Credit Constraints, Learning, and Spatial Misallocation

Elena Aguilar*

Princeton University

January 7, 2026

Updated regularly, please click here for the latest version.

Abstract

Human capital is a critical determinant of productivity in modern economies, and we have long understood credit market frictions to be a critical barrier for its accumulation. In this paper, I study location decisions as a particular form of human capital investment, where individuals trade long-term benefits for high upfront costs, most notably in the form of housing. I show that when locations differ in the learning opportunities they offer and agents are heterogeneous in their learning ability, credit frictions not only weaken positive sorting of learning ability across space but, under empirically relevant conditions, they will induce *negative* sorting among individuals that are credit constrained. That is, marginally better learners will optimally choose to reside in locations offering worse learning opportunities. I document the key mechanisms of the theory relying on a novel source of administrative data from Spain. I then build a dynamic heterogeneous agent spatial model to quantify the losses associated with the effect of credit frictions on the spatial distribution of labor. Importantly, these losses arise from distortions in the *composition* of skill in each city: the dominant source of inefficiency is the spatial misallocation of individuals with high learning-ability. In the presence of negative sorting, standard place-based policies strictly aimed to expand the size of productive cities may have limited effects, making it important to design policy that can better target the composition of heterogeneous workers across space.

*Email: elena.aguilar@princeton.edu. I thank my advisors, Gene Grossman and Richard Rogerson, for their help and guidance. I am also grateful to John S. Becko, Christine Blandhol, Thomas Bourany, Bo Honoré, Edward Jee, Michael Jenuwine, Hugo Lhuillier, Enric Martorell, Eduardo Morales, George Nikolakoudis, Steve Redding, Jordan Rosenthal-Kay, Esteban Rossi-Hansberg, Carolyn Tsao, Gianluca Violante, and Derek Wenning for their comments. This research benefited from financial support from the International Economics Section and the Simpson Center for Macroeconomics at Princeton University, the Fundación Ramón Areces and "la Caixa" Foundation. All errors are my own.

1 Introduction

Human capital is a critical determinant of productivity in modern economies. Understanding the forces that shape its accumulation is thus a central goal of economic research. Credit market frictions have long been recognized as a critical barrier. Because the returns to human capital investments are not realized immediately but often accrue over time, borrowing constraints may distort early-life educational decisions and continue to influence on-the-job accumulation later in the life cycle.

This paper studies a particular form of human capital investment: the decision of *where* to live. A growing body of evidence shows that locations differ substantially in the opportunities they provide for on-the-job skill development (De La Roca and Puga, 2017; Crews, 2024; Lhuillier, 2024). Cities like Madrid, New York, or Paris offer access to high-growth environments, better peer networks, and other learning opportunities not easily available in other parts of their respective countries. However, these long-term benefits are often associated with high upfront costs, most notably in the form of housing. This makes location choice an investment decision. In this paper, I study how credit constraints can distort location decisions and, consequently, the aggregate accumulation of human capital in the economy.

The analysis throughout the paper builds on a shared framework with four key features: opportunities for on-the-job human capital accumulation vary across space, access to a location requires the purchase of housing services, housing supply is inelastic in each location, and households face borrowing constraints. In this environment, I study the location decisions of individuals heterogeneous in their initial resources and their ability to learn. Aggregate human capital accumulation depends on how individuals are distributed across locations, being highest when those with higher learning ability locate in places with better opportunities for human capital accumulation. In equilibrium, however, workers' willingness to pay for better learning opportunities bids up housing prices in high-opportunity locations. In the presence of credit constraints, this creates a wedge between efficient and realized location choices, leading to inefficiently low levels of aggregate human capital accumulation.

My analysis proceeds in three steps. In the first part of the paper, I develop a tractable two-period version of this framework that allows me to fully characterize sorting in equilibrium. Then, using a novel source of administrative data combining information on wealth, income, and complete working histories, I present evidence on the key mechanisms of the model. To close the paper, I build and calibrate a dynamic quantitative spatial model with heterogeneous agents to understand the magnitude of these forces and their consequences.

The key contribution of the first part of the paper is to show theoretically that when individuals are heterogeneous in their learning ability, credit constraints will generally weaken the positive sorting of ability across space. Moreover, they will induce *negative* sorting among constrained individuals under empirically relevant parameterizations. That is, everything else held constant, credit constrained agents that are marginally better learners will optimally choose to reside in locations offering *worse* opportunities for on-the-job human capital accumulation.

This result reflects the net effect of two opposing forces, and is driven by the investment nature of location decisions. In the model, the supermodularity of the learning technology implies that agents with higher learning ability obtain higher returns from locations offering better opportunities for on-the-job human capital accumulation. This encourages better learners to choose better learning locations. On the other hand, however, high-ability agents anticipate higher future income, which increases the relative value of current consumption. Having exhausted their borrowing opportunities, constrained agents can only increase current consumption by saving on housing costs. This force leads them to choose cheaper locations offering worse learning opportunities in equilibrium. I show that when preferences exhibit strong consumption smoothing motives–more precisely, when the inter-temporal elasticity of substitution is smaller than one–the second effect dominates, generating negative sorting among constrained individuals.

These results are important because they shed light on the nature of credit-driven distortions in a spatial economy, guiding how housing and place-based policy can best address the inefficiencies driving this misallocation. The presence of negative sorting among constrained individuals implies that untargeted policies aimed at expanding access to high-opportunity cities—such as relaxing land use regulations—will be less effective at correcting spatial misallocation than standard models would suggest. This is because the marginal individual induced to move by lower housing prices will tend to have low learning ability. Effective housing and spatial policy must instead target constrained high-ability workers, facilitating their access to locations that match their learning potential.

In the second part of the paper, I present a novel source of administrative data from Spain recording detailed information on income, wealth, and full working and location histories for a representative sample of Spanish households. I use this dataset to provide evidence on the three key elements of the model: (i) that locations offer heterogeneous opportunities for on-the-job human capital accumulation, (ii) that individuals are heterogeneous in their learning ability, and (iii) that credit constraints distort the spatial distribution of labor in ways consistent with the model.

To separately identify location-specific learning opportunities and individual-specific learning ability, I implement an extended AKM framework (Abowd et al., 1999) that allows locations to affect both wage levels and wage growth, while simultaneously allowing workers to differ both in their initial human capital at labor market entry and their learning ability. This specification preserves the complementarity between individual and location characteristics at the core of the stylized framework, and allows me to recover estimates that can be mapped directly to the elements of the model.

Then, in the last part of the paper, I extend the stylized model to consider an OLG economy that, while preserving the key elements of the theory, also incorporates other drivers of location decisions commonly considered in the literature. I calibrate this economy using the previously estimated parameters, and implement a simulated method of moments aiming to match key features of the Spanish earnings distribution. I then quantify the output losses that can be directly attributed to the effect of borrowing constraints on location decisions. Importantly, these losses arise from distortions in the *composition* of skill in each city: the dominant source of inefficiency is the spatial misallocation of individuals with high-learning ability.

To highlight the role of negative sorting in these results, I study the effect of land use regulation designed to increase housing space in productive cities. I take as a benchmark the aggregate output that would arise under the efficient distribution of labor in the calibrated economy. I then compute what would be the required land expansion in Madrid and Barcelona, the two largest commuting zones in Spain, for the equilibrium economy to reach this same level of output.[1] I contrast this with a situation in which a planner can observe learning ability and reserve new housing exclusively for the highest-ability individuals currently living outside these cities.

[1] I assume the construction sector uses a Leontief technology, which implies the exogenous amount of land available in productive areas maps directly into population size.

In this scenario, the required land expansion is considerably smaller. I confirm this result by computing a ”recovery curve” that sequentially reallocates high-potential workers from low- to high-opportunity locations. Together, this exercise demonstrates that misallocation is concentrated among a relatively small group of high-ability constrained individuals, and that untargeted policies are considerably less effective than they would be in the absence of negative sorting.

I conclude by studying different second-best policies that highlight the importance of targeting constrained high-ability workers when learning potential is unobservable. I compare homeownership subsidies to rental subsidies in productive cities, and find that the marginal agent responding to rental subsidies has substantially higher learning ability than those responding to homeownership subsidies. This occurs because, in the calibrated economy, only credit-constrained agents use rental markets due to the presence of rental frictions. Rental subsidies therefore provide an indirect mechanism to target the population most affected by credit-driven misallocation, demonstrating that in the presence of negative sorting, effective spatial policy requires instruments that can differentially affect the composition of workers across space.

Related literature. This paper relates to several strands of literature. Most importantly, it contributes to the spatial literature studying dynamics in an economy with multiple locations (Kleinman et al., 2023; Greaney et al., 2025). In its characterization of location decisions in an environment with credit constraints, it is most similar to Bilal and Rossi-Hansberg (2021). My stylized model generalizes their environment in terms of preferences and individual heterogeneity. This is key to show that, under empirically relevant parameterizations, credit constraints have the potential to induce negative sorting of learning ability across space among constrained agents, thus amplifying the magnitude of spatial misallocation. Crews (2024) and Lhuillier (2024) study how cities affect human capital accumulation in a model in which agents cannot save. Here I show how these optimal choices are distorted in the presence of borrowing frictions.

This paper also relates to a number of studies analyzing homeownership in a spatial context (Oswald, 2019; Giannone et al., 2023; Greaney, 2023; Luccioletti, 2023; Díaz et al., 2023). Relative to these, I incorporate human capital accumulation and a more flexible characterization of housing markets, placing all the emphasis of tenure choices on the degree to which they can alleviate credit distortions. Finally, within the spatial literature, it contributes to the study of sorting in a static setting (Diamond, 2016; Fajgelbaum and Gaubert, 2020; Diamond and Gaubert, 2022), characterizing these decisions in a dynamic environment.

This paper also contributes to a literature on lifetime earnings inequality and human capital accumulation (Huggett et al., 2006, 2011; Heathcote et al., 2014; Bick et al., 2024), incorporating a spatial dimension. Using data from Spain, De La Roca and Puga (2017) show that location decisions affect life-cycle earnings. I model this source of wage dispersion and quantify its implications.

Finally, the mechanisms presented here are similar to those studied in the educational investment literature (see Lochner and Monge-Naranjo (2012) for a review). Within that literature, some papers have studied the role of location decisions on human capital through the residential choice of parents as an investment in their childrens' schooling (Benabou, 1993; Fernandez and Rogerson, 1996; Fernández and Rogerson, 1998; Eckert and Kleineberg, 2024; Fogli et al., 2024). These papers typically rule out the existence of capital markets allowing parents to borrow against their children's future income. To the extent that human capital accumulation does not cease once the individual stops studying, this paper extends similar arguments to later points in the life-cycle. Additionally, I show how the structure of housing markets provides heterogeneous access to on-the-job human capital accumulation.

The rest of this paper is organized as follows. Section 2 presents a stylized model of location decisions under credit constraints. Section 3 introduces the data and provides evidence of the key elements of the theory. To be able to quantify the degree of misallocation driven by credit frictions, Section 4 extends the stylized economy to a more general structure that can be taken to the data. Section 5 describes the calibration strategy. Section 6 presents the results of the quantification exercise and studies the type of policies that can better target the relevant source of misallocation. Finally, Section 7 concludes.

2 Stylized Model

This section develops a stylized model of location choice in the presence of credit constraints. In this stylized economy, locations differ only in the opportunities they offer for on-the-job human capital accumulation. The goal of this section is to characterize how and when limited access to credit can distort individual location decisions and, consequently, aggregate human capital in the economy. To best highlight the key mechanisms, this model abstracts from several standard features present in spatial economies. Section 4 will later embed the insights captured by this simple model into a richer quantitative spatial framework.

2.1 Environment

Households. There is a unit mass of households that live for two periods with preferences given by

$$U = u(c_1) + \beta u(c_2), \qquad u(c) = \frac{c^{1-\frac{1}{\sigma}} - 1}{1 - \frac{1}{\sigma}}$$

where c_t represents consumption of a freely tradable, final good in period t, $\beta \in (0,1]$ is a discount factor, and $\sigma > 0$ is the intertemporal elasticity of substitution (IES) for consumption. Households are endowed with one unit of time each period and do not value leisure.

Agents are heterogeneous along two dimensions: their initial endowment of the final good, $k \in [\underline{k}, \bar{k}]$, and their learning ability, $a \in [\underline{a}, \bar{a}]$. Heterogeneity is described by the cumulative distribution function $\Lambda(k,a)$, with $\underline{k} \geq 0$ and $\underline{a} > 0$. All agents enter the labor market with the same initial stock of human capital, denoted by h_1.[2] Learning ability affects how this stock evolves over time and therefore drives the dispersion of human capital in the second period. At any period t, agents earn labor income proportional to h_t.

Locations. In the first period, agents can freely choose where to live among a continuum of cities. Urban areas are indexed by the opportunities they offer for on-the-job human capital accumulation, $\ell \in [\underline{\ell}, \bar{\ell}] \subset \mathbb{R}^{++}$, with $\underline{\ell} \geq 1$.

The agent's learning ability a determines the returns to residing in each location. More precisely, an individual with initial human capital h_1 and learning ability a who chooses to reside in location ℓ will have human capital in the second period equal to $h_2 = \ell a h_1$. This law of motion is meant to highlight the role of cities in

[2] This assumption is imposed for expositional clarity. Appendix A shows that the key results presented here hold in the presence of heterogeneity in the initial endowment of human capital.

on-the-job human capital accumulation (De La Roca and Puga, 2017; Crews, 2024; Lhuillier, 2024). As I describe in detail below, the supermodularity embedded in this formulation generates positive sorting of ability across space when individuals are not credit constrained.

Space in each city is fixed, with local housing supply distributed according to a density function $g(\ell)$.[3] Each house can accommodate at most one person and the total mass of urban space is normalized to unity. In order to access a location, agents in the first period must purchase a house in a competitive market at price $p(\ell)$. Houses fully depreciate at the end of the second period, provide no direct utility, and have no residual value. The limited availability of space in each location will determine housing prices in equilibrium. Since all housing payments occur in the first period, housing costs behave as an (endogenous) fixed cost of accessing a location.

The assumption that all agents are homeowners is admittedly a strong assumption, but not a critical one. It allows me to derive sharp analytical results characterizing sorting patterns, which is the primary objective of this stylized framework.[4] In the quantitative model presented in Section 4, I introduce rental markets and a tenure choice. As will be discussed then, the specifics of the housing market will be important to determine the mass of agents that are credit constrained. However, I show that, conditional on constrained status, the same sorting patterns hold in a more general housing environment.

The assumption of a fixed housing supply in each location serves to isolate the effect of credit frictions on the *composition* of cities. As I discuss below, the sorting patterns characterized here hold under any increasing price schedule. The assumption of perfectly inelastic supply thus simplifies the analysis without driving the main results. I impose this assumption to emphasize the key insight of this paper: in a dynamic spatial economy with heterogeneous households and heterogeneous opportunities for on-the-job human capital accumulation, *who* lives in each location, and not just *how*

[3] I assume the associated CDF, $G(\ell)$, is strictly increasing and continuous in ℓ. Rather than endowing each location with a fixed mass of houses, we could alternatively model a construction sector with a Leontief technology using final goods and land in fixed supply. Assuming land is distributed according to $g(\ell)$ and owned by absentee landlords who absorb all profits would generate the same results.

[4] This assumption also connects the model to the educational investment literature, summarized in Lochner and Monge-Naranjo (2012). The individual problem resembles one in which agents invest in human capital through educational choices and given an exogenous price schedule (Lochner and Monge-Naranjo, 2011). I show that a spatial economy provides a natural setting to study these forces in equilibrium: the cost of accessing better opportunities for human capital accumulation is endogenous, determined by limited housing supply and aggregate sorting patterns.

many, matters for aggregate outcomes.

Production. In each city, a representative competitive firm produces a perfectly tradable good using a CRS technology with human capital as the only input. The price of the final good is normalized to one. Firms are homogeneous across space and production exhibits no complementarities. This structure pins down labor income as wh, where w represents the common productivity level.[5]

Financial Markets. Agents have access to a risk-free bond, b, with exogenous gross interest rate $R > 1$. Agents face an exogenous borrowing limit, $\underline{b} \geq 0$, such that bond holdings must satisfy $b \geq -\underline{b}$. A more general specification with collateralized borrowing is presented in the Appendix.

