Online Appendix
Land, Wealth, and Taxation

Roberto Brunetti* Carl Gaigné† Fabien Moizeau‡

Abstract

This appendix presents the theoretical derivations (Appendix A) and provides six extensions (Appendix B): (i) heterogeneity in labor force and wages are endogenous; (ii) heterogeneity in housing size; (iii) a construction industry; (iv) payment-to-income constraint instead of loan-to-value constraint; (v) the interest rate for borrowers is larger than the interest rate for lenders; (vi) housing bubbles.

Appendix A. Theorical derivations

A.1 Spatial wealth sorting

No mortgage market imperfections. Without borrowing constraints, the slope of the bid-rent curve, given by

$$\frac{\partial \Psi_t(x, g_t)}{\partial x} = -\frac{\partial \tilde{\kappa}(x)}{\partial x} + \mu^2 \frac{\partial p_{t+2}(x)}{\partial x}. \tag{A.1}$$

does not vary across agents which implies there is no spatial wealth sorting within the city. The price of land located at distance x is

$$\max\left\{\mathcal{K}_t - \tilde{\kappa}_t(x) + \mu^2 p_{t+2}(x), R_A\right\} \tag{A.2}$$

while the net lifetime wealth of young agent is

$$W(g_t) = w + g_t + \mu y_{t+1} - \mathcal{K}_t - \frac{G}{L_{0,t} + L_{1,t}} - \mu \frac{G}{L_{0,t} + L_{0,t+1}}. \tag{A.3}$$

Since $\mathcal{K}_t - \tilde{\kappa}_t(L_{0,t}) + \mu^2 R_A = R_A$ at the city limit, we have

$$\mathcal{K}_t = (1 - \mu^2) R_A + \tilde{\kappa}_t(L_{0,t}) \tag{A.4}$$

which is common to all types of agents, but depends on city size.

*LEMMA Université Paris Panthéon-Assas. Email: roberto.brunetti@assas-universite.fr
†INRAE, SMART, Rennes, France and Laval University, CREATE, Quebec, Canada. Email: carl.gaigne@inrae.fr
‡Univ Rennes, CNRS, CREM-UMR6211, F-35000 Rennes, France. Email: fabien.moizeau@univ-rennes1.fr

Proof of Proposition 1. From (18), we have for any g_t agent that

$$\psi'(x, g_t) = \begin{cases} 0 & \text{for } x \in [0, \widehat{x}_t(g_t)], \\ \Psi_t'(x, g_t) < 0 & \text{for } x \in (\widehat{x}_t(g_t), L_{0,t}] \end{cases}$$

considering that ψ is not differentiable at $\widehat{x}_t(g_t)$ (a prime denotes $\mathrm{d}/\mathrm{d}x$).

First, suppose that $\widehat{x}_t(g_t) \in [0, L^\star_{0,t}]$ for some g_t. Since $\widehat{x}_t(g_t)$ decreases with g_t, we can define the threshold $\widehat{g}_t \in [\underline{g}_t, \overline{g}_t]$ such that, at a given location x, we have $x = \widehat{x}(\widehat{g}_t)$. Considering a bid-rent function $\psi(x, g_t)$, we take a point on that curve $(\widehat{x}_t(\widehat{g}_t), \psi(\widehat{x}_t(\widehat{g}_t), g_t))$. By holding the value of ψ constant, we can then rank the bid-rent slopes according to the wealth level for any $x = \widehat{x}_t(\widehat{g}_t) \in (0, L^\star_{0,t}]$:

$$\psi'(x, g_t)\big|_{x=\widehat{x}(\widehat{g}_t), \psi=\text{const}} = \begin{cases} \Psi_t'(x, g_t) & \text{for any } g_t > \widehat{g}_t, \\ 0 & \text{for } g_t < \widehat{g}_t. \end{cases}$$

where, from (A.1), $\Psi_t'(x, g_t) = -\tilde{\kappa}_t'(x) + \mu^2 p_{t+2}'(x) < 0$. Steeper bid rents imply locations closer to the CBD, as agents with steeper bid rents bid away the agents with flatter bid rents. In other words, agents endowed with an initial wealth $g_t > \widehat{g}_t$ can outbid agents poorer than $\widehat{g}_t$ in more attractive areas, *i.e.*, in the area $[0, \widehat{x}(\widehat{g}_t)[$.

If $L^\star_{0,t} > \widehat{x}(\overline{g}_t) > 0$, then for any $x \in [0, \widehat{x}(\overline{g}_t)]$, bid-rent slopes are nil for any agent. However, as the bid rent $g_t/(1-\lambda)$ is strictly increasing with the agent's wealth and any constrained borrower strictly prefers to live close to the city center, there is perfect sorting in the area $[0, \widehat{x}(\overline{g}_t)]$.

If $0 < \widehat{x}(\underline{g}_t) < L^\star_{0,t}$, then for any $x \in [\widehat{x}(\underline{g}_t), L^\star_{0,t}]$, this implies that there is no $\widehat{g}_t \in [\underline{g}_t, \overline{g}_t]$ such that $x = \widehat{x}(\widehat{g}_t)$ for $x \in [\widehat{x}(\underline{g}_t), L^\star_{0,t}]$. No agent g_t is constrained by the borrowing constraint at any $x \in [\widehat{x}(\underline{g}_t), L^\star_{0,t}]$. There is no sorting in this area, as the bid-rent slope is the same whatever the agent g_t for locations $x \in [\widehat{x}(\underline{g}_t), L^\star_{0,t}]$.

Second, assume that there is no $\widehat{g}_t(g_t) \in [0, L^\star_{0,t}]$ for any g_t. This arises under two cases: (i) The wealth distribution is such that $(1-\lambda)\Psi_t(x, g_t) > \overline{g}_t$ for any $x \in [0, L_{0,t}]$, meaning that there does not exist $\widehat{x}_t(g_t) \in [0, L^\star_{0,t}]$ for any g_t. This implies that all agents are borrowing constrained, all agents bid $g_t/(1-\lambda)$ for land, and there is perfect sorting in the area $[0, L^\star_{0,t}]$, and (ii) the wealth distribution is such that $(1-\lambda)\Psi_t(0, g_t) < \underline{g}_t$, meaning that there does not exist $\widehat{x}_t(g_t) \in [0, L^\star_{0,t}]$ for any g_t. This implies that for no agent and no location the borrowing constraint is binding. One could not rank bid-rent slopes because all agents would be able to pay Ψ_t.

As bid-rent slopes can be ranked according to the wealth level for any $x = \widehat{x}_t(\widehat{g}_t) \in (0, L^\star_{0,t}]$ this implies the existence of a unique negative assortative matching between wealth levels and distance x. In other words, there is a unique one-to-one and decreasing relationship between g_t and x (Chiappori, 2017).

A.2. The internal structure of cities

The size, number and location of the areas hosting borrowing-constrained agents depend on the the shape of the wealth distribution. We illustrate the model by developing an example in which commuting costs are linear and increasing with distance and wealth g_t is Pareto distributed, truncated to the support $[\underline{g}_t, \overline{g}_t]$ with shape parameters $\omega > 0$: $F_t(g) = [1 - (g/\underline{g}_t)^{-\omega}]/\phi_t$ with $\phi_t \equiv 1 - (\overline{g}_t/\underline{g}_t)^{-\omega} \in (0, 1]$.

