) define spatial units so that the population is always constant, or *iii*) normalise obtained amenity estimates to the city size that was used.

G Additional results

G.1 Counterfactual policy, negative wage effects

Figure A8 reports the results of the counterfactual experiment in scenario 2, in which there is a *negative* income shock in the capital region in 2029 which becomes public knowledge in 2024.



(**a**) Compensating variation

(**b**) House prices (short-run)

(**c**) Prices

(**d**) Rents

Figure A8: Counterfactual policy, scenario 2: A negative income shock in the capital region in 2029, which becomes public knowledge in 2024.

Notes. Compensating variation is the consumption-equivalent income increase for year 2024, relative to the equilibrium with no tunnel. Price change is measured in 2024, relative to prices in 2024 in the equilibrium with no tunnel. Prices and rents in the lower panels are reported each year relative to their values in the equilibrium with no tunnel.

G.2 Alternative counterfactual policy

This appendix section reports the results of an alternative counterfactual policy experiment in which the government tries to mitigate regional house price divergence by regional subsidies.

The regional divergence of house prices has caused concern in the public debate in Finland, in particular because of the regional divergence of house values, but also because it is a part of a broader divergence phenomenon across big cities, attracting households, and deprived regions, declining in population. A similar concern for regional divergence is widespread across developed countries, as are different policies aimed at reviving declining regions, so-called *base-placed policies* (Kline & Moretti, 2014). As they highlight, there can be a rationale for place-based policies due to equity concerns or due to efficiency concerns following market imperfections such as missing markets for insurance against regional shocks.

I study the effects of a place-based policy that takes the form of a regional wage subsidy or tax (similar to Kline & Moretti (2014)). I assume that in the beginning of 2013, unexpectedly, the government would have announced a regional redistribution policy such that the incomes of households in "Other Finland" (Other Uusimaa, Other Southern, Northern, Western and Eastern Finland) would have been supplemented by a proportional income subsidy of fraction λ_o and this subsidy would have been financed via a proportional income tax on city residents of rate λ_c. I consider an income subsidy of 5% (λ_c=0.05), which is financed via a tax in other locations of rate such that the policy is revenue-neutral in 2013[24]. The policy does not change labor productivity (thus labor costs to developers are not affected by the policy). Thus, household budget constraint would write

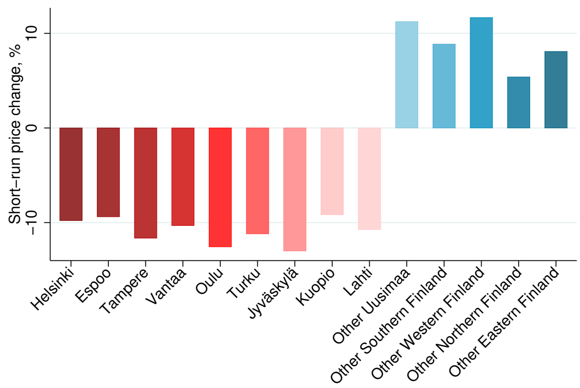
$$(1 + \lambda_l)w_{l,t} = c_{l,t} + r_{l,t}h_{l,t} \tag{98}$$

where $\lambda_l = \lambda_c$ for the cities and $\lambda_l = \lambda_o$ for other locations. To compute the new equilibrium under the regional wage subsidy, I assume that the location-specific economic fundamentals evolved otherwise as they did according to my estimates for 2013-2019.

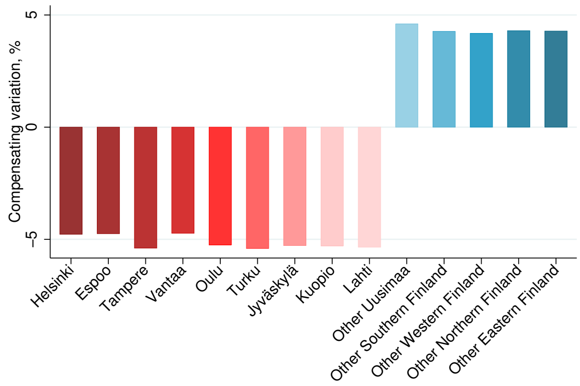
The tax rate needed to balance budget in 2013 is $\lambda_o = -0.072\%$[25]. Even if the policy is designed not to have effects on government budget in 2013, it soon becomes far from revenue-neutral. As soon as in 2014, the annual cost of the policy is already 30 million

[24]This figure, although large, is still modest relative to many of the existing policies. For example, Henkel et al. (2021) estimate that in Germany, fiscal transfers across local judiciaries range from a net contribution rate of 13.3 percent of local GDP in Frakfurt to a net benefiting rate of 23 % in some remote locations.

[25]In the beginning of 2013, the model-consistent population share of the 9 cities is 0.375 and 0.625 for other Finland. On the other hand, average incomes are higher in cities than in other Finland.



(**a**) House prices



(**b**) Renter welfare

Figure A9: The effects of a counterfactual policy experiment on short-run house prices and renter welfare.

Notes. The effects of a counterfactual policy experiment on renter welfare (measured in compensating variation) and house price change in 2013 relative to values if there was no base-placed policy.

euros.

The immediate effects of the place-based policy on renter and landlord welfare are displayed in the right panel of Figure A9. The subsidy on incomes in "Other Finland" immediately pushes up house prices in regions receiving the subsidy by orders of magnitudes of 5-10%. By design, the policy redistributes welfare from the big cities to the declining regions, and does mitigate the regional differences in house prices. However, if changes in economic fundamentals take place as estimated in section 5 form 2014 to 2019, the regional price patterns afterwards start soon reflecting a trend similar to Figure 4: prices, in particular in Eastern and Northern Finland, start to decline. Similarly, prices go down in cities as a response to the tax, but start soon increasing, in particular in the big cities.