Equilibrium. Using the final good as numeraire, an individual with initial endowment k and learning ability a solves the following problem

$$\begin{aligned} &\max_{c_1,c_2,b,\ell} u(c_1) + \beta u(c_2) && (1) \\ &\text{s.t. } c_1 = k + wh_1 - b - p(\ell) \\ &\qquad c_2 = wh_2 + Rb, \quad h_2 = \ell h_1 a \\ &\qquad b \geq -\underline{b} && (\mu) \end{aligned}$$

A critical feature of this environment is the misalignment between the timing of expenditures and the timing of income. Accessing better locations requires upfront housing payments $p(\ell)$ in period 1, but the returns to this decision are not delivered until period 2. When agents can borrow freely, financial markets resolve this temporal wedge. But when borrowing is constrained, location choice becomes a margin through which to transfer resources across time, distorting the distribution of learning ability across space.

This problem resembles the stylized model in Bilal and Rossi-Hansberg (2021) with one key difference: they impose logarithmic preferences ($\sigma = 1$), while I allow for a more general preference structure. This generalization is critical for the main

[5]This simplification focuses attention on the role of locations in human capital accumulation. Both the empirical exercises and the quantitative model in Section 4 consider heterogeneous productivity levels across space and therefore allow for spatial variation in returns to human capital.

analysis, since the intertemporal elasticity of substitution plays a key role to determine the sign of sorting.[6]

We now have all the necessary ingredients to define an equilibrium.

Definition 1. *Given a distribution Λ of initial endowments and learning ability $(k,a) \in [\underline{k},\bar{k}] \times [\underline{a},\bar{a}]$, a fixed distribution of housing in urban areas $g(\ell)$, and a gross interest rate R, an equilibrium is a list of policy functions $\{c_1(k,a), c_2(k,a), b(k,a), l(k,a)\}$, and a housing price schedule $p(\ell)$, such that*

1. *Given prices, policy functions solve the individual problem (1),*
2. *Final good firms maximize profits taking prices as given,*
3. *Local housing markets clear,*

$$\int_k \int_a \mathbb{1}[l(k,a) \leq \ell]\Lambda(dk,da) = G(\ell) \equiv \int_{\underline{\ell}}^{\ell} g(s)ds, \qquad \forall \ell \in [\underline{\ell},\bar{\ell}]. \tag{2}$$

The housing market clearing condition in equation (2) ensures that the mass of households that choose to live in urban areas offering opportunities for on-the-job human capital accumulation weakly less than ℓ is equal to the total mass of housing space available below that value.

The rest of this section proceeds in three steps. Subsection 2.2 begins by characterizing location decisions in an economy where agents face no effective borrowing limit. This serves to establish the efficient allocation of labor across space and provides a benchmark against which to evaluate the distortions induced by limited access to credit. Section 2.3 then introduces credit constraints and characterizes how borrowing limits distort location decisions for both constrained and unconstrained agents. Finally, Section 2.4 discusses general equilibrium properties.

2.2 Benchmark Economy

Consider a version of the previously described economy where $\underline{b} \to \infty$, that is, agents face no effective borrowing limit. This implies that the borrowing constraint $b \geq -\underline{b}$ never binds and all agents optimally allocate consumption intertemporally.

[6]They also include an exogenous endowment of goods k_2 received in the second period. Appendix A shows that the qualitative results hold in a more general environment with a standard value for the IES as long as labor income is the primary source of second-period resources.

I characterize individual location choices for a given price schedule $p(\ell)$. Locations offering better learning opportunities generate higher lifetime income. Since housing in each city is in fixed supply, it follows that housing prices must be increasing in ℓ in equilibrium.

Lemma 1. *The equilibrium price schedule $p^*(\ell)$ is increasing in ℓ.*

Proof. See Appendix A. □

Agents in the benchmark economy choose their optimal level of savings to satisfy the Euler equation

$$u'(c_1) = \beta R u'(c_2).$$

For expositional purposes I assume $\beta R = 1$ throughout this section, although this is not required for any of the proofs.[7] In this case, the Euler equation above states that agents will perfectly smooth consumption over time.

With $p'(\ell) > 0$, better learning opportunities come at the cost of higher housing prices. Since in this benchmark economy agents can optimally smooth consumption using financial markets, they choose their location to maximize lifetime income net of housing costs, and then save or borrow to achieve their desired consumption path. The location problem reduces to:

$$\max_{\ell \in [\underline{\ell}, \bar{\ell}]} \quad k + wh_1 + \frac{1}{R} w\ell h_1 a - p(\ell).$$

Whenever the problem returns an interior solution, from the first-order condition we obtain that location decisions are implicitly defined as

$$\underbrace{R}_{\substack{\text{Return on} \\ \text{financial assets}}} = \underbrace{\frac{wh_1 a}{p'[l^*(a)]}}_{\substack{\text{Marginal net return} \\ \text{on location invest.}}}, \tag{3}$$

where I use the notation $l^*(a)$ to denote location choices in the benchmark economy and to emphasize that these decisions do not depend on initial endowments k.

[7] Appendix A presents all the proofs in the paper without imposing this assumption.

Equation (3) implies that agents will choose the location that equates the marginal return on their location investment to the return on financial assets. The supermodularity embedded in the learning technology implies that higher learning ability increases the marginal return to better locations, generating positive sorting of ability across space. The following proposition formalizes these results.

Proposition 1. *For any increasing price schedule $p(\ell)$, location choices in the benchmark economy, $l^*(a)$, are:*

1. *Independent of initial endowments (k).*
2. *Increasing in learning ability (a).*

Proof. See Appendix A.1. □

Agents in the benchmark economy sort into cities on the basis of their learning ability alone, using credit markets to optimally smooth consumption over time.

Before analyzing this economy in the presence of binding credit constraints, it is useful to study how the spatial distribution of labor compares to the a first-best allocation. The problem of the planner can be decomposed into two steps. First, allocate agents across space to maximize total second-period output, subject to housing supply constraints. Second, redistribute resources to achieve the optimal distribution of consumption.

The planner chooses a distribution of agents across space, $\pi(a,\ell)$, to solve

$$\max_{\pi(a,\ell)} \int_a \int_\ell a\ell\, \pi(a,\ell)\, da\, d\ell \quad \text{subject to} \int_a \pi(a,\ell)\, da = g(\ell)\ \forall \ell, \quad \int_\ell \pi(a,\ell)\, d\ell = \lambda_a(a)\ \forall a, \tag{4}$$

where $\lambda_a(a)$ is the marginal density of learning ability and I drop the factor wh_1 from the objective, as it does not affect the optimal assignment.

As discussed in Appendix A.2, the solution to this problem features positive sorting. The planner assigns higher-ability agents to higher-opportunity locations to exploit the complementarity between learning ability and local opportunities. An implication of Proposition 1 is that, when agents face no effective borrowing limit, the equilibrium of the benchmark economy also features positive sorting of ability across space. Since both the planner's allocation and the competitive equilibrium are positively assortative and satisfy the same marginal constraints, they must coincide. Therefore, when agents face no effective borrowing limit, the benchmark economy replicates the efficient spatial allocation of labor.

2.3 Location Choices in the Presence of Credit Constraints

Consider now an economy in which $\underline{b}$ is finite and, given the observed price schedule $p(\ell)$, credit constraints bind for a positive mass of agents. The presence of binding borrowing constraints requires addressing two questions: which agents are credit constrained, and how do credit constraints affect location decisions for both constrained and unconstrained agents.

Lemma 2 addresses the first of these two concerns.

Lemma 2. *For any increasing price schedule $p(\ell)$, there exists a threshold function $a^*(k)$ such that agents with learning ability $a > a^*(k)$ are credit constrained, while agents with $a \leq a^*(k)$ are unconstrained. Additionally, $a^*(k)$ is increasing in k.*

To understand this result, notice that individuals with higher learning ability want to borrow *more* for two reasons. First, as established by Proposition 1, in the absence of binding credit constraints agents with higher a would optimally choose locations offering better learning opportunities. For any increasing price schedule, this implies higher housing payments in the first period and therefore requires higher amounts of borrowing for any k. Second, even conditional on location choice, higher learning ability implies higher future income, increasing the incentive to borrow to smooth consumption across periods. Both channels imply that, conditional on their good endowment k, agents with higher learning ability are more likely to exhaust their borrowing capacity.

Having established which agents are credit constrained, I now turn to characterizing how these frictions affect location decisions, taking the price schedule $p(\ell)$ as given. Agents whose desired borrowing is less than $\underline{b}$ face no binding constraint. Since these unconstrained agents can use financial markets to optimally smooth consumption across periods, they choose locations to maximize discounted lifetime income net of housing costs, then save or borrow to achieve their desired consumption path. Conditional on a fixed price schedule, these location choices exhibit the same properties established in Proposition 1: they are independent of initial endowments and increasing in learning ability.

Constrained agents face a different tradeoff. Having exhausted their borrowing capacity, they cannot use financial markets to smooth consumption and must rely on location choices to shift resources between today and tomorrow. This creates tension between two objectives: maximizing lifetime income (which favors high-ℓ cities) and increasing consumption in the period in which they have least resources (which favors low-cost, low-ℓ cities).

To characterize optimal location decisions, consider the first-order condition with respect to ℓ in Problem (1). For constrained agents with $b^*(k,a) = -\underline{b}$, substituting the budget constraints into this condition yields that, whenever an interior solution exists, optimal location choices $l(k,a)$ are implicitly defined by

$$u'\Big[k + wh_1 + \underline{b} - p\Big(l(k,a)\Big)\Big]p'\Big(l(k,a)\Big) = \beta u'\Big[w\ell a h_1 l(k,a) - R\underline{b}\Big]wah_1. \tag{5}$$

The left-hand side of equation (5) represents the marginal cost of choosing a better location: lower first-period consumption due to higher housing costs. The right-hand side represents the marginal benefit: higher second-period consumption from increased human capital accumulation. The following proposition characterizes how this tradeoff depends on learning ability and the intertemporal elasticity of substitution.

Proposition 2. *For any increasing price schedule $p(\ell)$, the location choices of constrained agents are decreasing in learning ability (a) when the intertemporal elasticity of substitution (σ) is smaller than one.*

Proposition 2 states that, whenever $\sigma < 1$ and conditional on a fixed level of initial endowments k, agents with marginally higher learning ability optimally choose to reside in locations offering worse learning opportunities.

This result reflects two opposing forces: a substitution and an income effect. On one hand, the supermodularity in the learning technology means that higher learning ability increases the return to locations offering higher learning opportunities. This pushes marginally better learners toward better locations. On the other hand, high-ability agents anticipate higher future income, increasing the relative value of first-period consumption due to diminishing marginal utility. Having exhausted their borrowing capacity, constrained agents can only increase first-period consumption by saving on housing costs, hence choosing locations offering worse learning opportunities. When preferences exhibit relatively strong consumption smoothing motives ($\sigma < 1$), the income effect dominates: high-ability agents prefer cheaper housing today, accepting lower learning opportunities in exchange of a smoother consumption path across periods.

When borrowing is not allowed ($\underline{b} = 0$), the characterization simplifies considerably. In this case, the direction of sorting depends exclusively on the value of the intertemporal elasticity of substitution (σ), with log utility representing a knife-edge case where the substitution and income effects exactly offset.

Corollary 1. *Assume agents cannot borrow ($\underline{b} = 0$) and utility is CRRA with intertemporal elasticity of substitution σ. Location choices of constrained agents are:*

- *Increasing in learning ability a if $\sigma > 1$.*
- *Independent of learning ability a if $u(c) = \log(c)$.*
- *Decreasing in learning ability a if $\sigma < 1$.*

It is worth highlighting that this result does not necessarily contradict evidence of positive sorting in human capital *levels* across space. Appendix A shows that, with heterogeneity in initial human capital (h_1), the model can generate positive sorting in levels among constrained agents even when sorting in learning ability remains negative, everything else held constant.[8]

To show these results graphically, Figure 1 plots location decisions across different levels of learning ability (a) holding initial good endowments (k) and the price schedule constant. For the later I use the equilibrium price schedule under the benchmark economy, $p^*(\ell)$.

The dashed orange line shows location choices in the benchmark economy, where borrowing constraints never bind ($\underline{b} \to \infty$). As established by Proposition 1, agents sort positively across space.

The different maroon lines plot location decisions when agents cannot borrow

[8]This last fact can also be inferred from Corollary 1, which establishes that the direction of sorting among constrained agents depends exclusively on σ, and therefore holds for any initial level of h_1.

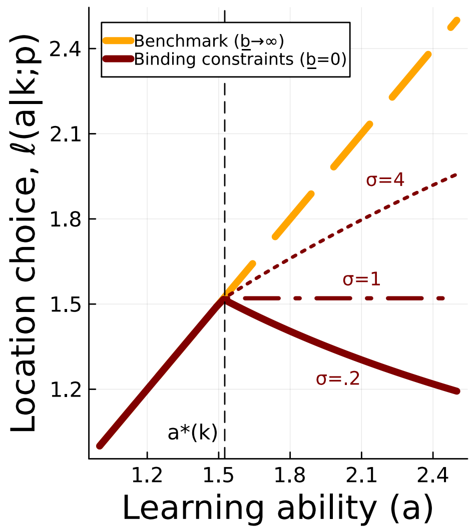


Figure 1: Optimal location decisions by learning ability.

($\underline{b} = 0$), corresponding to the economy described in Corollary 1. As stated in Lemma 2, agents with learning ability below $a^*(k)$ remain unconstrained. Among this subset, marginally better learners choose to reside in marginally better locations. Agents with $a > a^*(k)$ are credit constrained; among these constrained agents, the relationship between learning ability and location decisions depends on σ. When $\sigma < 1$, further increases in ability lead to *worse* location choices as the income effect dominates.

I highlight individual behavior when $\sigma < 1$ because it represents the empirically relevant scenario: a broad empirical literature estimates the intertemporal elasticity of substitution to be below one, with the standard estimate around 0.5 (Havránek, 2015; Thimme, 2017).

A useful implication for empirical work follows from comparing agents in the same location.

Lemma 3. *Given a fixed and increasing price schedule, constrained agents choose weakly worse locations than unconstrained agents with the same learning ability.*

This lemma delivers a testable prediction about within-location heterogeneity. Among agents residing in the same location ℓ, those who are credit-constrained must have weakly higher learning ability than unconstrained residents. To see this, note that a constrained agent at location ℓ would choose $\ell' > \ell$ if unconstrained (Lemma 3), which by Proposition 1 requires higher a. Therefore, conditional on location, constrained agents should experience higher subsequent wage growth.

Before moving on to study these forces in the data, the next section characterizes the economy in equilibrium.

2.4 Sorting Patterns in Equilibrium

The previous subsection characterized individual location choices for any increasing price schedule. A natural question is whether such schedules emerge in equilibrium and, if so, what sorting patterns they generate. Proposition 3 establishes properties that must hold in any equilibrium of this economy.

Proposition 3. *In any equilibrium of this economy the price schedule is increasing in ℓ. Conditional on their level of initial endowments, k,*

1. *Unconstrained agents sort positively across space: location choices are increasing in learning ability a.*
2. *If $\sigma > 1$, constrained agents sort positively across space: location choices are increasing in learning ability a.*
3. *If $\sigma < 1$, constrained agents sort negatively across space: location choices are decreasing in learning ability a.*

The proposition reveals how credit constraints and preferences interact to shape equilibrium sorting patterns. When $\sigma < 1$ (the empirically relevant case), the income effect dominates among constrained agents, generating negative sorting in learning ability conditional on initial endowments. Unconstrained agents continue to sort positively in a, producing a mixed pattern in equilibrium. When $\sigma > 1$, the substitution effect dominates: conditional on initial endowments k, all agents sort positively in learning ability.