A low value of ω means that the wealth distribution is close to uniform among agents, whereas that distribution gets more and more skewed towards low-wealth agents for larger values of ω. In addition, ϕ_t increases with $\overline{g}_t/\underline{g}_t$. In addition, when perfect sorting occurs, we can deduce that $g_t^\star(x)$ is decreasing and *convex* since the cumulative distribution function of a Pareto variable is an increasing and concave function. Figure 1 illustrates this case. From Proposition 1, assortative matching requires that the x% of the wealthiest agents be matched with the x% of the least distant locations.

Formally, the matching function $g_t(x)$ can be retrieved from the following condition:

$$L_{0,t}\int_{g_t}^{\overline{g}_t} f_t(g)\mathrm{d}g = \int_0^x \mathrm{d}x \tag{A.5}$$

The proportion of the population with wealth greater than or equal to g_t being $1 - F_t(g)$, a simple calculation then shows that the equilibrium wealth mapping is such that

$$g_t^\star(x) = \underline{g}_t\left[\frac{1}{1-\phi_t+(\phi_t/L_{0,t})x}\right]^{1/\omega}. \tag{A.6}$$

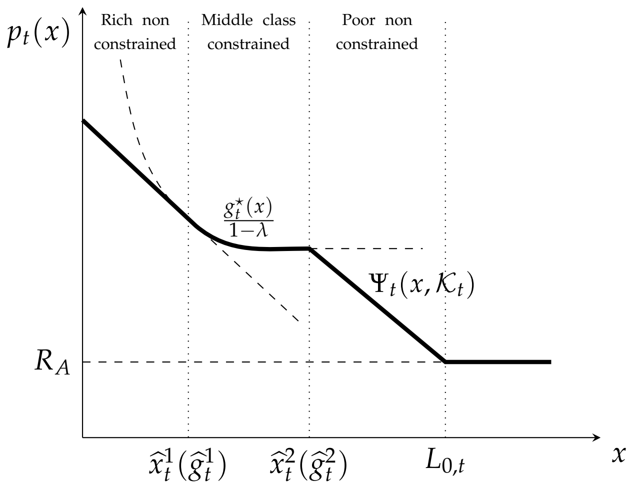

Figure 1: Equilibrium with Pareto distribution

Consider the case where the city structure is characterized by three areas, each one hosting agents with a particular status, constrained or non-constrained. The poorest individuals who are assumed to not face any borrowing constraint live at the least attractive places, *i.e.* the city fringe $[\widehat{x}_t^2, L_{0,t}^\star]$ and pay $\Psi_t(x, g_t)$ where the threshold location $\widehat{x}_t^2$ is such that $g_t^\star(\widehat{x}_t^2) = (1-\lambda)\Psi_t(\widehat{x}_t^2, g_t)$. The richest agents can afford to live in the most attractive locations $[0, \widehat{x}_t^1]$ without being borrowing contrained while the intermediate-wealth agents, the only agents who are borrowing constrained, reside in the area $[\widehat{x}_t^1, \widehat{x}_t^2]$ and pay $g_t^\star(x)/(1-\lambda)$. The house price paid by the wealthiest agents (with $g_t \geq \widehat{g}_t^1$ where $\widehat{g}_t^1$ is characterized below) is given by the bid rent $\Psi_t^1(x, g_t) = K_t^1 - \Upsilon_t(x)$ with

$$\Upsilon_t(x) \equiv \tilde{\kappa}(x) - \mu^2 p_{t+2}(x)$$

where K_t^1 is obtained from $K_t^1 - \Upsilon_t(\widehat{x}_t^1) = g_t^\star(\widehat{x}_t^1)/(1-\lambda)$ and the location cutoff $\widehat{x}_t^1$ is such that the marginal agent endowed with wealth $\widehat{g}_t^1 \equiv g_t^\star(\widehat{x}_t^1)$ cannot outbid credit-constrained agents, *i.e.*

$$V[\Psi_t^1(x, K_t^1), \widehat{g}_t^1] \geq V[g_t^\star(x)/(1-\lambda), \widehat{g}_t^1]$$

for any $x \in [\widehat{x}_t^1, \widehat{x}_t^2]$ where $V(.)$ is the indirect utility. This condition holds if $g_t^{\star\prime}(x) = -(1-\lambda)\Upsilon_t'(x)$ evaluated at $\widehat{x}_t^1$ and $g_t^\star(x)$ is strictly convex. Stated differently, *the gain in housing expenditures by marginally moving further from the CBD and paying a lower rent $g_t^\star(x)/(1-\lambda)$ must be lower than additional costs associated with distance in equilibrium*. Given Pareto wealth distribution, $g_t^{\star\prime}(x) = -(1-\lambda)\Upsilon_t'(x)$ becomes

$$\frac{\phi_t(\widehat{g}_t^1)^{1+\omega}}{\omega L_{0,t}\underline{g}_t^{\omega}} = (1-\lambda)\Upsilon_t'(\widehat{x}_t^1).$$

Since $\Upsilon_t(x)$ increases linearly with distance while $g_t^\star(x)$ is decreasing and convex, there is a single solution (see Figure 1). As a consequence, $\Psi_t^1(x, g_t)$ is the tangent to the curve $g_t^\star(x)/(1-\lambda)$ at $\widehat{x}_t^1$. Moreover, it is below the curve $g_t^\star(x)/(1-\lambda)$ and parallel to $\Psi_t(x, \mathcal{K}_t)$ (see Figure 1). Using the equilibrium condition for which the utility level of the individual living in $\widehat{x}_t^1$ is such that $V(\Psi_t^1, \widehat{g}_t^1) = V(\widehat{g}_t^1/(1-\lambda), \widehat{g}_t^1)$, we have $\Psi_t^1(\widehat{x}_t^1, g_t) = \widehat{g}_t^1/(1-\lambda)$ and $\Psi_t^1(x, g_t) < \Psi_t(x, \mathcal{K}_t)$ so that $K_t^1 < \mathcal{K}_t$ (see Figure 1). Further, for any $x \in [0, \widehat{x}_t^1]$, we have $p_t^\star(x) = \Psi_t^1(x, g_t) < g_t/(1-\lambda)$ for all g_t agents with $g_t \in [0, \widehat{g}_t^1]$, implying that they are not borrowing-constrained. *Even though the wealthiest agents are not credit-constrained, they gain from credit market imperfections* as their housing expenditures is reduced by the amount $\mathcal{K}_t - K_t^1$. Note that if $\widehat{x}_t^1$ does not exist, the city structure would be characterized by two areas: the wealthiest agents with $g_t \geq \widehat{g}_t^2$ reside at $x \in [0, \widehat{x}_t^2]$ and are borrowing constrained and the poorest agents live in $x \in [\widehat{x}_t^2, L_{0,t}^\star]$ and are not constrained.