To evaluate efficiency, consider again the first-best allocations determined in Problem (4). Recall that the planner assigns agents to locations based solely on learning ability, with higher-ability agents allocated to better locations. When $\sigma > 1$, Proposition 3 established that agents sort positively in learning ability conditional on their initial endowments. However, this does not necessarily imply that the equilibrium replicates the planner's efficient assignment. As established in Lemma 3, constrained agents choose weakly worse locations than unconstrained agents with the same learning ability. In the presence of heterogeneity in initial endowments, credit constraints can therefore distort the spatial allocation of skill: agents with high endowments but low learning ability can outbid constrained, high-ability agents for housing in productive cities. This is the case analyzed in Bilal and Rossi-Hansberg (2021). This crowding-out effect distorts the composition of cities, and the equilibrium allocation of skill does not necessarily match the planner's solution, even though higher-ability agents continue selecting better locations than lower-ability agents with the same k. When $\sigma < 1$, the losses are compounded. In addition to this crowding-out mechanism, the income effect characterized in Corollary 1 reverses location choices among constrained agents, generating negative sorting and systematically misallocating high-ability agents to low-opportunity cities, which further reduces aggregate human capital accumulation.

These findings have important implications for the design of spatial and housing policy. Standard place-based policies aim to expand the size of productive cities through housing construction, regulation, or infrastructure investment. While such policies would generate output gains in this economy, in the empirically relevant case where $\sigma < 1$ they may have limited effectiveness in correcting the misallocation of learning ability across space. In the presence of negative sorting, when housing supply expansions reduce prices in productive cities, the marginal individual induced to relocate will tend to be drawn from a pool of relatively low-ability individuals. Taking credit frictions as given, addressing the spatial misallocation of talent requires policies that facilitate access to productive cities specifically for constrained high-ability individuals. The challenge for policy design is thus not simply to make productive cities larger, but to improve their composition by targeting constrained workers with high learning potential.

The remainder of the paper evaluates these mechanisms empirically and quantitatively. Section 3 provides evidence consistent with the sorting patterns characterized here using administrative data from Spain. Section 4 develops a quantitative spatial model that embeds the mechanisms identified in this stylized framework within a richer environment featuring rental markets, regional productivity differences, and an overlapping generations structure with human capital accumulation over the life cycle. I use this model to quantify the losses derived from spatial misallocation and compare different placed-based interventions to highlight those that best target constrained individuals.

3 Sorting and Human Capital Accumulation in Spain

This section provides empirical evidence for the key elements of the stylized model. Using administrative data from Spain, I document three patterns central to the theory. First, locations offer heterogeneous opportunities for on-the-job human capital accumulation: workers in larger cities experience faster wage growth, particularly early in their careers. Second, individuals are heterogeneous in their learning ability: conditional on location and experience, some workers accumulate human capital faster than others. Third, credit constraints distort location decisions: among workers residing in the same location, those with lower initial wealth exhibit higher subsequent wage growth, consistent with model predictions highlighted on the previous sections.

I divide this section into three parts. First, I introduce the Spanish Household Panel, a novel administrative dataset linking wealth, income, and complete employment histories at the individual level. Then, I document how wage growth varies across locations and over the life cycle, establishing stylized facts that are consistent with the model's mechanisms and replicable in standard datasets. Finally, I implement an extended AKM framework to separately identify location-specific learning opportunities and individual-specific learning ability, recovering the heterogeneity and location-specific parameters that will discipline the quantitative model in Section 4.

3.1 Data Description: The Spanish Household Panel

The primary data source used in this paper is the newly developed Spanish Household Panel. This project links administrative records from different Spanish institutions, providing a representative sample of approximately 5% of all households residing in Spain between 2016 and 2019.[9]

For every person living in a selected household, I observe basic demographic information (age, gender, nationality), census block of residence, and all sources of individual income. A key feature of this dataset is that it provides information on both the stock and income flows from a comprehensive list of financial and real assets, excluding only business wealth. Financial institutions operating in Spain are

[9]The full dataset contains information up to 2023 and is expected to be updated yearly. I limit my analysis to 2019 in order to avoid the Covid crisis in Spain. A detailed description of the dataset can be found in Appendix D.

legally required to report this information, ensuring reliability through third-party verification rather than self-reporting.

Each individual in the sample is linked to employment records from the Spanish Continuous Sample of Employment Histories (*Muestra Continua de Vidas Laborales*, or MCVL, in Spanish). This dataset provides complete working (and location) histories at the individual level.[10] For each employment spell in the individual's working life, I can observe start and end dates for the spell, working hours expressed as a percentage of full-time equivalent job, type of contract, and establishment characteristics including municipality, sector, and number of employees. This information allows me to construct full workplace location histories at the individual level since labor market entry.

My baseline sample uses employment information for men attached to the standard employment regime in Spain between 2016 and 2019.[11] This criterion mostly excludes the self-employed and workers in the primary sector. I also exclude workers employed in public firms or public-dominant sectors, such as education and healthcare. These sectors are heavily regulated, and therefore I expect human capital accumulation to play a smaller role in wage determination. Finally and in order to ensure that I observe full working histories, for the wage regressions included in Section 3.3 I also limit my sample to individuals born between 1971 and 2001.[12] These individuals will be aged between 25 and 45 during my sample period.

I convert the dataset into a monthly panel. In those cases in which individuals hold more than one job in the same month, I select the one representing the highest source of income. Earnings are reported as full-time-equivalent monthly wages. All income variables are deflated by the Spanish CPI, with 2018 as the reference year.

This selected subsample contains about 1.6 million observations, following 34,072 individuals over a maximum of 48 months. For further details on the dataset, the construction of the main sample, or summary statistics, see Appendix D.

[10]The MCVL can be accessed as an independent source of information. It has been extensively used in previous research. As the closest reference, it is the data source used in De La Roca and Puga (2017).

[11]I define an individual as attached to the labor market if, for each year between 2016 and 2019, he or she works at least 91 full-time equivalent days per year. This follows Guvenen et al. (2022) and Bick et al. (2024), who consider individuals attached if they work at least 520 hours per year.

[12]Detailed information on employment spells characteristics is available since 1980, which would imply full working histories are provided for any person born after 1964. During the first few years, however, records are not well-populated. For this reason, and to also match the period length considered in the quantitative section, I only introduce in the regressions those individuals born after 1970.

3.2 Properties of Lifetime Wages

The stylized model in Section 2 has three key features: locations differ in the opportunities they offer for on-the-job human capital accumulation, individuals are heterogeneous in their learning ability, and accessing better locations represents an investment that can be constrained by limited resources. This subsection presents descriptive evidence consistent with these mechanisms.

Figure 2 plots average wage growth rates across space and over the life cycle. Wage growth is substantially higher in Madrid and Barcelona than in smaller cities, particularly for young workers early in their careers. Among workers aged 24-33, those in Madrid and Barcelona experience wage growth rates approximately 5 percentage points higher than workers in other locations. This gap narrows considerably for middle-aged workers (ages 34-44), and by ages 45-55, the spatial differential in wage growth largely disappears. This pattern is consistent with heterogeneous opportunities for on-the-job human capital accumulation, and suggests that location choices are particularly important for human capital accumulation early in the career. The stylized model highlights that these differences in observed wage growth reflect both location-specific learning opportunities and the sorting of high-ability workers into productive cities. The estimation exercise in Section 3.3 will disentangle these two components.

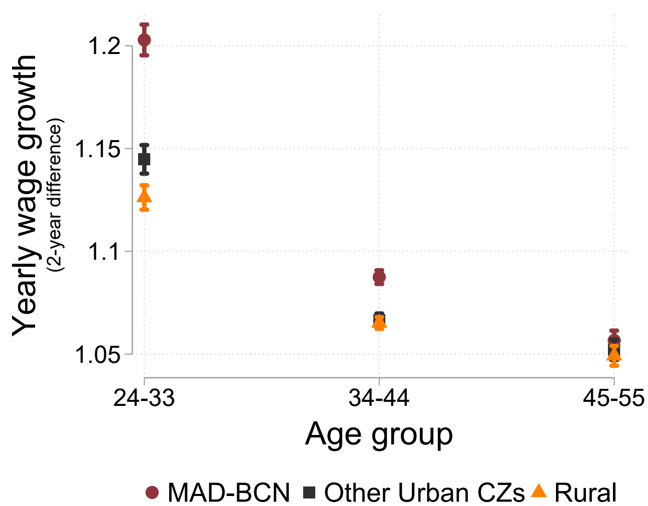


Figure 2: Average wage growth across space and over the life-cycle.

Note: Yearly wages are defined as total labor earnings in a given year divided by the full-time equivalent number of days worked. Figures plotting 1- and 3-year differences look qualitatively similar.

The dynamic benefits of cities are capitalized in housing prices, creating substantial upfront costs. Since young workers cannot immediately realize wage growth gains, location choices represent an investment decision. Figure 3 illustrates this tradeoff across Spanish cities. The solid line plots median monthly wages for workers aged 22-26, while the dashed line shows median rental prices. The median young worker in Madrid earns only 6.9% more than those in the median commuting zone (Santiago de Compostela), but faces rental costs that are 57% higher. Young workers in high opportunity cities thus sacrifice current consumption for higher future earnings, making location choices an investment.

Figure 3: Median income and rental prices across commuting zones in Spain.

Having documented these patterns, I turn to Lemma 3 from the previous section for evidence on the role of credit constraints on location decisions. Proposition 1 showed that unconstrained agents sort positively into locations according to their learning ability. Lemma 3 established that a constrained agent with the same learning ability will choose instead a location offering worse opportunities for on-the-job human capital accumulation. This implies that, conditional on observing both constrained and unconstrained agents *in the same location*, those who are constrained should have higher learning ability. Otherwise, they would have chosen a location offering better opportunities. This difference in learning ability manifests as faster wage growth among constrained individuals within a given location.

The main caveat when testing this implication is that constrained status is not directly observable in the data, as it depends on unobservable learning ability.[13] I

[13]For a constant initial endowment, agents with marginally higher learning ability will have higher

therefore proxy for constrained status using net wealth. Figure 4 plots wage growth rates by location and wealth quantile for young workers in the Household Panel, where I define wealth quantiles within age-location cells. The figure reveals that, outside of the most productive locations, workers in the bottom wealth quartile experience 5 percentage points higher wage growth than those in the top quartile. This pattern is consistent with the model's prediction: low-wealth (likely constrained) workers have higher learning ability conditional on location, suggesting credit-driven distortions in the allocation of skill across space.

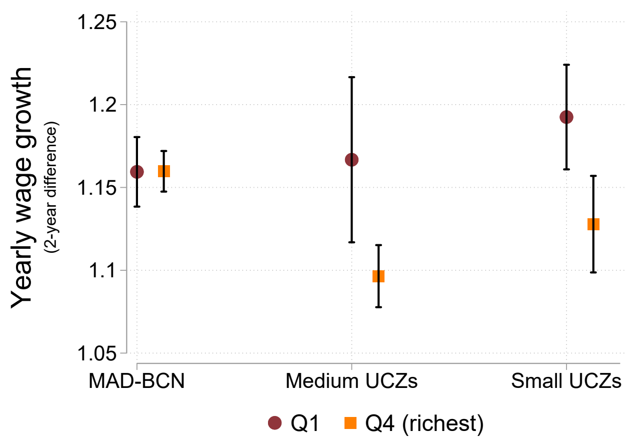


Figure 4: Wage growth by location and wealth quantile.

Note: Point estimates in this figure are obtained regressing 2-year growth rates on quantiles of wealth, controlling for homeownership and the individual's initial wage. Excluding these controls does not substantially affect the results. As the relevant measure of wealth, defined within age-location cells, I include assets of high and medium liquidity (excluding pensions, long-term insurance, and housing) as reported when the individual first appears in the sample. Confidence intervals are at the 90% level.

Finally, Figure 5 plots the log-wage distribution across the age distribution, using the individual mean of yearly wages between 2016 and 2019. Consistent with findings in the literature studying inequality over the life cycle (Huggett et al., 2006, 2011; Heathcote et al., 2014; Bick et al., 2024), this figure presents an increasing cross-sectional dispersion of log wages as individuals age.[14]

Although the model attributes part of this dispersion to location decisions, I find substantial heterogeneity in income profiles even at the local level. This within-location dispersion is suggestive of heterogeneity in individual learning ability, which

income in the future, which means they will be inclined to transfer more resources to the present to smooth a higher level of lifetime consumption.

[14]Similar patterns hold even if we partition the sample by gender and education levels.

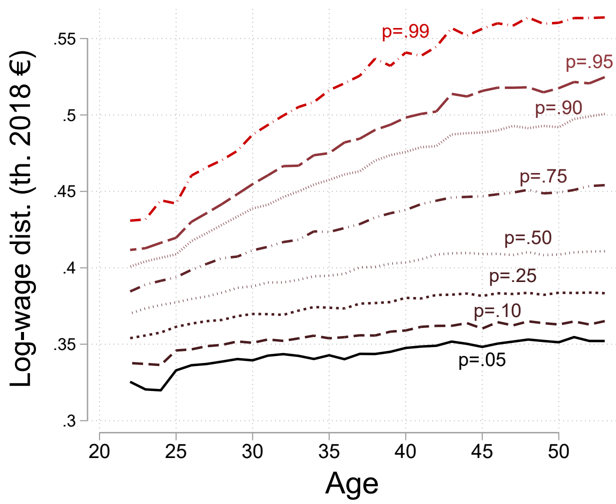


Figure 5: Percentiles of the log-wage distribution.

will motivate the estimation strategy in the next subsection.

3.3 Wage Dynamics and the Impact of Learning Ability on Earnings

The previous subsection documented substantial heterogeneity in wage growth across Spanish commuting zones. For the purposes of this paper, it is key to decompose this spatial variation into two components: location-specific learning opportunities and individual-specific learning ability. To do so, I implement an extended AKM framework that allows locations to affect both wage levels and wage growth, while simultaneously allowing workers to differ in their initial human capital at labor market entry and their learning ability.

3.3.1 A Model of Human Capital Accumulation

Building on the learning technology in the stylized model, consider a law of motion for human capital that incorporates both individual heterogeneity and location-specific learning opportunities. Suppose human capital evolves according to

$$h_{i,j+1} = h_{i,j}\left(\psi_{\ell(i,j)} a_i\right)^{\delta_j}, \tag{6}$$

where $h_{i,j}$ is the human capital of individual i at age j, $\psi_{\ell(i,j)}$ captures on-the-job learning opportunities present in the location where individual i was residing at age j, a_i is the individual's learning ability, and δ_j is an age-specific parameter that allows the contribution of learning to human capital accumulation to vary over the life cycle.

The age-specific parameter δ_j is motivated by the well-documented fact that wage growth declines over the life cycle. Following Keane and Wolpin (1997) and Imai and Keane (2004), this specification allows learning opportunities to have differential effects at different ages.[15]

This specification is useful because it allows me to write current human capital as a function of individual fixed effects and complete location histories. Iterating equation (6) backward from age j to age 25, the logarithm of human capital can be written as

$$\log h_{i,j} = \underbrace{\log h_{i,25}}_{\substack{\text{initial human}\\ \text{capital (age 25)}}} + \underbrace{\log(a_i)\sum_{s=25}^{j-1}\delta_s}_{\substack{\text{cumulative effect}\\ \text{of learning ability}}} + \underbrace{\sum_{s=25}^{j-1}\delta_s \log\psi_{\ell(i,s)}}_{\substack{\text{local contributions}\\ \text{to human capital acc.}}}. \tag{7}$$

Conditional on observing full location histories, this specification allows me to separate location-specific learning opportunities from compositional differences driven by worker sorting.

3.3.2 Estimating Equation

To bring equation (7) to the data, assume that earnings for individual i at time t are given by $w_{i,t} = z_{\ell(i,t)}h_{i,t}$, where z_ℓ is a location-specific productivity shifter that captures static differences in wage levels across locations. Taking logarithms and using the decomposition in (7), log earnings can be written as

$$\log w_{i,t} = \log z_{\ell(i,t)} + \log h_{i,25} + \log(a_i)\sum_{s=25}^{age(i,t)-1}\delta_s + \sum_{s=25}^{age(i,t)-1}\delta_s \log\psi_{\ell(i,s)}.$$

[15] When δ_j is decreasing in j, a year of experience accumulated early in the career contributes more to human capital than the same year accumulated later in the life cycle, generating concave income profiles. An alternative approach introduces curvature directly into the learning technology as a function of cumulative experience (Blandin, 2018).

Define age-weighted cumulative experience as $E_{i,t} \equiv \sum_{s=25}^{t} \delta_s$ and location-specific experience as $E_{i,\ell,t}$, measuring the total age-weighted experience accumulated in location ℓ up to time t, so that $\sum_\ell E_{i,\ell,t} = E_{i,t}$. This yields a more familiar estimating equation,

$$\log w_{i,t} = \log h_{i,25} + \log z_{\ell(i,t)} + \log(a_i) E_{i,t-1} + \sum_{\ell=1}^{\mathcal{L}-1} \log(\psi_\ell) E_{i,\ell,t-1} + u_{i,t}, \tag{8}$$

where $\mathcal{L}$ is the total number of locations and $u_{i,t}$ captures measurement error, which I assume to be mean independent of current and lagged location choices.

This specification extends the canonical AKM framework (Abowd et al., 1999) in two important ways. First, following De La Roca and Puga (2017), it allows locations to affect both wage levels (z_ℓ) and wage growth (ψ_ℓ), capturing both static productivity differences and dynamic learning opportunities. Second, similar to Gregory (2023), it introduces individual-specific returns to experience through a_i, allowing workers to differ in how they transform experience into human capital. [16]

Identification. As in the canonical AKM framework, identification of the location effects ($\log z_\ell, \log \psi_\ell$) in equation (8) relies critically on workers who move across location partitions. Intuitively, individual fixed effects $\log h_{i,25}$ and $\log a_i$ are identified from within-person wage variation over time, while location effects are separately identified from wage changes experienced by movers when they switch locations. I now discuss the key behavioral assumption underlying identification, the functional form restrictions imposed by the model structure, and two practical estimation challenges that arise in implementation.

Strict exogeneity. The exogeneity condition imposed by this model parallels the canonical AKM assumption, ruling out systematic endogenous mobility in response to transitory wage shocks after controlling for individual heterogeneity and location histories. In other words, workers do not "move on a shock" beyond what is captured

[16] Gregory (2023) finds that heterogeneity in firm learning environments accounts for 40% of the increase in the cross-sectional earnings variance over the life cycle. To the extent that we would expect high-learning firms to be geographically concentrated, the distinction between firm-specific and location-specific learning opportunities is not important for the novel mechanism considered in this paper. In both cases, learning opportunities are capitalized into local housing costs in equilibrium, creating the same barrier to entry and returning the same sorting implications for credit-constrained workers.

by the components of the model.

Functional form restrictions. The specification in equation (8) imposes log-additivity between individual and location effects, limiting the degree of complementarity between worker and location characteristics, both in the production and learning technology. On the other hand, this specification introduces an important dynamic dimension absent from the canonical AKM framework, allowing current wages to depend on the worker's complete location history. This feature allows for experience acquired in different locations to have persistent effects on earnings, even after the worker relocates.

Incidental parameter bias. The standard AKM literature has focused on incidental parameter bias arising from the firm (location) side, which occurs when a large number of fixed effects are identified from a small number of movers (Andrews et al., 2008). This leads to imprecise estimates and bias in the variance decomposition. It is equally important to address incidental parameter bias arising from the worker side. Both individual-specific parameters, determining initial wage levels, $\log h_i$, and individual wage growth, $\log a_i$, are high-dimensional relative to the typical panel length. With short panels, the sampling variance in individual slope estimates can substantially inflate the estimated cross-sectional dispersion in learning ability, overstating the true heterogeneity in the population.

In my spatial context, geographic aggregation mitigates this concern: commuting zones represent larger units than firms, increasing the number of movers. To further address this concern and following De La Roca and Puga (2017), I partition Spain's commuting zones into groups based primarily on population size. I introduce (1) Madrid and (2) Barcelona separately, as they host nearly 40% of all workers in my sample. I divide the remaining commuting zones into five partitions of approximately equal size, ordered in terms of population. This partition yields a fully connected mobility graph.

On the worker side, I address this problem by grouping individuals into 100 latent types, following Bonhomme and Manresa (2015). To cluster them, I use their estimated intercepts and returns to experience obtained from an individual fixed-effects regression. This grouped fixed-effects estimator reduces the concerns for incidental parameter bias by pooling information across similar individuals, while preserving the key dispersion in initial human capital and learning ability necessary to discipline the model.

Implementation. Ultimately, I estimate the following regression:

$$w_{i,\ell,t} = \log h_{g(i)} + \log z_\ell + \log a_{g(i)} E_{i,t} + \sum_{c=1}^{7} \log \psi_c E_{i,c,t} + u_{i,\ell,t}, \tag{9}$$

where ℓ denotes each of the commuting zone partitions and $g(i)$ defines the latent type of individual i.

Estimating this specification requires addressing one additional challenge: recovering the location effects $(\log z_\ell, \log \psi_\ell)$ and individual heterogeneity parameters $(\log h_{g(i)}, \log a_{g(i)})$ requires knowledge of the life cycle profile $\{\delta_j\}_{j=25}^{44}$, which determines how experience at different ages contributes to human capital accumulation. However, this profile is not directly observable and must itself be estimated from wage growth data using these same parameters.

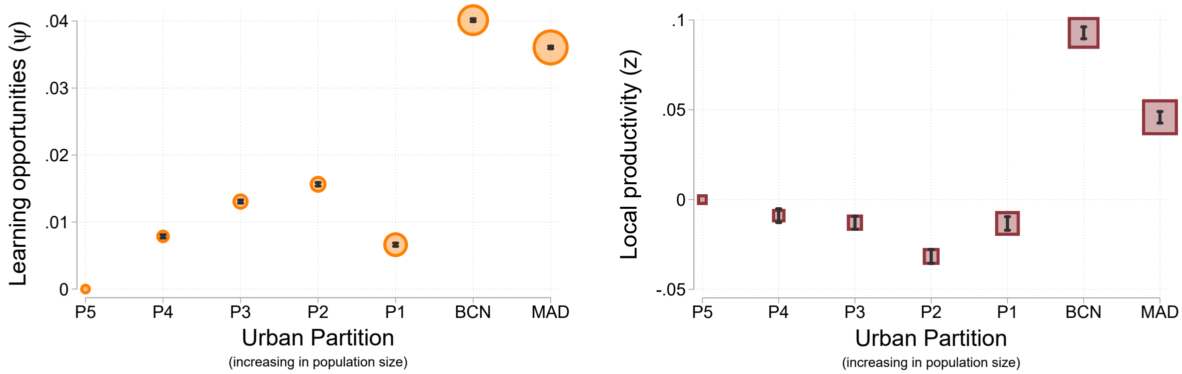
I address this challenge through an iterative procedure that alternates between estimating the regression parameters conditional on a given age profile, and updating the age profile conditional on the recovered parameters. Starting from an initial guess based on observed average wage growth at different ages, I construct age-weighted experience measures and estimate equation (9) using the grouped fixed-effects procedure described above. I then use the recovered parameters to update the age profile via the relationship $\Delta w_{i,j} = \delta_j(\log a_{g(i)} + \log \psi_{\ell(i,j)})$, exploiting wage growth among non-movers. This procedure iterates until convergence. A complete description of the iterative algorithm is provided in Appendix E.1.

3.3.3 Results

The estimated location-specific coefficients associated to equation (9) are presented graphically in Figure 6.

These results indicate that a year of experience is not equally valuable in all locations. For instance, the average year of experience in Madrid and Barcelona raises earnings by 3.7% relative to having worked that same year in a city belonging to the smallest location partition.[17] In Panel (a) we observe that, generally, these gains are

[17] Recall the age profile parameters δ_j have been normalized so that their average is equal to one.



(a) Learning opportunities (ψ_ℓ)

(b) Local productivity (z_ℓ)

Figure 6: Estimated location-specific parameters.

increasing with location size. Panel (b) presents the estimated local productivities, once again with Madrid and Barcelona providing significantly higher returns.

The estimated life cycle profile (presented in Figure 10 in the Appendix) further establishes that these gains are not evenly distributed over the life cycle. Experience acquired at 25 is about 4 times more valuable for human capital accumulation than the same year of experience at age 45.

Finally, Table 1 presents the results linked to the estimated group fixed effects $(\log h_i, \log a_i)$ in the regression that will be used in the quantitative exercise in Section C.

Parameter	Value
$\hat{\mu}_a$	.0141
$\hat{\sigma}_a$	.0297
$\hat{\sigma}_h$	.2656
$\widehat{corr}_{a,h}$	-.3240

Table 1: Distribution parameters

Note: To reduce the amount of noise introduced by outliers, I recover the standard deviations of a and h as $\sigma \approx IQR/1.349$, exploiting the relationship between the inter-quantile range and the standard deviation in normal distributions.

It is worth noticing that, despite using different specifications, these *average* estimates are similar in magnitude to those reported in De La Roca and Puga (2017).

4 Quantitative Spatial Model

The stylized model studied in Section 2 served to highlight the key economic mechanisms through which credit constraints can distort the allocation of workers across space. In the following sections, I quantify these effects. To do this, it is important to first enrich the simplified model in Section 2 so that it better connects with the data. I extend the stylized economy to an overlapping generations setting in which the initial wealth and skill of the young are related to those of the previous generation. I also add additional sources of heterogeneity for both individuals and locations, and allow individuals to choose between renting and owning their home.

4.1 Preliminaries

Time is discrete, indexed by t, and continues forever. There is no aggregate uncertainty. The economy is populated by a continuum of individuals, each living for three periods indexed by $j \in \{1,2,3\}$. Each period lasts 20 years, with the first period corresponding to ages 25-44, the second to ages 45-64, and the third to ages 65-84. At each date, a new cohort of measure one enters the economy. Individuals work in the first two periods of life ($j \in \{1,2\}$) and retire in the last period ($j = 3$).

I focus on steady state equilibrium outcomes. For this reason, whenever possible, I omit time subscripts to lighten notation.

Building on the stylized framework in Section 2, this quantitative model features heterogeneous households who choose consumption, savings, location, and housing tenure (whether to own or rent their home) over their life cycle. During working years, human capital evolves through learning-by-doing. Differences in location-specific learning opportunities, combined with endogenous sorting of learning ability across space, generates spatial variation in wage growth. At retirement, agents make bequests to the next generation, linking skill and wealth across cohorts. Competitive production and construction sectors determine wages and housing prices in each location, while a rental sector intermediates housing subject to frictions. The government finances social security payments and transfers through labor income taxes. I describe each of these components below.

4.2 Households in an OLG Economy

Preferences. As in the stylized economy, agents have CRRA preferences over consumption,

$$u(c) = \frac{c^{1-\frac{1}{\sigma}} - 1}{1 - 1/\sigma},$$

with intertemporal elasticity of substitution $\sigma > 0$.

Life Cycle - Working Age. Agents in the model start making choices at age 25. At that point, they observe their initial state vector $x_1 = (k_1, h_1, a, \ell_1)$. This includes the agent's net wealth, k_1, learning ability, a, and initial location of residence, ℓ_1. Relative to the stylized model in Section 2, I also allow individuals to be heterogeneous in their initial human capital at labor market entry, h_1. This is meant to capture differences in educational attainment, early career experiences, or other factors affecting productivity at age 25. Notice that while k_j, h_j, and ℓ_j are allowed to vary over the life cycle, learning ability is constant over time.

Having observed their state vector, x_j, working-age individuals make their decisions sequentially. At the very beginning of the period, they choose their location of residence, $\ell_{j+1} \in \mathcal{L}$, among a *discrete* number of locations, subject to migration frictions described below. This choice determines their wage, as labor markets are segmented by location. Second, conditional on their chosen residence, agents decide whether to rent their home at price q_ℓ or buy a house at price p_ℓ. In this model, the tenure decision, $o_j \in \{0, 1\}$, is purely financial: homeownership provides access to collateralized borrowing and represents lower discounted housing costs over the life cycle, but requires a substantial upfront payment that many young and asset-poor agents cannot afford. Third, having paid for housing, agents allocate all remaining resources between consumption, c_j, and savings, b_j. At the end of the period, human capital accumulates via learning-by-doing, with accumulation influenced by the agent's chosen location. For notational purposes, it is convenient to define the reduced set $\tilde{x}_j = (k_j, h_j, a)$.

When choosing their location of residence, agents face age-dependent migration frictions. With probability π_j, they can freely choose whether to stay in their current location or move to a new commuting zone. With probability $1 - \pi_j$, they must remain in place, with $\ell_{j+1} = \ell_j$. To focus on the role of credit frictions shaping initial location choices, I assume $\pi_1 = 1$, implying that all young agents are free to move.

This structure implies that working agents solve the following problem:

$$V_j(x_j) = (1-\pi_j)v_j(\tilde{x}_j, \ell_j) + \pi_j \max_n \Big\{v_j(\tilde{x}_j, n)\Big\},$$

where $v_j(\tilde{x}_j, \ell)$ is the ℓ-choice-specific value function,

$$v_j(\tilde{x}_j, \ell) = \max_{c_j, b_j, o_j} \quad u(c_j) + \beta V_{j+1}(\{\tilde{x}_{j+1}, \ell\})$$

$$\text{s.t. } c_j + (1-o_j)q_\ell + o_j p_\ell + b_j = (1-\tau)w_\ell h_j + Rk_j + T_j$$

$$k_{j+1} = b_j + \frac{1}{R} o_j p_\ell (1-d^k) \tag{10}$$

$$h_{j+1} = \psi_\ell h_j a \tag{11}$$

$$o_j \in \{0,1\}, \quad b_j \geq -\lambda_j o_j p_\ell, \quad c_j > 0 \tag{12}$$

Every period, individual resources are composed of labor income net of taxes, $(1-\tau)w_\ell h_j$, net wealth, Rk_j, and government transfers, T_j, which I describe in detail later. These resources are used for consumption, c_j, housing payments, and net financial savings, b_j. Housing costs depend on tenure choice. Renters pay the rental rate q_ℓ each period. Homeowners will purchase their home at price p_ℓ. At the end of the period, they will have to pay a maintenance cost proportional to the value of the house, $d^k p_\ell$.[18] The remaining of their housing equity will be incorporated into their net wealth at the beginning of the following period.[19]

The borrowing constraint in (12) limits borrowing by homeowners to a fraction of the value of their home. Renters, on the other hand, are not allowed to borrow.[20] The choice variable b_j represents financial savings. Along with homeownership status, this decision will determine future net wealth through the law of motion in (10).

While working, human capital evolves according to the learning technology de-

[18]This maintenance cost offsets physical depreciation, ensuring that housing values remain constant over the household's life cycle.

[19]Note that although this specification implicitly assumes homeowners purchase a house every period, embedded in k_j is the housing equity of continuing homeowners. Since in steady state prices are constant, this makes the flow cost of homeownership equal to the per-period maintenance cost.

[20]According to the Spanish Survey of Household Finances in 2017, 85.4% of all outstanding debt held by households was originated for the purchase of real estate. Among the remaining 15%, the main reasons for incurring other debt are the purchase of vehicles and other durable goods, home improvements, debt repayment, and business financing. In the empirical section, I exclude business owners. This, along with the fact that home improvements can also be linked to real estate, means the constraints on borrowing incorporated in the model are empirically grounded.

scribed in (11). As in the stylized model, this specification captures the complementarity between individual learning ability and location-specific learning opportunities. In a later subsection, I describe how the parameters governing this process, ψ_ℓ, can be mapped to the estimates recovered in Section 3.3.

Life Cycle - Retirement. At the end of period 2, individuals retire. During the last period of their life, they receive Social Security payments in lieu of labor income, with benefits ω common to all individuals in the economy.

At this point, their location decisions are determined exogenously according to location-specific migration parameters $\pi_3^{\ell,\ell'}$, denoting the probability that a retiree currently in location ℓ moves to location ℓ'. Location decisions by retirees are by themselves a topic of interest, as can be seen for example in Maroto et al. (2024). I impose this exogeneity assumption to simplify the problem and focus the quantification exercise on the distribution of young skill across space, while still matching a realistic age distribution. The choices of retirees are therefore reduced to a simple consumption-savings problem with a tenure choice.

Retirees die at the end of age 3 with probability one. They derive utility from leaving a bequest to the next generation, which I denote by B. This is the only saving motive among retirees. I assume this takes the form of a warm-glow bequest motive, defined by the function $\Phi(B)$. I show the problem faced by retirees in this economy in Appendix B.