General case. We now characterize of the spatial pattern of 'constrained' and 'non-constrained' areas for any inherited wealth distribution. *If there exists a cutoff location $\widehat{x}_t^0 \in [0, L_{0,t}^\star]$ such that $g_t^\star(\widehat{x}_t^0)/(1-\lambda) = \Psi_t(\widehat{x}_t^0, \mathcal{K}_t)$, amounting to have*

$$(1-\lambda)\Psi_t(0, \mathcal{K}_t) > \underline{g}_t > (1-\lambda)\Psi_t\left(L_{0,t}^\star, \mathcal{K}_t\right), \tag{A.7}$$

then the city is partitioned into distinct areas defined by cutoff locations $\widehat{x}_t^j$, $j = 1, ..., J$ with $0 < \widehat{x}_t^J < ... < \widehat{x}_t^j < ... < \widehat{x}_t^0 < L_{0,t}^\star$ such that

(i) Residents living in the area $[\widehat{x}_t^0, L_{0,t}^\star]$ are not borrowing constrained and pay the rent $\Psi_t(x, \mathcal{K}_t)$.

(ii) If there exists $\widehat{x}_t^1 \in [0, \widehat{x}_t^0]$ such that

$$-g_t^{\star\prime}(\widehat{x}_t^1) = (1-\lambda)\Upsilon_t'(\widehat{x}_t^1) \quad \textit{and} \quad g_t^\star(x) \textit{ is convex at } \widehat{x}_t^1, \tag{A.8}$$

then residents in area $x \in [\widehat{x}_t^1, \widehat{x}_t^0]$ are borrowing constrained and pay the rent $g_t^\star(x)/(1-\lambda)$. Otherwise, all residents are borrowing constrained in the area $x \in [0, \widehat{x}_t^0]$.

(iii) Provided that (A.8) *holds. If there exists $\widehat{x}_t^2 > 0$ such that*

$$\frac{g_t^\star(\widehat{x}_t^2)}{1-\lambda} = \Psi_t^1(\widehat{x}_t^2, g_t) \quad \textit{with} \quad K_t^1 = \Upsilon_t(\widehat{x}_t^1) + \frac{g_t^\star(\widehat{x}_t^1)}{1-\lambda} \tag{A.9}$$

then residents in the area $[\widehat{x}_t^2, \widehat{x}_t^1]$ are not borrowing constrained and pay the rent $\Psi_t^1(x, g_t) = K_t^1 - \Upsilon_t(x)$. Otherwise, residents living in the area $[0, \widehat{x}_t^1]$ are not borrowing constrained.

(iv) Provided cutoff locations with counting number $z = 0, ..., j$ exist, if $\widehat{x}_t^j$ is given by (A.8), *respectively* (A.9),*then the area $[\widehat{x}_t^j, \widehat{x}_t^{j+1}]$ hosts borrowing constrained, respectively unconstrained residents.*

The city hosts at most two types of agents in equilibrium, those who are credit-constrained and those who are not, under condition (A.7). The linearity of Ψ_t is obtained if $\kappa(x)$ is linear and the rent is stationary. This is not necessary for our results and this is made for expositional purpose. Condition (A.7) is easily satisfied as one can always find an agricultural rent so that the housing price at the city fringe is very low and even the poorest resident is not borrowing constrained. By contrast, it is relevant to consider that the willingness to pay at the CBD is high enough so that it prevents some poor agents from borrowing without facing any limit.

Let us now consider particular wealth distributions.

Case 1. Concave wealth distributions. Let us consider the case where g_t is distributed according to the truncated Pareto distributed as above. The mapping $g_t^\star(x) = F_t^{-1}\left(1 - \frac{x}{L_{0,t}^\star}\right)$ is given by (A.6). If the distribution is such that

$$- \left.\frac{\mathrm{d}g_t^\star(x)}{\mathrm{d}x}\right|_{x=L_{0,t}^\star} = -\frac{1}{\omega}\frac{\phi_t}{L_{0,t}^\star}\underline{g}_t < \Upsilon_t'(L_{0,t}^\star) \tag{A.10}$$

and that condition (A.7) is satisfied then the urban equilibrium is characterized as follows:

(i) Agents with wealth $\underline{g}_t \leq g_t(x) \leq \widehat{g}_t^2$ reside in the area $\left[\widehat{x}_t^2(\widehat{g}_t^2), L_{0,t}^\star\right]$ and are not borrowing-constrained.

(ii) Agents with wealth $\widehat{g}_t^2 \leq g_t(x) \leq \widehat{g}_t^1$ reside in the area $\left[\widehat{x}_t^1(\widehat{g}_t^1), \widehat{x}_t^2(\widehat{g}_t^2)\right]$ and borrow up to their borrowing limit.

(iii) Agents with wealth $\widehat{g}_t^1 \leq g_t(x) \leq \overline{g}_t$ reside in the area $\left[0, \widehat{x}_t^1(\widehat{g}_t^1)\right]$ and are not borrowing-constrained.

(iv) The equilibrium rent is such that

$$p_t^\star(x) = \begin{cases} \Psi_t^1\left(x, g_t^\star(x)\right) = K_t^1 - \Upsilon_t(x) & \text{for } x \in \left[0, \widehat{x}_t^1\left(\widehat{g}_t^1\right)\right] \\ \frac{g_t^\star(x)}{1-\lambda} & \text{for } x \in \left[\widehat{x}_t^1\left(\widehat{g}_t^1\right), \widehat{x}_t^2(\widehat{g}_t^2, \mathcal{K}_t)\right] \\ \Psi_t\left(x, \mathcal{K}_t\right) = \mathcal{K}_t - \Upsilon_t(x) & \text{for } x \in \left[\widehat{x}_t^2(\widehat{g}_t^2, \mathcal{K}_t), L_{0,t}^\star\right]. \end{cases}$$

with $\widehat{x}_t^1\left(\widehat{g}_t^1\right)$ such that

$$\left.\frac{\mathrm{d}g_t^\star(x)}{\mathrm{d}x}\right|_{x=\widehat{x}_t^1} = (1-\lambda)\Upsilon'(\widehat{x}_t^1),$$

K_t^1 such that

$$K_t^1 - \Upsilon_t(\widehat{x}_t^1) = \widehat{g}_t^1(\widehat{x}_t^1)/(1-\lambda),$$

and $\widehat{x}_t^2(\widehat{g}_t^2, K)$ such that

$$\mathcal{K}_t - \Upsilon_t(\widehat{x}_t^2) = \widehat{g}_t^2(\widehat{x}_t^2)/(1-\lambda).$$

Let us characterize the threshold distances and the rent function. First, assuming (A.7) holds implies that the poorest agent does not need to borrow at the city fringe. The bid rent at $L_{0,t}^\star$ is equal to $\Psi_t\left(L_{0,t}^\star, \mathcal{K}_t\right) = R_A$. Hence, $\mathcal{K}_t - \Upsilon_t(L_{0,t}^\star) = R_A$ leading to $\mathcal{K}_t = \kappa\left(L_{0,t}^\star\right) + R_A\left(1-\mu^2\right)$. Hence, the rent function for non-constrained agents living at $\left[\widehat{x}_t^2(\widehat{g}_t^2, \mathcal{K}_t), L_{0,t}^\star\right]$ is $\Psi_t\left(x, \mathcal{K}_t\right) = \mathcal{K}_t - \Upsilon_t\left(x\right)$.