Intergenerational linkages. At each date, a new cohort of individuals enters the economy with an initial state $x_1 = (k_1, h_1, a, \ell_1)$. I now describe the assumptions placed on this initial distribution.

I assume upon death, a *parent* household with state vector x_3^P is replaced by a newborn with state x_1^C, where superscripts P and C denote parent and child generations respectively. Each new agent will be born in the same location in which the parent died, $\ell_1^C = \ell_3^P$, and inherit as net wealth the bequest left by their predecessor, $k_1^C = B(x_3^P)$. Initial human capital and learning ability are drawn from a distribution that depends on the parent's own learning ability. More precisely,

$$\begin{aligned} \log a^C &= \rho_a \log a^P + \varepsilon_a, \qquad \text{with } \varepsilon_a \sim \mathcal{N}(\mu_a, \sigma_a) \\ \log h_1^C &= \rho_h \log a^C + \varepsilon_h, \qquad \text{with } \varepsilon_h \sim \mathcal{N}(\mu_h, \sigma_h) \end{aligned} \tag{13}$$

The parameter ρ_a governs the degree of intergenerational persistence in learning ability. The noise terms ε_a and ε_h are normally distributed with means μ_a and μ_h and standard deviations σ_a and σ_h, capturing both imperfect skill transmission and other factors affecting human capital at labor market entry.

By linking both skill and wealth transmission between parents and children, this generates an endogenous correlation between the key states of the problem. In the stylized economy presented in Section 2, I showed that negative sorting of learning ability across space is a key contributor to aggregate misallocation. Negative sorting, however, only appears among constrained individuals. If high-a agents are more likely to be born with higher inheritances, the mass of individuals with potential for negative sorting will be reduced. This means that allowing for this correlation is important for quantifying the potential for credit-driven misallocation.

4.3 Technology, Markets, and Government Policy

Production. All final goods are produced by representative local firms with a constant returns to scale technology. Relative to the stylized model, I assume firms in each commuting zone have access to an exogenous level of local productivity z_ℓ.

The representative firm in each location produces using labor as its sole input. Defining $\mathcal{H}_\ell$ as total human capital residing in commuting zone ℓ, the representative firm produces output according to the technology

$$Y(\mathcal{H}_\ell) = z_\ell \mathcal{H}_\ell$$

Labor markets are segmented by commuting zone. The location-specific zero profit condition pins down wages as $z_\ell h$.

Housing Markets. A competitive construction sector operates a Leontief technology, using final goods and land owned by absentee landlords who absorb all profits. Every period, a constant mass d^k of houses depreciates and is immediately re-built and sold at price p_ℓ. To maintain constant home values over the life cycle, homeowners must pay a per-period maintenance cost equal to $d^k p_\ell$, fully offsetting the physical depreciation of the dwelling.

A competitive rental sector owns housing units in each location and rents them out to households. Rental companies can frictionlessly buy and sell units on the housing market. When renting out housing units, they incur a rental friction, χ, representing a proportional cost that drives a wedge between rental rates and the user cost of housing. Although I do not specify what this friction is exactly, it can be thought to represent any force distorting price-to-rent ratios, such as tenant protection laws, search costs, or taxes.

Rental companies are subject to the same depreciation as households and need to renew a constant fraction of their housing stock every period. Given these ingredients, the equilibrium rental rate will be a constant fraction of local house prices,

$$q_\ell = \chi\left(1 - \frac{1-d^k}{R}\right)p_\ell.$$

The problem of both the construction and rental firm can be found in Appendix B.

Financial Markets. Agents have access to a risk free bond returning an exogenous gross interest rate R. Relative to the stylized model, I extend borrowing opportunities to allow homeowners to use their home as collateral. I model this as a loan-to-value (LTV) constraint, allowing owners to borrow up to a fraction $\lambda_j \leq 1$ of their home value. Renters, on the other hand, are not allowed to borrow.

Borrowing constraints are age-specific. I assume that $\lambda_j = \lambda$ for $j \in \{1,2\}$, while $\lambda_3 = 1$. This assumption allows full extraction of home equity in the last period, serving as a reverse-mortgage, and ensures agents have no incentive to switch to renting right before death.

Government. The government levies flat taxes on labor income, τ. These resources are used to subsidize social security payments to retirees, ω, and individual transfers designed to ensure a minimum consumption level for all households in the economy.

To prevent strategic use of these transfers, I assume subsidy recipients must be renters and must remain in the same location.[21] Government transfers are therefore defined as the minimum amount of income necessary for agents to reach a minimum

[21]Without these conditions, some agents might be inclined to use governmental transfers to either subsidize homeownership or access locations providing high opportunities for human capital accumulation at little to no cost.

level of consumption, $\underline{c}$.

$$T_j = \max\Big\{0, \mathbb{1}\{\ell_{j+1} = \ell_j\}\mathbb{1}\{o_j = 0\}\Big(\underline{c} - \Big[Rk_j + (1-\tau)w_{\ell_j}h_j - q_{\ell_j}\Big]\Big)\Big\}$$

4.4 Equilibrium

In equilibrium, individuals maximize their expected lifetime utility by choosing their homeownership status, consumption, and savings. They also choose their location of residence while working, and bequests at the end of their life cycle. Final good, construction, and rental firms in each location maximize profits by choosing their corresponding inputs. Prices clear all markets. The government budget constraint is balanced period by period.

I solve for the stationary equilibrium of the economy numerically. Stationarity implies that both prices and the cross-sectional allocations for any given cohort of age j are time-invariant.

5 Model Estimation

I calibrate this model in three steps. A first set of parameters are either fixed externally to standard values in the literature, such as the intertemporal elasticity of substitution, or recovered directly from the data, namely, location transitions when old. Second, I map the estimates recovered in Section 3.3 to the parameters driving individual heterogeneity and returns to skill in the quantitative model. Third, all remaining parameters are estimated internally using the simulated method of moments, conditional on the values recovered in the two previous steps.

5.1 Preliminaries and Externally Calibrated Parameters

The model features $15 + 3\mathcal{L} + \mathcal{L}(\mathcal{L}-1)$ parameters, where $\mathcal{L}$ is the (discrete) number of locations. In this quantification, I set $\mathcal{L} = 2$, which brings this number to 23 parameters. I define a *high opportunity* location (with returns to skill parameters z_H, ψ_H) that will encompass the commuting zones of Madrid and Barcelona. These areas alone host around 30% of the Spanish population. The rest of the country will represent the *low* *opportunity* location.[22]

There are three common preference parameters (β, σ, ϕ), two credit market parameters (R, λ), five housing parameters $(\chi, \tilde{\eta}, f_L, f_H, d^k)$, two parameters representing migration frictions (π_3^{LH}, π_3^{HL}), six parameters characterizing the distribution of inherited skill endowments $(\mu_h, \sigma_h, \rho_h, \mu_a, \sigma_a, \rho_a)$, and five returns to skill parameters $(\omega, z_L, z_H, \psi_L, \psi_H)$.

Nine of these parameters are either set externally to standard values in the literature or recovered directly from observables in the data. These are summarized in Table 2, with additional detail provided in Appendix C.

To briefly highlight some of them, since all borrowing in this economy is linked to housing, I choose the gross interest rate R to match average rates for mortgage loans, adjusted for a 20 year gap to account for the 3-period structure. The inter-temporal elasticity of substitution is set to $\sigma = .5$. As discussed in Section 2, this matches the standard value in the literature. I assume land in the low-opportunity location, f_L, is in excess supply, which under the Leontief assumption fixes its price to a positive constant, $1/\tilde{\eta}$. Land in the high-opportunity location, is chosen to match population shares in the commuting zones of Madrid and Barcelona. The parameter λ is set to match average LTV ratios between 2016 and 2019, once again accounting for the 20-year interval. Finally, pensions are set to match approximately 55% of average earnings to reflect the empirical difference between gross pensions and gross labor costs in Spain, and I recover location transitions among retirees, $\pi_3^{\ell,\ell'}$, using their empirical counterpart in the Household Panel.

The model requires two normalizations. Average labor productivity $(\sum_\ell L_\ell z_\ell)$ cannot be separately identified from the average endowment of human capital in the economy. Similarly, average local learning opportunities $(\sum_\ell L_\ell \psi_\ell)$ cannot be separately identified from average learning ability. As such, I normalize $z_L = 1$ and $\mu_a = 0$.

[22] The regions of Navarra and the Basque Country are excluded from the Household Panel due to particularities in their tax system. I also exclude the cities of Ceuta and Melilla, bordering with Morocco in the northern coast of Africa.

Param	Interpretation	Value	Source
β	Discount factor	.684	$1/R$
σ	Inter-temporal elast. of subst.	0.5	Standard
R	Gross interest rate	1.463	Long-term rate comp. over 20 years
λ	Collateral constraint	.33	Spanish legislation
$\tilde{\eta}$	Construction tech.	4.51	Avg. unit per building ratio
f_L	Low-opportunity land	∞	Excess supply of rural land
f_H	High-opportunity land	.088	Population share in 5 biggest UAs
d^k	Housing depreciation	.131	Resid. 40% land value after 100 years
μ_a	Mean of log learning ability	0.0	Normalization
ρ_a	Intergen. a transmission	.25	Intergen. wage correlation (standard)
π_3^{LH}	Retiree transitions ($L \to H$)	.004	Household Panel
π_3^{HL}	Retiree transitions ($H \to L$)	.017	Household Panel

Table 2: Externally Calibrated Parameters and Normalizations

5.2 Local returns and the distribution of unobserved heterogeneity

To inform the location-specific parameters and the distribution of unobserved heterogeneity, I use the objects recovered from the empirical exercise in Section 3.3. In that section, I obtained a vector of productivity and learning opportunity parameters $\{\hat{z}_\ell, \hat{\psi}_\ell\}_{\ell=1}^{7}$, along with a set of parameters characterizing the distribution of initial human capital and learning ability, $\{\hat{\mu}_a, \hat{\sigma}_a, \hat{\sigma}_h, \hat{\rho}_h\}$, and an income profile $\{\hat{\delta}_j\}_{j=25}^{44}$. I use these parameters to inform their counterparts in the quantitative model, conditional on the selected normalizations $\mu_a = 0$ and $z_L = 1$.

This transformation is not straightforward. The regression in Section 3.3 estimates individual-level heterogeneity and local characteristics using monthly data, whereas a period in the quantitative model in Section 4 represents 20 years. Additionally, the empirical exercise uses a total of 7 location partitions that need to be aggregated into two. To do this, define $\hat{z}_H$ as the population weighted average of the estimates associated to Madrid and Barcelona (partitions 1 and 2), and $\hat{z}_L$ as the population weighted average of the local estimates in all remaining partitions.

Local productivity. I normalize $z_L = 1$. Local productivity in Madrid and Barcelona is therefore $\log z_H = \log \hat{z}_H - \log \hat{z}_L$.

Local learning opportunities. The local parameters ψ_L and ψ_H aggregate 20 years of human capital accumulation into a single period transition. Relative to the estimates from Section 3.3, I perform two types of adjustments.

The first one is to adjust the selected normalization. In the empirical section, $\hat{\mathbb{E}}[\log a_i] = \hat{\mu}_a$ and $\hat{\psi}_7 = 0$, with all other learning opportunity parameters being relative to the outside option. For computational purposes, I wish to normalize $\mu_a = 0$, which requires rescaling ability to zero-mean and shifting all local learning parameters by the same constant:

$$\log \hat{a}_i' = \log \hat{a}_i - \hat{\mu}_a, \qquad \log \hat{\psi}_\ell' = \log \hat{\psi}_\ell + \hat{\mu}_a$$

Note that this only implies a change in the selected normalization, but both the relative gaps in learning ability and the variance of $\log \hat{a}_i$ are unaffected.

Second, having defined $\hat{\psi}_L$ as the weighted sum of these renormalized $\hat{\psi}_\ell'$, I map this new estimate of local learning parameters to its 20-year counterpart. To do that, consider an individual who remains in location ℓ throughout the young period (ages 25-44). From equation (6), their human capital at age 45 is:

$$\log \hat{h}_{i,45} = \log \hat{h}_{i,25} + \bar{\delta} \log \hat{\psi}_\ell + \bar{\delta} \log \hat{a}_i',$$

where $\bar{\delta} = \sum_{s=25}^{44} \hat{\delta}_s$. In the quantitative model in Section 4, this same transition is represented as $\log h_{i,2} = \log h_{i,1} + \log \psi_\ell + \log a_i$. Therefore, the mapping between the parameters in both sections is

$$a_i = (\hat{a}_i')^{\bar{\delta}} \qquad \psi_\ell = (\hat{\psi}_\ell)^{\bar{\delta}}$$

In practice, I compute these objects using average (age weighted) experience for workers in the sample at age 44 to estimate $\bar{\delta}$. Although I could've used the sum of the estimated age-profile coefficients, $\sum_{s=25}^{44} \hat{\delta}_s = 20$, I do this to better match the empirical wage growth over the life cycle. Using the latter would only increase the relevance of the local learning mechanism in the calibrated economy.

Skill distribution parameters. According to the relationships described in equation (13), learning ability follows an intergenerational AR(1) process. In steady state, the stationary distribution of log learning ability is:

$$\log a \sim \mathcal{N}(m_a, s_a^2), \qquad m_a = \frac{\mu_a}{1-\rho_a}, \quad s_a^2 = \frac{\sigma_a^2}{1-\rho_a^2}, \tag{14}$$

The (unconditional) stationary distribution of $\log h_1$ will take the form

$$\log h_1 \sim \mathcal{N}(m_h, s_h^2), \qquad m_h = \rho_h m_a + \mu_h, \quad s_h^2 = \rho_h^2 s_a^2 + \sigma_h^2. \tag{15}$$

The estimated group fixed effects can be used to recover information on the parameters behind these distributions. Keeping in mind the previous transformation,

$$\hat{s}_a^2 = \text{Var}[\log a_i] = \bar{\delta}^2 \text{Var}[\log \hat{a}_i'].$$

All other parameters can be directly recovered without the need to perform any additional transformation. More precisely, from the estimated group fixed effects, I compute $\hat{s}_h = \sqrt{\text{Var}[\log \hat{h}_{i,25}]}$, along with their correlation $\widehat{corr}_{a,h} = \text{Corr}[\log \hat{a}_i', \log \hat{h}_{i,25}]$.[23] Given the normalization $\mu_a = 0$ and the externally calibrated value of ρ_a, I recover the innovation variances and the transmission coefficient from learning ability to initial human capital as:

$$\sigma_a = \hat{s}_a \sqrt{1-\rho_a^2}, \qquad \rho_h = \widehat{corr}_{a,h} \frac{\hat{s}_h}{\hat{s}_a}, \qquad \sigma_h^2 = \hat{s}_h^2 - \rho_h^2 \hat{s}_a^2.$$

Table 3 presents the final mapping from regression objects to model parameters.

5.3 Method of Moments Estimation

As the last step in calibration, I estimate the remaining four parameters $(\mu_h, \chi, \phi, \pi_2)$, using the simulated method of moments.

Notice that although I have previously recovered estimates for the distribution of initial human capital in section 3.3, the *level* associated to this distribution is by itself

[23] I compute these moments using only 25-year-old individuals to ensure both dispersion measures correspond to the same cohort. According to the model, these distributions are time-invariant. In practice, however, the estimated average initial human capital is increasing in age. This is to be expected, as the 4-year labor force attachment criteria imposed in sample selection imposes a gradually smaller restriction as potential working life length increases.

Param	**Interpretation**	**Value**
z_L	Returns to	1.076
z_H	human capital	1.000
ψ_L	Local opportunities for	1.514
ψ_H	human capital accumulation	2.789
σ_a	Dispersion of a innovation	0.409
σ_h	Dispersion of h_0 innovation	0.251
ρ_h	Leaning ability to h_0 transmission	-0.204

Table 3: Model-based regression estimates.

Note: These parameters have been transformed as described in section 5.2.

meaningless. The parameter μ_h still needs to be calibrated to ensure that the average wages in the economy match the level of all other relevant quantities in the model.

To inform these moments, I target (1) median net wealth owned by young agents between 25 and 30 years of age, as a share of median house prices, (2) the ratio of median earnings to median house prices, (3) the average rental market share in the economy, and (4) the share of working agents switching commuting zones at least once between 45 and 64. I provide further information on how I recover these moments from the data in Appendix C.

Although these parameters jointly inform all four moments, it is useful to briefly discuss how each of them guides the calibration. Rental market shares are disproportionately influenced by the parameter governing rental market frictions (χ), driving a wedge between tenure choices. The parameter ϕ governs the bequest motive, and therefore the average net wealth inherited by each cohort. I tie this statistic to median housing prices to ensure that all components of the model are in comparable units. Similarly, conditional on the recovered estimates for local productivity, the parameter μ_h mostly determines the mean of the earnings distribution relative to median house prices. Finally, π_2 determines whether middle-aged agents are allowed to move. Higher values will therefore increase increase the share of commuting zone switchers.

Table 4 displays the results of this final step. Despite the non-linearity embedded in the model, all moments are matched exactly.

Param	Interpretation	Moment	Value
μ_h	Avg. initial human capital	Avg. wage to housing expenditure ratio	.0138
χ	Rental friction	Rental market share	1.499
ϕ	Bequest motive	Net wealth of young to med. house price	.001
π_2	Migration friction	Share of movers aged 45 to 64	.247

Table 4: Internally Calibrated Parameters

6 Counterfactuals

6.1 Estimates of Misallocation and Credit-Driven Distortions

Efficient Allocation. To quantify the output losses derived from credit market imperfections, I compare the competitive equilibrium to the constrained efficient allocation of a social planner that can freely transfer consumption over time and across space, but is subject to the same migration frictions, technology, and capacity constraints as the competitive equilibrium.

In order to focus on the key mechanism of the paper, I also assume that the planner must respect the age distribution present in equilibrium. That is, the planner can reshuffle young agents across space, but cannot distort the overall age composition within a given location. In the absence of this last condition, and given the estimated local parameters, the planner would find it optimal to concentrate as many young agents as possible in Madrid and Barcelona. While the spatial age composition is itself of interest, this model is targeted to study sorting, and for that reason I abstract from this. Appendix B presents the planner problem.

Using the calibrated model, I find that credit constraints induce losses equivalent to 6.89% of aggregate output. These output losses are driven by inefficient city composition among young agents, which are in turn caused by credit-related distortions on individual location decisions. Inefficient sorting of skill across space generates inefficient human capital accumulation.

6.2 The Role of Negative Sorting in Learning Ability

One of the most common counterfactuals in the literature on spatial misallocation considers the relaxation of land based policies, aiming to increase the size of specific cities. As was previously mentioned, all of the losses in my model arise from the *composition*, rather than the *size* of cities. Nevertheless, and in order to highlight the role of negative sorting and its implications for spatial policy, I consider a land expansion exercise that increases the land area available in Madrid and Barcelona.

I use the model to answer the following question: how much additional land, f_H, would Madrid and Barcelona need to have for the equilibrium economy to reach the level of aggregate human capital that we would observe under the initially efficient distribution of labor?

To compute this, I gradually increase the amount of land in these two cities until the economy reaches the same level of period 2 human capital as the efficient allocation in the previous subsection. Note that, as I expand the size of Madrid and Barcelona, the efficient distribution of labor (and therefore aggregate human capital) also changes. Yet, this exercise compares the equilibrium levels of human capital to that in the *initially* efficient distribution of labor.

Given the calibrated parameters, I find that the commuting zones of Madrid and Barcelona would require a 52.8% increase in available land in order to reach the same level of human capital that the economy would achieve in the absence of credit frictions. In population terms, under this policy Madrid and Barcelona would host 47.1% of the Spanish population versus the originally calibrated 30.8%.

The magnitude of this difference is driven by two main factors. First, the new land in equilibrium will not only be populated by previously misallocated young agents, but is shared across the entire age distribution. Second, negative sorting in learning ability implies that the marginal agent that moves from the low- to high-productivity area as new land is introduced has low learning ability, relative to other agents that choose to remain in the location offering worse learning opportunities.

To reflect this last point, Figure 7 gradually swaps the highest $(a \times h)$ agents living in low-productivity areas with the lowest $(a \times h)$ living in Madrid and Barcelona, computing the share of aggregate output losses that would be recovered with each swap. What we can see here is that reallocating 25% (50%) of the initially misallocated agents would already return 52% (80%) of the output losses. Misallocation is therefore driven by a relatively small mass of individuals with high learning potential.

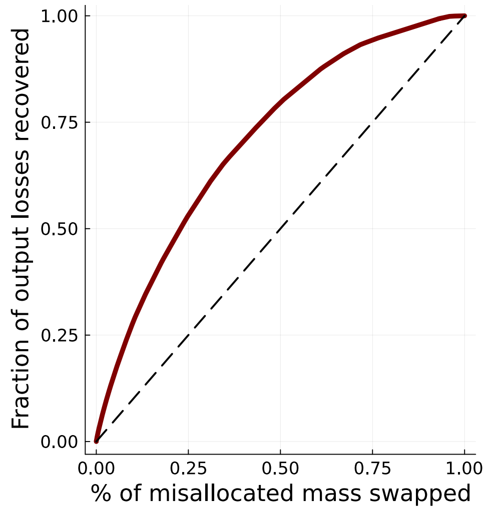


Figure 7: Recovery curve

6.3 Second-Best Policies: Targeting the Marginal Agent

So far this section has attempted to highlight the critical role of negative sorting in spatial misallocation. Although the optimal policy in this model would directly relax borrowing frictions, I now consider the type of second-best spatial and housing policies that are better abled to target the relevant source of misallocation.

To do this, I consider two well known types of housing policies–homeownership subsidies and rental market policy–and identify the characteristics of the marginal agent responding to each of them. More precisely, I consider a marginal decrease in p_H and q_H separately,[24] representing the housing price and rental market rate in Madrid and Barcelona. I identify the agents that would move in response of each of these policies, and plot their learning ability distribution in Figure 8.

From this Figure we can see that a marginal decrease in house prices achieved through a homeownership subsidy, attracts individuals with relatively low learning ability to the high-opportunity location. The opposite is true in the case of rental policy.

The reason for this is that the existence of rental market frictions in the calibrated economy implies that only constrained agents will use rental markets. Because agents with high learning ability are more likely to be constrained, this means that rental

[24]The counterfactual exercise considers a 1% decrease in the equilibrium prices of Madrid and Barcelona, and stores the characteristics of those agents that move from low- to high-productivity areas in response to the policy.

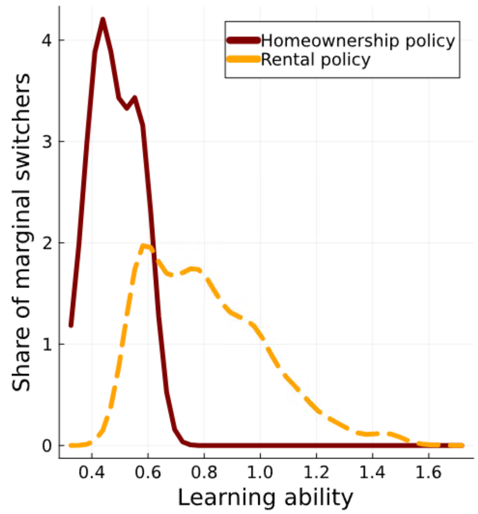


Figure 8: Learning ability dist. Marginal movers in response to local housing policy.

policy is able to attract those agents that are most critical for aggregate human capital accumulation to the high productivity location.

This exercise highlights how, in the presence of negative sorting, it becomes critical to design policies that can better target the relevant source of misallocation.

7 Conclusion

This paper studies how credit constraints distort location decisions and, consequently, the spatial allocation of human capital. I show that when locations differ in the learning opportunities they offer and individuals are heterogeneous in their learning ability, credit frictions not only weaken positive sorting of ability across space, but under empirically relevant conditions generate *negative* sorting among those individuals that are credit-constrained.

Using administrative data from Spain linking wealth, income, and complete working histories, I provide evidence consistent with this mechanism. I document substantial spatial variation in wage growth, particularly early in workers' careers, and show that low-wealth individuals experience faster subsequent wage growth than wealthier neighbors residing in the same location. This pattern is consistent with credit-constrained high-ability workers being systematically allocated to suboptimal locations. I implement an extended AKM framework to separately identify location-specific learning opportunities and individual-specific learning ability, recovering the complementarity between these factors that drives the theoretical results.

I embed these mechanisms in a quantitative spatial model calibrated to Spain and find that credit constraints reduce aggregate output by 6.3% through inefficient sorting of skill across space. Importantly, these losses arise entirely from distortions in the *composition* rather than the *size* of productive cities. The presence of negative sorting has important implications for spatial policy, as standard place-based policies aimed at expanding the size productive cities through land use deregulation are less effective than would be expected in the absence of negative sorting. I demonstrate that effective spatial policy must instead target constrained high-ability workers. Comparing homeownership subsidies to rental market interventions, I show that rental subsidies disproportionately attract high-ability individuals to productive cities. This occurs because rental markets are used predominantly by credit-constrained agents in the calibrated economy, providing an indirect mechanism to target the population most affected by credit-driven misallocation.

These findings contribute to our understanding of the forces shaping human capital accumulation over the life cycle and highlight the critical role of housing markets in mediating access to opportunity. They underscore that in dynamic spatial economies, both the size and composition of cities matter for aggregate outcomes, and that effective policy design requires careful attention to which workers are induced to move in response to different types of policy interventions.

References

Abowd, J. M., F. Kramarz, and D. N. Margolis (1999). High Wage Workers and High Wage Firms. *Econometrica 67*(2), 251–333. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/1468-0262.00020.

Andrews, M. J., L. Gill, T. Schank, and R. Upward (2008). High Wage Workers and Low Wage Firms: Negative Assortative Matching or Limited Mobility Bias? *Journal of the Royal Statistical Society Series A: Statistics in Society 171*(3), 673–697.

Becker, G. S. (1973). A Theory of Marriage: Part I. *Journal of Political Economy 81*(4), 813–846. Publisher: University of Chicago Press.

Benabou, R. (1993). Workings of a City: Location, Education, and Production. *The Quarterly Journal of Economics 108*(3), 619–652. Publisher: Oxford University Press.

Bick, A., A. Blandin, and R. Rogerson (2024). Hours Worked and Lifetime Earnings Inequality. *NBER Working Paper #32997*.

Bilal, A. and E. Rossi-Hansberg (2021). Location as an Asset. *Econometrica 89*(5), 2459–2495.

Blandin, A. (2018, May). Learning by Doing and Ben-Porath: Life-cycle Predictions and Policy Implications. *Journal of Economic Dynamics and Control 90*, 220–235.

Bonhomme, S. and E. Manresa (2015). Grouped Patterns of Heterogeneity in Panel Data. *Econometrica 83*(3), 1147–1184.

Crews, L. (2024). A Dynamic Spatial Knowledge Economy. *Working Paper*.

De La Roca, J. and D. Puga (2017). Learning by Working in Big Cities. *The Review of Economic Studies 84*(1), 106–142.

Diamond, R. (2016, March). The Determinants and Welfare Implications of US Workers' Diverging Location Choices by Skill: 1980-2000. *American Economic Review 106*(3), 479–524.

Diamond, R. and C. Gaubert (2022, August). Spatial Sorting and Inequality. *Annual Review of Economics 14*(Volume 14, 2022), 795–819. Publisher: Annual Reviews.

Díaz, A., A. Jáñez, and F. Wellschmied (2023, February). Geographic Mobility Over the Life-cycle. *ICAE Working Paper #2301*.

Eckert, F. and T. Kleineberg (2024). The Geography of Opportunity: Education, Work, and Intergenerational Mobility Across US Counties. *Working Paper*.

Fajgelbaum, P. D. and C. Gaubert (2020, May). Optimal Spatial Policies, Geography, and Sorting*. *The Quarterly Journal of Economics 135*(2), 959–1036.

Favilukis, J., S. C. Ludvigson, and S. Van Nieuwerburgh (2017, February). The Macroeconomic Effects of Housing Wealth, Housing Finance, and Limited Risk Sharing in General Equilibrium. *Journal of Political Economy 125*(1), 140–223. Publisher: The University of Chicago Press.

Fernandez, R. and R. Rogerson (1996). Income Distribution, Communities, and the Quality of Public Education. *The Quarterly Journal of Economics 111*(1), 135–164. Publisher: Oxford University Press.

Fernández, R. and R. Rogerson (1998). Public Education and Income Distribution: A Dynamic Quantitative Evaluation of Education-Finance Reform. *The American Economic Review 88*(4), 813–833. Publisher: American Economic Association.

Fogli, A., V. Guerrieri, M. Ponder, and M. Prato (2024). Scaling Up the American Dream: A Dynamic Analysis. *Working Paper*.

Galichon, A. (2016). Chapter 3: One-Dimensional Case. In *Optimal Transport Methods in Economics*, pp. 34–44. Princeton University Press.

Giannone, E., Q. Li, N. Paixao, and X. Pang (2023). Unpacking Moving: A Quantitative Spatial Equilibrium Model with Wealth. *Working Paper*.

Greaney, B. (2023). Homeownership and the Distributional Effects of Uneven Regional Growth. *Working Paper*.

Greaney, B., A. Parkhomenko, and S. Van Nieuwerburgh (2025). Dynamic Urban Economics. *Working Paper*.

Gregory, V. (2023). Firms as Learning Environments: Implications for Earnings Dynamics and Job Search. *Working Paper*.

Guvenen, F., G. Kaplan, J. Song, and J. Weidner (2022, October). Lifetime Earnings in the United States over Six Decades. *American Economic Journal: Applied Economics 14*(4), 446–479.

Havránek, T. (2015, December). Measuring Intertemporal Substitution: The Importance of Method Choices and Selective Reporting. *Journal of the European Economic Association 13*(6), 1180–1204.

Heathcote, J., K. Storesletten, and G. L. Violante (2014, July). Consumption and Labor Supply with Partial Insurance: An Analytical Framework. *American Economic Review 104*(7), 2075–2126.

Huggett, M., G. Ventura, and A. Yaron (2006, March). Human capital and earnings distribution dynamics. *Journal of Monetary Economics 53*(2), 265–290.

Huggett, M., G. Ventura, and A. Yaron (2011, December). Sources of Lifetime Inequality. *American Economic Review 101*(7), 2923–2954.

Imai, S. and M. P. Keane (2004). Intertemporal Labor Supply and Human Capital Accumulation. *International Economic Review 45*(2), 601–641. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1468-2354.2004.00138.x.

Kaplan, G., K. Mitman, and G. L. Violante (2020, September). The Housing Boom and Bust: Model Meets Evidence. *Journal of Political Economy 128*(9), 3285–3345. Publisher: The University of Chicago Press.

Keane, M. P. and K. I. Wolpin (1997). The Career Decisions of Young Men. *Journal of Political Economy 105*(3), 473–522. Publisher: The University of Chicago Press.

Kleinman, B., E. Liu, and S. J. Redding (2023). Dynamic Spatial General Equilibrium. *Econometrica 91*(2), 385–424.

Lhuillier, H. (2024). Should I Stay or Should I Grow? How Cities Affect Learning, Inequality and Aggregate Productivity. *Working Paper*.

Lochner, L. and A. Monge-Naranjo (2012, September). Credit Constraints in Education. *Annual Review of Economics 4*, 225–256.

Lochner, L. J. and A. Monge-Naranjo (2011, October). The Nature of Credit Constraints and Human Capital. *American Economic Review 101*(6), 2487–2529.

Luccioletti, C. (2023). Should Governments Subsidize Homeownership? *Working Paper*.

Maroto, M. B., B. Faber, A. Levy, and M. Munoz (2024). Senior Migration, Local Economic Development and Spatial Inequality. *Working Paper*.

Oswald, F. (2019). The effect of homeownership on the option value of regional migration. *Quantitative Economics 10*(4), 1453–1493.

Shimer, R. and L. Smith (2000). Assortative Matching and Search. *Econometrica 68*(2), 343–369. Publisher: [Wiley, Econometric Society].