Since the mapping $g_t^\star(x)$ given by (A.6) is decreasing and convex, and that $\Psi_t(x, \mathcal{K}_t)$ is a decreasing function, we consider that the wealth distribution is such that there exists $\widehat{x}_t^2$ satisfying

$$\Psi_t\left(\widehat{x}_t^2, \mathcal{K}_t\right) = \frac{\widehat{g}_t^2}{1-\lambda} \text{ and } \widehat{g}_t^2 = F_t^{-1}\left(1 - \frac{\widehat{x}_t^2}{L_{0,t}^\star}\right),$$

and given (A.7), we have

$$\frac{g_t^\star(x)}{1-\lambda} \geq \Psi_t(x, \mathcal{K}_t), \text{ for any } x \in \left[\widehat{x}_t^2, L_{0,t}^\star\right],$$

implying that the agents are not borrowing-constrained.

Second, let us consider the area $\left[\widehat{x}_t^1(\widehat{g}_t^1), \widehat{x}_t^2(\widehat{g}_t^2, \mathcal{K}_t)\right]$. We define $\widehat{x}_t^1(\widehat{g}_t^1)$ such that it satisfies

$$\frac{\mathrm{d}g_t^\star(\widehat{x}_t^1)}{\mathrm{d}x} = (1-\lambda)\mathrm{Y}_t'(\widehat{x}_t^1).$$

It is unique given $g_t^\star(x)$ is decreasing and convex and that $(1-\lambda)\mathrm{Y}'$ is assumed to be constant ($\Psi_t(x, \mathcal{K}_t)$ is assumed to be linear and decreasing). Hence, for any $x \in \left[\widehat{x}_t^1(\widehat{g}_t^1), \widehat{x}_t^2(\widehat{g}_t^2, \mathcal{K}_t)\right]$

$$\frac{g_t^\star(x)}{1-\lambda} \leq \Psi_t(x, \mathcal{K}_t),$$

and given (A.10), we have for $x \in \left[\widehat{x}_t^1(\widehat{g}_t^1), \widehat{x}_t^2(\widehat{g}_t^2, \mathcal{K}_t)\right]$

$$-\frac{\mathrm{d}g_t^\star(x)/(1-\lambda)}{\mathrm{d}x} \leq -\frac{\mathrm{d}\Psi_t(x, \mathcal{K}_t)}{\mathrm{d}x} = \frac{\mathrm{d}\mathrm{Y}_t(x)}{\mathrm{d}x}.$$

It turns out that all residents are borrowing-constrained and they are sorted across space. At $\widehat{x}_t^1$, the agent $\widehat{g}_t^1 = g_t^\star(\widehat{x}_t^1)$ has no interest to deviate to another location $x \in \left[\widehat{x}_t^1(\widehat{g}_t^1), \widehat{x}_t^2(\widehat{g}_t^2, \mathcal{K}_t)\right]$ paying $g_t^\star(x)/(1-\lambda)$, that is

$$V\left[W_t(\Psi_t^1(\widehat{x}_t^1, g_t), \widehat{g}_t^1)\right] \geq V\left[W_t(g_t^\star(x)/(1-\lambda), \widehat{g}_t^1)\right],$$

since

$$-\frac{\mathrm{d}g_t^\star(x)/(1-\lambda)}{\mathrm{d}x} \leq \frac{\mathrm{d}\mathrm{Y}(x)}{\mathrm{d}x}$$

any deviation leads to a decrease of W_t.

The bid rent $\Psi_t(x, K_t^1)$ at $\widehat{x}_t^1$ is such that the borrowing constraint is binding. This allows us to get K_t^1

$$K_t^1 - \mathrm{Y}_t(\widehat{x}_t^1) = \widehat{g}_t^1(\widehat{x}_t^1)/(1-\lambda).$$

Note that as $g_t^\star(x)/(1-\lambda) \leq \Psi_t(x, \mathcal{K}_t)$ for any $x \in \left[\widehat{x}_t^1(\widehat{g}_t^1), \widehat{x}_t^2(\widehat{g}_t^2, \mathcal{K})\right]$, we deduce that $K_t^1 < \mathcal{K}_t$.

Third, residents in $x \in \left[0, \widehat{x}_t^1(\widehat{y}_t^1)\right]$ are sorted according to wealth along space. The rent is $\Psi_t^1(x, g_t) = K_t^1 - \mathrm{Y}_t(x)$. As $g_t^\star(x)$ is convex and Ψ linear we deduce that $g_t^\star(x)/(1-\lambda) > \Psi_t^1(x, g_t)$ implying that inhabitants are not borrowing-constrained.

Remark that the Pareto distribution could be such that there does not exist any $\widehat{x}_t^1$. If this was the case then the urban configuration would be characterized by two types of urban areas:

(i) Agents with wealth $\underline{g}_t \leq g_t(x) \leq \widehat{g}_t^2$ reside in the area $\left[\widehat{x}_t^2(\widehat{g}_t^2), L_{0,t}^\star\right]$ and are not borrowing-constrained.

(ii) Agents with wealth $\widehat{g}_t^2 \leq g_t(x) \leq \overline{g}_t$ reside in the area $\left[0, \widehat{x}_t^2(\widehat{g}_t^2)\right]$ and borrow up to their borrowing limit.

This proof could be extended to any other concave c.d.f. leading to a decreasing and convex mapping $g_t^\star(x)$.

Case 2. Convex wealth distributions. Let us consider that g_t is distributed according to a convex c.d.f. F_t on the support $[\underline{g}_t, \overline{g}_t]$ leading to a concave mapping $g_t^\star(x)$. If condition (A.7), then the urban equilibrium is characterized as follows

(i) Agents with wealth $\underline{g}_t \leq g_t(x) \leq g_t^\star(\widehat{x})$ reside in the area $\left[\widehat{x}_t, L_{0,t}^\star\right]$ and are not borrowing-constrained.

(ii) Agents with wealth $g_t^\star(\widehat{x}) \leq g_t(x) \leq \overline{g}_t$ reside in the area $[0, \widehat{x}_t]$ and are borrowing-constrained.