Thimme, J. (2017). Intertemporal Substitution in Consumption: A Literature Review. *Journal of Economic Surveys 31*(1), 226–257. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/joes.12142.

Topkis, D. M. (1998). *Supermodularity and Complementarity*. Princeton University Press.

A Mathematical Appendix

This Appendix contains the proofs of all Propositions and Lemmas stated in the main text. It characterizes sorting patterns using global arguments based on monotone comparative statics. These results establish that, whenever an equilibrium exists, whether better learners sort into better locations depends on (i) their constrained status and (ii) the intertemporal elasticity of substitution.

Recall that given any price schedule $p(\ell)$, wage w, and interest rate R, an individual with initial states (k, a) solves

$$\max_{c_1, c_2, b, \ell} u(c_1) + \beta u(c_2) \tag{16}$$

$$\text{s.t. } c_1 = k + wh_1 - b - p(\ell)$$

$$c_2 = wh_2 + Rb, \quad h_2 = \ell h_1 a$$

$$b \geq -\underline{b} \qquad (\mu)$$

where $\underline{b} \geq 0$ and $\mu(k, a)$ is the Lagrange multiplier associated with the borrowing constraint.

A.1 Proofs for the Benchmark Economy

Substituting the value of consumption in Problem 16 and taking first-order conditions with respect to bond savings, we obtain

$$-u'(c_1) + \beta R u'(c_2) + \mu(k, a) = 0$$

In the benchmark economy, I set $\underline{b} \to \infty$ so that the credit constraint never binds. This means the Lagrange multiplier satisfies $\mu(k, a) = 0 \ \forall (k, a)$ and using the CRRA assumption, we obtain

$$c_1 = (\beta R)^{-\sigma} c_2. \tag{17}$$

Aggregating the budget constraint to obtain the present value of lifetime resources,

$$\begin{aligned} c_1 + \frac{c_2}{R} &= k + wh_1 + \frac{w}{R}\ell h_1 a - p(\ell) \\ &\equiv \mathcal{I}_\ell(k,a) \end{aligned} \tag{18}$$

where I define $\mathcal{I}_\ell(k,a)$ as the discounted lifetime income, net of housing costs, of an individual with initial states (k,a) and choosing to live in location ℓ.

Combining the relationships in equations (17) and (18), first period consumption can be written as a function of $\mathcal{I}_\ell(k,a)$. This also implies that we can write the indirect utility function $V_\ell^U(k,a)$, such that

$$\begin{aligned} V_\ell^U(k,a) &\equiv u(c_1) + \beta u(c_2) \\ &= u(c_1) + \beta u(c_1(\beta R)^\sigma) \\ &= \left(1 + \beta(\beta R)^{\tilde{\sigma}}\right) u(c_1) \\ &= \left(1 + \beta(\beta R)^{\tilde{\sigma}}\right) u\left(\kappa \mathcal{I}_\ell(k,a)\right) \end{aligned} \tag{19}$$

where κ and $\tilde{\sigma}$ are constants.

The indirect utility function $V_\ell^U(k,a)$ is strictly increasing in $\mathcal{I}_\ell(k,a)$. Therefore for any (k,a), maximizing $V_\ell^U(k,a)$ is equivalent to choosing the location that maximizes discounted lifetime income, net of housing costs. This implies Problem (16) reduces to one in which agents simply choose the location that maximizes discounted lifetime income net of housing costs,

$$\max_{\ell \in [\underline{\ell}, \bar{\ell}]} \quad k + wh_1 + \frac{1}{R} w\ell h_1 a - p(\ell). \tag{20}$$

Lemma 1. The equilibrium price schedule $p^*(\ell)$ is increasing in ℓ.

Proof. Suppose, by contradiction, that $p^*(\ell)$ were decreasing and take any $\ell < \ell'$ such that $p^*(\ell') \leq p^*(\ell)$. For any $a > 0$,

$$\mathcal{I}_{\ell'}(k,a) - \mathcal{I}_\ell(k,a) = \frac{w}{R} h_1 a(\ell' - \ell) - [p^*(\ell') - p^*(\ell)] > 0,$$

because $\ell' > \ell$, $a > 0$, and $w, R, h_1 > 0$. This means $V_{\ell'}^U(k,a) > V_\ell^U(k,a)$ for all k and all

$a > 0$, so no agent strictly prefers ℓ.

By assumption, the housing supply distribution $g(\ell)$ is continuous and strictly positive on $[\underline{\ell}, \bar{\ell}]$, so that every location has strictly positive housing supply. I also assumed that total housing space is equal to total population. If no agent chooses ℓ in equilibrium, it must be that another location $\ell'' \in [\underline{\ell}, \bar{\ell}]$, is in excess demand. This means the local housing market at ℓ'' does not clear, contradicting the equilibrium condition. Therefore, in any equilibrium where all locations are populated, we must have $p^*(\ell') > p^*(\ell)$ whenever $\ell' > \ell$. □

Proposition 1. For any increasing price schedule $p(\ell)$, location choices in the benchmark economy, $l^*(a)$, are:

1. Independent of initial endowments (k).
2. Increasing in learning ability (a).

Independence from initial endowments. In Problem (20), k is an additive constant and does not affect the argmax over ℓ. Therefore, location choices are independent of initial endowments.

Increasing in learning ability. Let the choice set be a complete lattice $\mathcal{L} = [\underline{\ell}, \bar{\ell}]$, and consider the parameter $a \in [\underline{a}, \bar{a}]$, both with the usual order on $\mathbb{R}$. Define

$$f(\ell, a) \equiv \frac{w}{R} a h_1 \ell - p(\ell),$$

where I drop the additive constant $k + w h_1$, as it does not affect the arg max. I now show that $f(\ell, a)$ has increasing differences in ℓ and a in order to make use of Topkis' theorem.

For any $\ell < \ell'$ and $a < a'$,

$$\begin{aligned} f(\ell', a') + f(\ell, a) - f(\ell', a) - f(\ell, a') \\ &= \frac{w h_1}{R}(a'\ell' + a\ell - a\ell' - a'\ell) \\ &= \frac{w h_1}{R}(a' - a)(\ell' - \ell) > 0 \end{aligned}$$

Hence, f has increasing differences in (ℓ, a).

Suppose for each a the set $\Phi(a) \equiv \arg\max_{\ell \in [\underline{\ell}, \bar{\ell}]} f(\ell, a)$ is nonempty. Then, by Topkis'

Monotonicity Theorem, Φ is non-decreasing in the strong set order; in particular, there exists a non-decreasing selection $l^*(a) \in \Phi(a)$. If in addition the maximizer is unique and interior, then $l^*(a)$ is single-valued and strictly increasing in a. □

The arguments above are conditional on the set of optimal locations $\Phi(a)$ being nonempty for each a. What they illustrate is that, whenever an equilibrium exists, location decisions in the benchmark economy must be non-decreasing in a.

Although in the main text I assume the learning technology $\psi(a,\ell,h_1)$ to be log-linear, the argument extends to any learning technology that is supermodular in (a,ℓ). In that case, future labor income net of housing costs can be written as $\tilde{f}(a,\ell) = \frac{w}{R}\psi(\ell,a,h_1) - p(\ell)$. Supermodularity of $\psi(\cdot)$ in (a,ℓ) implies that $\tilde{f}$ has increasing differences in (a,ℓ), so the same Topkis argument delivers that optimal location decisions are non-decreasing in learning ability.

Finally, and looking ahead to the economy with possibly binding credit constraints, it is useful to prove that optimal savings in the benchmark (or unconstrained) economy are decreasing in learning ability.

Proposition 1b. For any increasing price schedule $p(\ell)$, optimal savings in the benchmark economy, $b^*(a,k)$, are:

1. Increasing in initial endowments (k).
2. Decreasing in learning ability (a).

Proof. We previously concluded that optimal period 1 consumption is a constant fraction of lifetime income net of housing costs:

$$c_1^*(a,k) = \theta \cdot \mathcal{I}_{l^*(a)}(a), \qquad \text{where } \theta \equiv \frac{1}{1+\beta^\sigma R^{\sigma-1}} \in (0,1).$$

Net savings $b^*(a,k)$ are defined as period 1 income net of consumption and housing costs. Substituting optimal consumption from the expression above, we obtain

$$b^*(a,k) = (1-\theta)\left[k + wh_1 - p(l^*(a))\right] - \frac{\theta}{R} w\ell a h_1. \tag{21}$$

From the closed-form presented above, it is easy to see that $b^*(a,k)$ is increasing in k. Since $l^*(a)$ is increasing in a by Proposition 2 and $p(\ell)$ is increasing, we also conclude that $b^*(a,k)$ is decreasing in a. □

This Proposition in turn helps us to characterize the set of constrained agents whenever $\underline{b} \geq 0$ is finite and binds in equilibrium.

LEMMA 2. For any increasing price schedule $p(\ell)$, there exists a threshold function $a^*(k)$ such that agents with learning ability $a > a^*(k)$ are credit constrained, while agents with $a \leq a^*(k)$ are unconstrained. Additionally, $a^*(k)$ is increasing in k.

Proof. An agent is constrained if and only if desired net savings exceed the credit limit ($b^*(a,k) \geq -\underline{b}$). Since $b^*(a,k)$ is monotonically increasing in a by Proposition 1b, the set of constrained agents $\mathcal{A}^C$ forms an upper set. In other words, there exists a unique threshold $a^*(k)$ such that agents with learning ability $a > a^*(k)$ are credit constrained, while agents with $a \leq a^*(k)$ are unconstrained. This threshold is implicitly defined by the condition

$$b^*(a^*(k),k) = -\underline{b}. \tag{22}$$

By the Implicit Function Theorem and using the properties in Proposition 1b:

$$\frac{da^*(k)}{dk} = -\frac{\partial b^*/\partial k}{\partial b^*/\partial a} > 0.$$

□

A.2 Properties of the Planner's problem

Consider the planner's problem described in Section 2, where the planner chooses a joint distribution $\pi(a,\ell)$ to solve

$$\max_{\pi(a,\ell)} \int_{\mathcal{A}} \int_{\mathcal{L}} a\ell \, \pi(a,\ell) \, d\ell \, da \tag{23}$$

subject to the constraints that the marginal distributions of a and ℓ are equal to the given densities $\lambda_a(a)$ and $g(\ell)$, respectively:

$$\int_{\mathcal{L}} \pi(a,\ell) \, d\ell = \lambda_a(a) \quad \forall a,$$
$$\int_{\mathcal{A}} \pi(a,\ell) \, da = g(\ell) \quad \forall \ell,$$

where $\mathcal{A}$ and $\mathcal{L}$ denote the supports of the ability and local housing distributions, respectively.

The planner's problem in (23) can be seen as an optimal transport problem with surplus function $s(a,\ell) = a\ell$ and given marginal distributions for a and ℓ. Since the surplus function is strictly supermodular and there are no mass points in the ability or housing distributions, the efficient allocation is unique and features perfect positive assortative matching (Becker, 1973; Shimer and Smith, 2000; Galichon, 2016). That is, any solution to the planner's problem assigns higher-ability agents to higher-opportunity locations.

This implies that, under the assumptions of the benchmark economy, any competitive equilibrium replicates the planner's efficient allocation of labor across space.

A.3 Proofs for the Economy with Binding Credit Constraints

Consider an economy in which $\underline{b} \geq 0$ is finite so that credit constraints bind for a positive mass of agents. This section contains the proofs to Lemma 2, Proposition 2, and Corollary 1, characterizing how location choices vary with learning ability given an increasing price schedule $p(\ell)$ and, in particular, the equilibrium one, as stated in Proposition 3.

Let $\hat{l}(a,k)$ denote the location choice for an agent solving Problem 16. I analyze individual decisions in three steps. First, I characterize the solution to a *relaxed* problem where the constraint is ignored, matching the benchmark economy in the previous section, with location decisions denoted as $l^U(a)$. Second, I characterize the solution when the constraint binds strictly, with location decisions denoted $l^C(a,k)$. Third, I establish that the set of constrained agents forms an upper set in ability ("no re-entry" condition).

Step 1: Unconstrained location choices, $l^U(a)$, are increasing in a.

This is immediate from the results in the previous section. Let $V_\ell^U(a,k)$ denote the indirect utility of an unconstrained agent in location ℓ. As shown in equation (19) for the benchmark economy, this function is strictly increasing in lifetime income $\mathcal{I}_\ell(a,k)$. Consequently, the unconstrained optimal location $l^U(a) = \arg\max_\ell V_\ell^U(a,k)$ is equivalent to the income-maximizing location choice.[25] Since the learning technology is supermodular in (a,ℓ), the objective function satisfies increasing differences in (a,ℓ). By Topkis' Monotonicity Theorem, the unconstrained location choice $l^U(a)$ is increasing in a. $\square$

Step 2: Constrained location choices, $l^C(a,k)$, are decreasing in a if IES < 1.

When the credit constraint binds in Problem 16 ($b = -\underline{b}$), consumption is given by

$$\hat{c}_1(\ell) = k + wh_1 + \underline{b} - p(\ell),$$
$$\hat{c}_2(\ell,a) = w(\ell a h_1) - R\underline{b}.$$

Let $V^C_\ell(a)$ be the payoff of a constrained agent with ability a (and the omitted endowment k) that chooses to live in location ℓ:

$$V^C_\ell(a) \equiv u\Big(\hat{c}_1(\ell,a)\Big) + \beta u\Big(\hat{c}_2(\ell,a)\Big).$$

Define the intertemporal elasticity of substitution as $\eta(c) \equiv -\frac{u'(c)}{c\,u''(c)}$. I first show that, if $\eta(c) \leq 1$, then $V^C_\ell(a)$ has *decreasing differences* in (ℓ,a). Consider $\ell < \ell'$ and $a < a'$. We want to show that

$$V^C_{\ell'}(a') + V^C_\ell(a) \leq V^C_{\ell'}(a) + V^C_\ell(a')$$

Operating with this expression,

$$u\left(k + wh_1 + \underline{b} - p(\ell')\right) + \beta u\left(w\ell' a' h_1 - R\underline{b}\right) + u\left(k + wh_1 + \underline{b} - p(\ell)\right) + \beta u\left(w\ell a h_1 - R\underline{b}\right) \leq$$
$$u\left(k + wh_1 + \underline{b} - p(\ell')\right) + \beta u\left(w\ell' a h_1 - R\underline{b}\right) + u\left(k + wh_1 + \underline{b} - p(\ell)\right) + \beta u\left(w\ell a' h_1 - R\underline{b}\right),$$

which reduces to

$$\beta u\left(w\ell' a' h_1 - R\underline{b}\right) + \beta u\left(w\ell a h_1 - R\underline{b}\right) \leq \beta u\left(w\ell' a h_1 - R\underline{b}\right) + \beta u\left(w\ell a' h_1 - R\underline{b}\right).$$

This implies that decreasing differences of $V^C(\cdot)$ reduces to decreasing differences

[25] Note that I define unconstrained location choices as $l^U(a)$ in order to clearly separate the constrained and unconstrained problems. However, this is the same object as the location choice in the benchmark economy, previously defined as $l^*(a)$. This is also the reason why $l^U(a)$ does not depend on k, as we previously showed that unconstrained location decisions are independent of initial endowments.

in the term $U_2^C \equiv \beta u(w\ell' a' h_1 - R\underline{b})$. Because u is twice continuously differentiable, decreasing differences are equivalent to a negative cross-partial derivative.[26] Operating in steps, we obtain

$$\begin{aligned}
\frac{\partial U_2^C}{\partial \ell} &= u'(w\ell' a' h_1 - R\underline{b}) w h_1 a \\
\frac{\partial^2 U_2^C}{\partial \ell \partial a} &= u'(c_2) w h_1 + u''(c_2) w h_1 a \cdot w \ell h_1 \\
&= u'(c_2) w h_1 \left[1 + \frac{u''(c_2)}{u'(c_2)} w \ell a h_1 \right] \\
&= u'(c_2) w h_1 \left[1 - \frac{1}{\eta(c_2)} \frac{w \ell a h_1}{w \ell a h_1 - R\underline{b}} \right] \\
&\leq u'(c_2) w h_1 \left[1 - \frac{1}{\eta(c_2)} \right],
\end{aligned}$$

where the third line of the cross-partial derivative uses the definition of the intertemporal elasticity of substitution, and the fourth line uses the fact that $\underline{b} \geq 0$.