(iv) The equilibrium rent is such that

$$p_t^\star(x) = \begin{cases} \frac{g_t^\star(x)}{1-\lambda} & \text{for } x \in [0, \widehat{x}_t] \\ \Psi_t(x, \mathcal{K}_t) = \mathcal{K}_t - \Upsilon_t(x) & \text{for } x \in \left[\widehat{x}_t, L_{0,t}^\star\right]. \end{cases}$$

with

$$g_t^\star(x) = F_t^{-1}\left(1 - \frac{x}{L_{0,t}^\star}\right)$$

$\widehat{x}_t$ such that

$$g_t^\star(\widehat{x}) = F_t^{-1}\left(1 - \frac{\widehat{x}}{L_{0,t}^\star}\right) = \Psi_t(\widehat{x}, \mathcal{K}_t).$$

A convex c.d.f. F_t leads to a concave wealth mapping $g_t^\star(x) = F_t^{-1}\left(1 - x/L_{0,t}^\star\right)$. Given (A.7), there exists $\widehat{x}_t$ such that

$$\begin{gathered} g_t^\star(\widehat{x}_t) = F_t^{-1}\left(1 - \widehat{x}_t/L_{0,t}^\star\right) = (1-\lambda)\Psi_t(\widehat{x}_t, \mathcal{K}_t), \\ g_t^\star(x) \leq (1-\lambda)\Psi_t(x, \mathcal{K}_t) \text{ for any } x \in [0, \widehat{x}_t], \\ \text{and } g_t^\star(x) \geq (1-\lambda)\Psi_t(x, \mathcal{K}_t) \text{ for any } x \in \left[\widehat{x}_t, L_{0,t}^\star\right]. \end{gathered}$$

We thus deduce that for all agents with wealth $\underline{g}_t \leq g_t^\star(x) \leq g_t^\star(\widehat{x}_t)$ residing in the area $\left[\widehat{x}_t, L_{0,t}^\star\right]$ they are not borrowing-constrained. By definition of $\Psi_t(x, \mathcal{K}_t)$, no agent in this area has an interest to move to another location in $\left[\widehat{x}_t, L_{0,t}^\star\right]$. In addition, no one living in $\left[\widehat{x}_t, L_{0,t}^\star\right]$ would be able to outbid residents of $x \in [0, \widehat{x}_t]$ who have wealth $g_t^\star(\widehat{x}_t) \leq g_t^\star(x) \leq \overline{g}_t$ and are borrowing-constrained. Take an agent $g = g_t^\star(x)$ residing at $x \in [0, \widehat{x}_t]$. First, given that she is borrowing-constrained, she is not able to outbid richer agents at any location $z < x$. Second, she has no interest to deviate to another location $z > x, z \in [0, \widehat{x}_t]$ paying a lower bid rent $g_t(z)/(1-\lambda)$, that is

$$V\left[W_t(g/(1-\lambda), g)\right] \geq V\left[W_t(g_t^\star(z)/(1-\lambda), g)\right],$$

because since the mapping $g_t^\star(x)$ is concave we have

$$-\frac{\mathrm{d}g_t^\star(x)/(1-\lambda)}{\mathrm{d}x} - \Upsilon_t'(x) < 0$$

implying that the benefit of paying a lower rent $g_t^\star(z)/(1-\lambda)$ is outweighed by the extra transportation cost. Third, since any borrowing-constrained agent enjoys greater utility rather than paying $\Psi_t(x, \mathcal{K}_t)$, the agent g has no interest to live at a location $x' \in [\widehat{x}_t, L_{0,t}]$.

Remark that if (A.7) does not hold, it turns out that there does not exist any $\widehat{x}$. The urban equilibrium is such that no resident is borrowing-constrained and the rent is $p_t^\star(x) = \Psi_t(x, \mathcal{K}_t)$ for any $x \in [0, L_{0,t}]$.

Case 3. Distributions with concave and convex portions. A log-normal distribution or a Frechet distribution fall into this case. We still assume condition (A.7) holds. We adapt the above proofs of cases 1 and 2. We focus on areas where the mapping is such that $g_t^\star(x)/(1-\lambda) \leq \Psi_t(x, \mathcal{K}_t)$.

First, assume that the mapping is such that there exists a unique tangency point $\widetilde{x}$ such that

$$-\frac{\mathrm{d}g_t^\star(\widetilde{x})/(1-\lambda)}{\mathrm{d}x} = \Upsilon_t'(\widetilde{x}) \text{ and } g_t^\star(x)/(1-\lambda) \text{ is convex at } x = \widetilde{x}. \tag{A.11}$$

The proof is similar to case 1 and leads to the following urban configuration:

(i) Agents with wealth $\underline{g}_t \leq g_t(x) \leq \widehat{g}_t$ reside in the area $\left[\widehat{x}_t(\widehat{g}_t), L_{0,t}^\star\right]$ and are not borrowing-constrained.

(ii) Agents with wealth $\widehat{g}_t \leq g_t(x) \leq g_t^\star(\widetilde{x}_t)$ reside in the area $[\widetilde{x}_t, \widehat{x}_t(\widehat{g}_t)]$ and borrow up to their borrowing limit.

(iii) Agents with wealth $g_t^\star(\widetilde{x}_t) \leq g_t(x) \leq \overline{g}_t$ reside in the area $[0, \widetilde{x}_t]$ and are not borrowing-constrained.

(iv) The equilibrium rent is such that

$$p_t^\star(x) = \begin{cases} \Psi_t^1(x, g_t^\star(x)) = \widetilde{K}_t^1 - \Upsilon_t(x) & \text{for } x \in [0, \widetilde{x}_t] \\ \frac{g_t^\star(x)}{1-\lambda} & \text{for } x \in [\widetilde{x}_t, \widehat{x}_t(\widehat{g}_t, \mathcal{K}_t)] \\ \Psi_t(x, \mathcal{K}_t) = \mathcal{K}_t - \Upsilon_t(x) & \text{for } x \in \left[\widehat{x}_t(\widehat{g}_t, \mathcal{K}_t), L_{0,t}^\star\right]. \end{cases}$$

with $\widetilde{x}_t$ satisfying (A.11) and $\widetilde{K}_t^1$ such that

$$\widetilde{K}_t^1 - \Upsilon_t(\widetilde{x}_t) = g_t^\star(\widetilde{x}_t)/(1-\lambda),$$

and $\widehat{x}_t(\widehat{g}_t, \mathcal{K}_t)$ such that

$$\mathcal{K}_t - \Upsilon_t(\widehat{x}_t) = \widehat{g}_t(\widehat{x}_t)/(1-\lambda).$$

Second, assume w.l.o.g. that the mapping is such that there are two points $\widetilde{x}$ and $\widetilde{z}$, $\widetilde{x} < \widetilde{z}$, such that

$$-\frac{\mathrm{d}g_t^\star(.)/(1-\lambda)}{\mathrm{d}x} = \Upsilon_t'(.) \text{ and } g_t^\star(.)/(1-\lambda) \text{ convex, for } x = \widetilde{x}, \widetilde{z}. \tag{A.12}$$

We define the constant $\widetilde{K}_t^z - \Upsilon_t(\widetilde{z}) = g_t^\star(\widetilde{z})/(1-\lambda)$. Two cases must be considered.

If $g_t^\star(x)/(1-\lambda) > \Psi_t^z(x, g_t) = \widetilde{K}_t^z - \Upsilon_t(x)$ for any $x \in [0, \widetilde{z}]$ than all agents with wealth $g_t \geq g_t^\star(\widetilde{z})$ are not borrowing-constrained when paying the rent $\Psi_t^z(x, g_t)$. Following the same reasoning as in proof of case 1, we have the same type of urban configuration as in case 1:

(i) Agents with wealth $\underline{g}_t \leq g_t(x) \leq \widehat{g}_t$ reside in the area $\left[\widehat{x}_t(\widehat{g}_t), L_{0,t}^\star\right]$ and are not borrowing-constrained. The location $\widehat{x}_t$ is such that $\Psi_t(\widehat{x}_t, \mathcal{K}) = g_t^\star(\widehat{x}_t)/(1-\lambda)$.

(ii) Agents with wealth $\widehat{g}_t \leq g_t(x) \leq g_t^{\star}(\widetilde{z})$ reside in the area $[\widetilde{z}, \widehat{x}_t(\widehat{g}_t)]$ and borrow up to their borrowing limit.

(iii) Agents with wealth $g_t^{\star}(\widetilde{z}) \leq g_t(x) \leq \overline{g}_t$ reside in the area $[0, \widetilde{z}]$, are not borrowing-constrained and pay the rent $\Psi_t^z(x, g_t)$.

If there exist locations $x \neq \tilde{z}$ such that $g_t^{\star}(x)/(1-\lambda) = \Psi_t^z(x, g_t)$, take the highest solution $\overline{x}$. We know that $\Psi_t^z(x, g_t)$ intersects from above $g_t^{\star}(x)/(1-\lambda)$ at $\overline{x}$ and that individuals with wealth $g_t^{\star}(x) \in [g_t^{\star}(\overline{x}), g_t^{\star}(\widetilde{x}')]$ are not borrowing-constrained, that is $g_t^{\star}(x)/(1-\lambda) > \Psi_t^z(x, g_t)$. At $\widetilde{z}$, we can then apply the same reasoning as in case 1 and deduce that all residents in area $[0, \widetilde{z}]$ pay the rent $\Psi_t^z(x, g_t)$ where $\widetilde{K}_t^z$ such that $\widetilde{K}_t^z - \Upsilon_t(\widetilde{z}) = g_t^{\star}(\widetilde{z})/(1-\lambda)$.

The reasoning remains valid with several points satisfying (A.12).

Appendix A.3. Long-run land price and wealth

Rent dynamics and long-run wealth for non-constrained dynasties. We define a non-constrained dynasty a sequence of generations living at the city fringe such that the active agent and all her descendants are never constrained. Let us consider the dynamics of land price $p_t(x) = \Psi_t(x, \mathcal{K})$ expressed as follows

$$p_t(x) = \mathcal{K} - \tilde{\kappa}(x) + \mu^2 p_{t+2}(x), \tag{A.13}$$

where $\mathcal{K} \equiv \tilde{\kappa}(L_0) + (1-\mu^2)R_A$. We consider that there are some locations x such that these dynamics hold at any date t. In particular, if at some date t, $\underline{g}_t > R_A(1-\lambda)$, and if the sequence $\{\underline{g}_\varsigma\}_{\varsigma=t}^{\infty}$ is monotonously increasing, then these rent dynamics prevail at the city fringe. In this case, the rent can thus be solved by iterating forward the system. Hence,

$$p_t(x) = \lim_{\varsigma \longrightarrow \infty} \mu^{2(\varsigma+1)} p_{t+2+\varsigma}(x) + \lim_{\varsigma \longrightarrow \infty} \sum_{j=0}^{\varsigma} \mu^{2j} \left[\mathcal{K} - \tilde{\kappa}(x)\right].$$

In order to find a solution, we abstract from the presence of any housing bubble and we impose a transversality condition, that is, $\lim_{t\to\infty} p_t(x) < \infty$. As $\mu < 1$, we obtain

$$p_\infty(x) = \frac{\mathcal{K} - \tilde{\kappa}(x)}{1-\mu^2}. \tag{A.14}$$

From (21), the wealth dynamics for non-constrained agents are given by

$$g_{t+1}^{\star} = \gamma\Lambda\left(w + \frac{\alpha+\gamma}{\gamma} g_t - \mathcal{K} - \tilde{\tau}\right) \tag{A.15}$$

where

$$\tilde{\tau} = \frac{(1+\mu)G}{L} \quad \text{and} \quad \Lambda \equiv \frac{\delta^{1/\sigma}\mu^{(1+\alpha)(1-1/\sigma)-1}}{1+\delta^{1/\sigma}\mu^{(1+\alpha)(1-1/\sigma)}}.$$

Given (A.15), if $(\alpha+\gamma)\Lambda < 1$, then the wealth of non-constrained dynasties living at the city fringe converges to

$$g_\infty = \frac{\gamma\Lambda(w - \tilde{\tau} - \mathcal{K})}{1-(\alpha+\gamma)\Lambda} \equiv g_\infty^{\mathcal{N}}. \tag{A.16}$$

At the steady state, these dynasties reside at $x \in [\widehat{x}_\infty, L_0^\star]$, where $\widehat{x}_\infty$ is given by $(1-\lambda)\, p_\infty(\widehat{x}_\infty) = g_\infty^{\mathcal{N}}$. Note that if $\widehat{x}_\infty$ exists, it is necessarily unique. By rearranging terms, we finally obtain

$$\kappa(\widehat{x}_\infty^\star) = \mathcal{K} - \frac{w - \tilde{\tau} - \mathcal{K}}{\rho(\lambda)} \quad \text{with} \quad \rho(\lambda) \equiv \frac{1-(\alpha+\gamma)\Lambda}{\gamma\Lambda}\frac{1-\lambda}{1-\mu^2}. \tag{A.17}$$

Note that $b_\infty = \mu(y_\infty - p_\infty)$ is positive when $\alpha g_\infty/(\gamma\mu) > \mathcal{K}/(1-\mu^2)$. Since we must have $g_\infty/(1-\lambda) > \mathcal{K}/(1-\mu^2)$ (*i.e.* $R_A < \underline{R}_A$). When $1-\lambda > \gamma\mu/\alpha$, we have $b_\infty > 0$ when $R_A < \underline{R}_A$. Hence, we assume that

$$\lambda < 1 - \frac{\gamma\mu}{\alpha} \equiv \bar{\lambda} \tag{A.18}$$

Long run wealth and rent of constrained dynasties. Consider now constrained dynasties. They pay the rent $p_t^\star(x) = g_t^\star(x)/(1-\lambda)$ at each t. Note that $b_{t+1}(x) > 0$ if and only if $y_{t+2}(x) - g_{t+2}(x)/(1-\lambda) > 0$ or, equivalently, $\lambda < \bar{\lambda}$. Thus, the dynamics of the equilibrium rent $p_t^\star(x)$ follows the dynamics of $g_t^\star(x)$. From (21), constrained agents follow the dynamics:

$$g_{t+1}(x) = \gamma\Lambda\left[w + \frac{\alpha+\gamma}{\gamma}g_t + \mu^2\frac{g_{t+2}(x)}{1-\lambda} - \tilde{\kappa}_t(x) - \frac{g_t}{1-\lambda} - \tilde{\tau}_t\right].$$

These dynamics can be rewritten as the following second-order difference equation

$$g_{t+2}(x) + a_1 g_{t+1}(x) + a_2 g_t + a_3 = 0$$

with $a_1 \equiv -\frac{(1-\lambda)}{\mu^2\gamma\Lambda}$, $a_2 \equiv \frac{1-\lambda}{\mu^2}\left(\frac{\alpha+\gamma}{\gamma} - \frac{1}{1-\lambda}\right)$ and $a_3 \equiv (1-\lambda)\left(w - \tilde{\kappa}(x) - \tilde{\tau}_t\right)/\mu^2$.