It follows that whenever $\eta(c) \leq 1$, the term in brackets is negative and thus we have $\frac{\partial^2 U_2^C}{\partial a \partial \ell} < 0$. This in turn implies that V^C has decreasing differences in (ℓ, a). By Topkis' Monotonicity Theorem, the correspondence of optimal constrained locations $l^C(\cdot)$ is decreasing in a. □

Note that, even though I assume CRRA utility on the main text, the proof used in this step holds for any utility function that satisfies the Inada conditions.[27] It's also easy to see that, under CRRA utility, when borrowing is not allowed ($\underline{b} = 0$), the direction of sorting is entirely determined by σ, and behaves as described in Corollary 1 in the main text. Whenever the credit constraint allows for strictly positive borrowing ($\underline{b} > 0$), we can also see that $l^C(\cdot)$ will be decreasing in a even when $\sigma = 1$ (log utility).

Step 3: The set of constrained agents is an upper set (no re-entry).

By Lemma 2, the set of constrained agents $\mathcal{A}^C$ forms an upper set. This means that if agent a is constrained, any other agent with $a' > a$ will also be constrained. □

[26] See e.g. Topkis (1998), Theorem 2.6.1.

[27] More precisely, I need that the utility function $u(\cdot)$ is strictly increasing, strictly concave, and satisfies $\lim_{c \to 0} \partial u(c)/\partial c = +\infty$. This last condition is necessary to ensure that optimal choices will never lead to negative consumption.

Given these three steps we are now ready to prove Proposition 2.

Proposition 2. For any increasing price schedule $p(\ell)$, the location choices of constrained agents are decreasing in learning ability (a) when the intertemporal elasticity of substitution (σ) is smaller than one.

Proof. Combining the three steps above, we can conclude that individual location decisions, $l^*(a,k)$, are characterized by a unique threshold $a^*(k)$ such that $l^*(a,k)$ follows the unconstrained path for $a \leq a^*$ and the constrained path for $a > a^*$. Hence,

$$l^*(a,k) = \begin{cases} l^U(a) & \text{if } a \leq a^* \quad \text{with } l^U(a) \text{ increasing in } a, \\ l^C(a,k) & \text{if } a > a^* \quad \text{with } l^C(a,k) \text{ decreasing in } a \text{ if IES} \leq 1. \end{cases}$$

This completes the proof of Proposition 2. □

Finally, the proof of Proposition 2 establishes how location choices depend on learning ability for any increasing price schedule. Since the equilibrium price schedule is increasing in ℓ, we can conclude that these are the sorting patterns that will arise in equilibrium. Therefore, the proof of Proposition 3 is immediate.

B Details Related to the Quantitative Spatial Model

B.1 Retiree Problem

This section describes the retirement problem, first introduced in page 33.

Agents start the last period of their lives ($j = 3$) as retirees. They receive a pension from the government, ω, and face a warm-glow bequest motive. Their location decisions are determined exogenously according to the transition matrix $\pi_3^{\ell,\ell'}$. They choose tenure status, o, consumption, c, and bequests, B. Their problem is defined as

$$V_3(x_3) = \sum_n \pi_3^{\ell_3,n} v_3(\tilde{x}_3, n),$$

with ℓ-choice-specific value function

$$
\begin{aligned}
v_3(\tilde{x},\ell) = \max_{c,b,o} \;& u(c) + \Phi(B) \\
\text{s.t. } c + (1-o)q_\ell + op_\ell + b &= (1-\tau)\omega + Rk + T \\
B &= Rb + op_\ell \\
o \in \{0,1\}, \quad b &\geq -op_\ell(1-d^k), \quad c, B > 0.
\end{aligned}
$$

The bequest function $\Phi(B)$ is defined as

$$\Phi(B) = \phi \frac{B^{1-1/\sigma}}{1-1/\sigma}.$$

B.2 Housing Markets

This section describes the optimization problems of both the construction and rental sectors, first introduced in page 34.

Construction sector. A competitive construction sector operates a Leontief technology, using final goods (y) and land (f_ℓ) owned by absentee landlords. Similar to Favilukis et al. (2017) and Kaplan et al. (2020), I assume landlords charge a competitive land rental rate that extracts all economic profits from development. The representative local developer solves the static problem

$$\max_y p_\ell I_\ell - y - r_\ell f_\ell \qquad \text{s.t.} \quad I_\ell = \tilde{\eta} \min\{y, f_\ell\}$$

With land in fixed supply at f_ℓ and competitive construction, zero profits pin down the land rent.

Every period, a constant fraction d^k of houses depreciates and is immediately rebuilt. Defining L_ℓ as the mass of households living in location ℓ, steady-state housing investment satisfies

$$d^k L_\ell = I_\ell = \tilde{\eta} f_\ell$$

Given the exogenous land supply, f_ℓ, this equation pins down the steady-state housing stock in each location, L_ℓ. House prices p_ℓ are determined in general equilibrium to satisfy the housing market-clearing condition.

Rental markets. A competitive rental sector owns housing units in each location and rents them out to households. Rental companies can frictionlessly buy and sell units on the housing market at price p_ℓ. When renting out housing, they incur a rental friction $\chi \geq 1$. For each unit rented at rate q_ℓ, the rental company receives q_ℓ/χ in net revenue. The parameter χ represents any wedge distorting price-to-rent ratios, such as search and matching frictions, management and regulatory costs, or tenant protection laws. Rental companies are subject to the same depreciation as households and need to renew a constant fraction d^k of their housing stock every period.

The problem of the representative local rental company is as follows

$$V(\mathcal{R}_\ell) \;=\; \max_{\mathcal{R}'_\ell} \; \frac{q_\ell}{\chi}\mathcal{R}'_\ell - p_\ell\left[\mathcal{R}'_\ell - (1-d^k)\mathcal{R}_\ell\right] + \frac{1}{R}V(\mathcal{R}'_\ell)$$

The first-order condition in steady state implies that the equilibrium rental rate is a constant markup over the user cost of housing:

$$q_\ell = \chi\left(1 - \frac{1-d^k}{R}\right)p_\ell.$$

C Calibration Details

D Data Sources

E Estimation Details

As described in the main text, the log earnings of individual i, residing in location partition $\ell \in \{M, V, O\}$ can be expressed as follows

$$w_{i,\ell,t} = \eta_{g(i)} + \tilde{z}_\ell + \alpha_{g(i)}E_{i,t-1} + \sum_{c\in\{M,V\}} \tilde{\psi}_c E_{i,c,t-1} + u_{i,\ell,t}, \tag{24}$$

where $g(i)$ defines the latent group of individual i and $E_{i,j} \equiv \sum_j \delta_j^a$ is a measure of age-weighted accumulated years of experience. To match the first period considered in the quantitative model, I run this regression using individuals aged 25 to 45.

There are two challenges in this specification: (i) how to select the number of groups, and (ii) how to recover the right estimate for the individual income profile $\{\delta_j^a\}_{j\geq 25}$. In this Appendix, I describe in detail how I address each of these challenges.

E.1 Iterative estimation of life cycle profiles

The estimation of equation (24) faces a fundamental simultaneity problem: recovering the location effects $(\tilde{z}_\ell, \tilde{\psi}_\ell)$ and individual heterogeneity parameters $(\eta_{g(i)}, \alpha_{g(i)})$ requires knowledge of the age-weighting profile $\{\delta_j^a\}_{j=25}^{44}$. This profile, however, is not directly observable. Moreover, estimating it from the data would rely on these same parameters. To be precise, notice that taking first differences in equation (24), wage growth of individuals that stay in the same location between period $t-1$ and t is

$$\triangle w_{i,\ell,j} = \delta_j^a \left(\alpha_{g(i)} + \tilde{\psi}_\ell \right) \tag{25}$$

I address this problem through an iterative procedure that alternates between estimating the regression parameters conditional on a given age profile, and updating the age profile conditional on the regression parameters.

Step 0: Initial guess for the age profile. I construct an initial age profile $\{\hat{\delta}_j^{a,0}\}_{j=25}^{44}$ by normalizing observed average wage growth across all individuals:

$$\hat{\delta}_j^{a,0} = \overline{\Delta w_j} \left(\frac{20}{\sum_{s=25}^{44} \overline{\Delta w_s}} \right)$$

where $\overline{\Delta w_j} \equiv \mathbb{E}_i[\Delta w_{i,j} \mid \ell_{i,j} = \ell_{i,j-1}]$ denotes the average wage growth at age j among non-movers.[28] I normalize the estimates to have mean one. I pick this normalization to facilitate the interpretation of subsequent estimates.[29] The resulting initial profile

[28]Note that equation (25) holds only for individuals staying in the same location partition. For this reason, I do not use the wage growth of migrants at the period in which they move. However, I do include them in all other periods to minimize the potential for selection bias.

[29]On average, across the 20 years included in the regression, one year of experience is worth one unit of age-weighted experience. The profile $\{\delta_j^a\}$ therefore determines how much more or less valuable experience is at age j relative to this average. At the estimation stage, this normalization implies that the parameter ψ_ℓ should be interpreted as the additional wage growth per year associated to location ℓ.

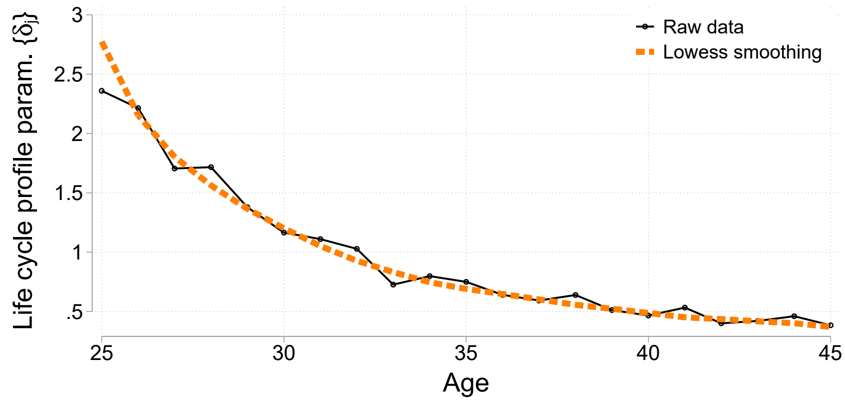


Figure 9: Initial guess for life cycle profile, $\{\delta_j^a\}_{j=25}^{45}$.

can be seen in Figure 9.

It is easy to see that this initial estimate is biased, since observed wage growth confounds three distinct components:

$$\overline{\Delta w_j} = \delta_j^a \left[\mathbb{E}[\alpha_{g(i)}] + \sum_{\ell} \tilde{\psi}_\ell \cdot \Pr(\ell_{i,j} = \ell \mid j) \right]$$

The bias arises if the spatial distribution of workers varies systematically with age, which we know is a very robust pattern in the data.[30] This means that the second term inside the brackets will be on average higher for young workers, causing $\hat{\delta}_j^{a,0}$ to overstate the true returns to experience early in the life cycle.

Step 1: Estimate regression conditional on age profile. Given the current estimate $\{\hat{\delta}_j^{a,n}\}$, I construct age-weighted experience measures:

$$E_{i,t-1}^n = \sum_{j=25}^{age(i,t)-1} \hat{\delta}_j^{a,n}, \qquad E_{i,\ell,t-1}^n = \sum_{j=25}^{age(i,t)-1} \hat{\delta}_j^{a,n} \cdot \mathbb{1}[\ell_{i,j} = \ell]$$

I then estimate equation (24) using the two-stage grouped fixed effects procedure described in Section 3.3, obtaining estimates $\hat{\theta}^n = (\hat{\eta}_{g(i)}^n, \hat{\alpha}_{g(i)}^n, \hat{\tilde{z}}_\ell^n, \hat{\tilde{\psi}}_\ell^n)$.

[30]This is also a pattern incorporated in the model: young workers have longer horizons over which to reap returns. As such, they will have higher incentives to choose locations offering high learning opportunities. Similarly, if older workers systematically return to lower-growth locations later in the life cycle, it would further bias the age profile.

Step 2: Update age profile conditional on regression parameters. Given $\hat{\theta}^n$, I recover an updated age profile by exploiting the relationship described in equation (25). I do this by regressing individual wage growth $\Delta w_{i,j}$ on age-specific interactions with the constructed variable $(\hat{\alpha}^n_{g(i)} + \hat{\tilde{\psi}}^n_{\ell_{i,j}})$, using only non-movers at age j.

To ensure stable convergence, I apply two smoothing procedures at this point. First, I apply lowess (locally weighted scatterplot) smoothing to the raw age-specific estimates to reduce sampling noise while preserving the profile's shape. I use a bandwidth of 0.25, which provides sufficient smoothing without over-constraining the functional form. An example of how this works can be seen in Figure 9. Second, I implement damped updating by setting $\hat{\delta}^{a,n+1}_j = 0.3 \cdot \hat{\delta}^{a,n}_j + 0.7 \cdot \hat{\delta}^{a,n+1}_{\text{smoothed}}$. This prevents overshooting and stabilizes convergence.

This returns a new estimated profile $\{\hat{\delta}^{a,n+1}_j\}^{44}_{j=25}$, which I once again normalize to ensure that average age-weighted experience is equal to one.

Convergence. I iterate between Steps 1 and 2 until $|\hat{\theta}^{n+1} - \hat{\theta}^n| < \epsilon$ for a pre-specified tolerance $\epsilon = 10^{-4}$. In practice, the algorithm converges within 5-10 iterations.

Importantly, the group assignment $g(i)$ is determined only once using the initial age profile $\{\hat{\delta}^{a,0}_j\}$ and remains fixed throughout all iterations. Re-computing types in each iteration creates a feedback loop wherein the clustering algorithm adapts to the current age profile, re-parameterizing age effects as ability heterogeneity. Empirically, this manifests as types becoming increasingly correlated with age as iterations proceed.[31] Fixing types ensures that $g(i)$ represents stable individual characteristics, allowing the procedure to separately identify the age profile δ^a_j from individual learning ability α_g.

Discussion. The iterative procedure resolves the initial simultaneity problem by exploiting different sources of variation at each step. In Step 1, conditional on $\{\hat{\delta}^{a,n}_j\}$, the location effects $(\tilde{z}_\ell, \tilde{\psi}_\ell)$ are identified from workers with identical experience profiles $(E^n_{i,t}, E^n_{i,\ell,t})$ but different location histories. Conversely, in Step 2, conditional on the location parameters, the age profile is identified from comparing wage growth of workers in the same location but at different ages. These two identification strategies are complementary: movers provide the variation needed to separate location effects from individual heterogeneity (Step 1), while within-location age variation purges the

[31] I check that, conditional on the initial profile, grouped fixed effects are uncorrelated with age.

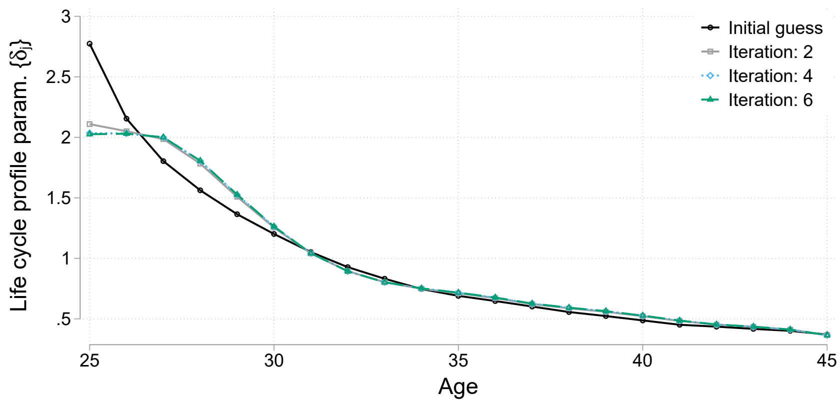


Figure 10: Iterations of estimated life cycle profile, $\{\delta_j^a\}_{j=25}^{45}$.

location-composition bias from the age profile (Step 2).

Figure 10 plots three iterations of the estimated life cycle profiles associated with the exercise in the main text.