A particular solution of the above equation is

$$g_\infty(x) = \frac{-a_3}{1+a_1+a_2} = \frac{\gamma\Lambda}{1-(\alpha+\gamma)\Lambda}\frac{w-\tilde{\tau}-\tilde{\kappa}(x)}{1+1/\rho(\lambda)} \equiv g_\infty^{\mathcal{C}}(x) \tag{A.19}$$

The deviations from this value will be given by the general solution of the homogenous equation

$$g_{t+2}(x) + r_1 g_{t+1}(x) + r_2 g_t = 0.$$

The characteristic equation is $r^2 + a_1 r + a_2 = 0$. The necessary and sufficient conditions so that the roots z_1 and z_2 of the characteristic equation are real and distinct are such that the discriminant

$$a_1^2 - 4a_2 = \frac{(1-\lambda)^2}{\mu^4\gamma^2\Lambda^2} - \frac{4(1-\lambda)}{\mu^2}\frac{\alpha+\gamma}{\gamma} + \frac{4}{\mu^2}$$

is positive which holds as long as $1 > (\alpha+\gamma)\Lambda$. Indeed, $a_1^2 - 4a_2$ achieves its minimum value when $1-\lambda = 2(\alpha+\gamma)\gamma\Lambda^2\mu^2$. Plugging this equality into $a_1^2 - 4a_2$ implies sign $\{a_1^2 - 4a_2\} =$ sign $\{1-\mu^2(\alpha+\gamma)\Lambda\}$. As $\mu^2 < 1$, $a_1^2 - 4a_2 > 0$ when $1 > (\alpha+\gamma)\Lambda$. Further, as $1 + a_1 + a_2 < 0$, there is at least one real root which is larger than 1. It remains to check whether the smallest real root, denoted by r_1, is lower than 1 in absolute value which is equivalent to $1 - a_1 + a_2 > 0$. This condition is satisfied when $\lambda \in (0, \bar{\lambda})$. Hence, households with wealth $g_\infty^{\mathcal{C}}(x)$ live in the area $[0, \widehat{x}_\infty]$, with $\widehat{x}_\infty$ defined by $g_\infty^{\mathcal{C}}(\widehat{x}_\infty) = (1-\lambda)\Psi_\infty(\widehat{x}_\infty, \widehat{K}_\infty)$ and it can be checked that it leads to (A.17).

The long-run return to location x for constrained agents is

$$\varrho_\infty^{\mathcal{C}}(x) = \frac{\frac{g_\infty^{\mathcal{C}}(x)}{1-\lambda} - R_A}{\frac{g_\infty^{\mathcal{C}}(x)}{1-\lambda} + \tilde{\kappa}(x) - [R_A + \tilde{\kappa}(L)]} \tag{A.20}$$

where $\varrho_\infty^{\mathcal{C}}(x) > \mu^{-2}$ is equivalent to

$$(1-\mu^2)R_A + \tilde{\kappa}(L) - \tilde{\kappa}(x) > \frac{g_\infty^{\mathcal{C}}(x)}{1-\lambda}. \tag{A.21}$$

Because the left-hand side of this inequality is equal to $\Psi_\infty(x, \mathcal{K})$, which is higher than $g_\infty^{\mathcal{C}}(x)/(1-\lambda)$, $\varrho_\infty^{\mathcal{C}}(x)$ is higher than $(1+r)^2$ when agents are credit constrained.

Long-run urban configurations. **Claim** *In the long run, (i) non-constrained agents always live further away than constrained agents and (ii) there can be at most one threshold location $\widehat{x}_\infty$.*

(i) By contradiction, assume a long run urban equilibrium with non-constrained agents living in $[0, \widehat{x}_\infty]$ and constrained-agents living in $[\widehat{x}_\infty, L_0^\star]$. The non-constrained agents would pay the rent $\Psi_\infty(x, \widehat{K}_\infty)$ with $\widehat{K}_\infty$ such that $g_\infty(\widehat{x}_\infty) = \Psi_\infty(\widehat{x}_\infty, \widehat{K}_\infty)(1-\lambda)$ and, given the dynamics (A.15), they would all end up with the same long-run wealth level $g_\infty^{\mathcal{N}} = g_\infty(\widehat{x}_\infty) = \Psi_\infty(\widehat{x}_\infty, \widehat{K}_\infty)(1-\lambda)$. By assumption, for any $x \in [0, \widehat{x}_\infty[$, $g_\infty^{\mathcal{N}} > \Psi_\infty(x, \widehat{K}_\infty)(1-\lambda)$ which is a contradiction as Ψ_∞ is stricly decreasing with x.

(ii) By contradiction, assume, w.l.o.g., a long run urban equilibrium with 3 threshold-locations denoted by $\widehat{x}_\infty^1, \widehat{x}_\infty^2, \widehat{x}_\infty^3$ such that there are constrained residents in $\left[0, \widehat{x}_\infty^1\right]$ and $\left[\widehat{x}_\infty^2, \widehat{x}_\infty^3\right]$ and non-constrained agents in between. Applying the same logic as in item (i) there cannot be constrained agents further away from the CBD than non-constrained agents. According to this claim, the long run spatial sorting is that constrained agents live close to the CBD and non-constrained further away.

Agregate land rent. The present value of aggregate differential land rents denoted by ALR_∞ equals

$$\begin{aligned}
\text{ALR}_\infty &= \frac{1}{1-\delta}\int_0^{\widehat{x}_\infty^\star}\left[\frac{g_\infty^{\mathcal{C}}(x)}{1-\lambda} - R_A\right]\mathrm{d}x + \frac{1}{1-\delta}\int_{\widehat{x}_\infty^*}^{L^o/2}\left[\Psi_\infty(x,\mathcal{K}) - R_A\right]\mathrm{d}x,\\
&= \frac{1}{1-\delta}\left\{\int_0^{L^o/2}\left[\Psi_\infty(x,\mathcal{K}) - R_A\right]\mathrm{d}x - \int_0^{\widehat{x}_\infty^\star}\left[\Psi_\infty(x,\mathcal{K}) - \frac{g_\infty^{\mathcal{C}}(x)}{1-\lambda}\right]\mathrm{d}x\right\},\\
&= \frac{1}{1-\delta}\left\{\int_0^{L^o/2}\left[\frac{\tilde{\kappa}(L^o/2) - \tilde{\kappa}(x)}{1-\mu^2}\right]\mathrm{d}x - \int_0^{\widehat{x}_\infty^\star}\Theta(x)\mathrm{d}x\right\},
\end{aligned}$$

where $\widehat{x}_\infty^*$ is defined so that $\Psi_\infty(\widehat{x}_\infty^\star, \mathcal{K}) = g_\infty^{c}(\widehat{x}_\infty^\star)/(1-\lambda)$ and $g_\infty^{\mathcal{C}}(x)$ is the wealth of credit-constrained agent.

Equilibrium city size. When there is no borrowing constraint, we must have $U(g_t) = U^\star(g_t)$ at the spatial equilibrium where $U^\star(g_t)$ is the common level of utlity of g_t-agents. The indifference condition $U(g_t) = U^\star(g_t)$ is equivalent to $W(g_t) = W^\star(g_t)$, $W^\star(g_t)$ being the common level of wealth of g_t-agents while the net lifetime wealth of young agent is

$$W(g_t) = w + g_t\left(1 + \frac{\alpha}{\gamma}\right) - \mathcal{K}_t - \frac{G}{L_{0,t} + L_{1,t}} - \mu\frac{G}{L_{0,t} + L_{0,t+1}}. \tag{A.22}$$

where $\mathcal{K}_t = (1-\mu^2)R_A + \tilde{\kappa}_t(L_{0,t})$ which is common to all types of agents, but depends on city size.

$W(g_t) = W^\star(g_t)$ can be solved for the equilibrium city size $L^\star_{0,t}$ as a function of $W^\star(g_t)$. This equilibrium is stable only if the marginal utility decreases with city size for all cities with a positive equilibrium population (Behrens and Robert-Nicoud, 2015). The relationship between the utility level of agents and population size follows the same inverted-U curve pattern. Indeed, it is straightforward to check that

$$\text{sign}\,\frac{\partial W(g_t)}{\partial L_{0,t}} = \frac{G}{(L_{1,t}+L_{0,t})^2} + \frac{\mu G}{(L_{0,t}+L_{0,t+1})^2} - \frac{\partial \tilde{\kappa}_t(L_{0,t})}{\partial L_{0,t}} \tag{A.23}$$

which is positive (resp. negative) when $L_{0,t}$ is low (resp. high) enough and $W(g_t)$ reaches its maximum value when the population size is such that $(1+\mu)G/(2L_{0,t})^2 = \partial\tilde{\kappa}_t/\partial L_{0,t}$ which has a single solution. If, for a given $W^\star(g_t)$, $W(g_t) = W^\star(g_t)$ has two solutions which is common to all types of young agents, the larger solution $L^\star_{0,t}$ is the unique stable equilibrium as the resulting equilibrium population level is only stable when utility is decreasing with city size. Therefore, even though there exist many decentralized equilibria that simultaneously satisfy the indifference condition $W(L^\star, g^{\mathcal{N}}_\infty(L^\star)) = W^\star$, and the stability condition $\partial V/\partial L < 0$ at $L = L^\star$, cities are oversized in any equilibrium with free migration because individuals do not take into account the negative impact they impose on other agents when making their location decisions (Behrens and Robert-Nicoud, 2015).

Appendix A.4. Decentralization and land taxation

Second best. The policymaker acquires the land needed for each city from landowners (the agents) at the amount $\mathcal{R}_t \geq R_A$ and chooses the population size L and the taxation scheme in each city as well as agents' location, the consumption of the composite good (c_t and d_{t+1}) and the bequests (g_{t+1} and b_{t+1}). Formally, the policy maker maximizes $\sum_{t=0}^{\infty}\delta^t U(c_t, D_{t+1})$ under the per capita intertemporal resource constraint of the economy is:

$$w + g_t + \mu y_{t+1} = c_t + \mu\left(d_{t+1} + g_{t+1} + b_{t+1}\right) + \frac{\mathcal{R}_t}{2} + \mu\frac{\mathcal{R}_{t+1}}{2} + (1+\mu)\,\Omega(L) \tag{A.24}$$

with $\delta \leq 1$ being the social discount factor and where $y_{t+1} = \mu^{-1}b_t + \mathcal{R}_{t+1}$, where $\mathcal{R}_{t+1}$ is the value of the land property right. When bequests are positive, the first order conditions are

$$\mu\frac{\partial U_t}{\partial c_t} = (1-\alpha-\gamma)\frac{D_{t+1}}{d_{t+1}}\frac{\partial U_t}{\partial D_{t+1}}, \tag{A.25}$$

$$\mu\frac{\partial U_t}{\partial c_t} = \gamma\frac{D_{t+1}}{g_{t+1}}\frac{\partial U_t}{\partial D_{t+1}} + \delta\frac{\partial U_{t+1}}{\partial c_{t+1}} \tag{A.26}$$

$$\mu\frac{\partial U_t}{\partial c_t} = \frac{\alpha}{\mu}\frac{D_{t+1}}{y_{t+2}}\frac{\partial U_t}{\partial D_{t+1}} + \delta\frac{\partial U_{t+1}}{\partial c_{t+1}} \tag{A.27}$$

A higher voluntary bequest (g_{t+1}, b_{t+1}) reduces the consumption of donors, generating a utility loss, while it raises not only the current welfare of donors (joy-of-giving effect) but also the welfare of the next generation as the consumption of donees increases. Given a CRRA utility function and assuming bequests are positive, the optimality conditions associated with d_{t+1}, g_{t+1}, b_{t+1} imply

$$g^o_\infty = \frac{\gamma}{1-\alpha-\gamma}\frac{d^o_\infty}{1-\delta/\mu}, \quad y^o_\infty = \frac{\alpha}{1-\alpha-\gamma}\frac{1}{\mu}\frac{d^o_\infty}{1-\delta/\mu}, \quad g^o_\infty = \frac{\gamma\mu}{\alpha}y^o_\infty \tag{A.28}$$

at the steady state as well as

$$d^o_\infty = (1-\alpha-\gamma)\delta^{1/\sigma}\mu^{\alpha(1-1/\sigma)-1/\sigma}(1-\delta/\mu)^{(\alpha+\delta)/(\sigma-1)}c^o_\infty. \tag{A.29}$$

Given the specification of our utility function (impure altruism), the decentralized city implies that the ratio g_∞/c_∞ and g_∞/d_∞ is lower than that of the planner. Indeed, agents do not take fully into account the infinite stream of their descendants' utilities. To attain an efficient allocation of resources between consumption and intergenerational transfer, the city government has to implement a standard Pigouvian tax on consumption equal to $1/(\mu/\delta-1) \equiv \tau^c_\infty$, where tax revenues are recycled in a lump-sum transfer scheme.[1] This is a well-known feature of models with impure altruism (Michel and Pestieau, 2004). By contrast, in the presence of pure altruism (parents derive utility from the offspring's utility), the ratio of inherited wealth to consumption in the *laissez-faire* equilibrium is efficient.

Lump sum tax and transfer. When the economy is decentralized, the governement's tax instruments consist of a tax on inherited land wealth τ_2, a location-specific tax $\tau_1(x)$ paid by old agents, and a location-specific tax $\tau_0(x)$ paid by borrowing-constrained young agents to cover public expenditures G. In addition, like in OLG models, lump-sum taxes and transfers are needed to achieve the optimal configuration. The optimal transfer to each young agent, denoted by S, is financed by a lump sum tax paid by each old agent at the same period, denoted by T. The fiscal scheme is such that the household budget constraint is now given by

$$w + g_t - p_t(x) + S = c_t + s_t + \kappa_0(x) + \mathbb{1}_C\tau_0(x), \tag{A.30}$$

$$\mu^{-1}s_t + y_{t+1} = d_{t+1} + g_{t+1} + b_{t+1} + \kappa_1(x) + \tau_1(x) + T, \tag{A.31}$$

where $\mathbb{1}_C = 1$ if the agent is credit constrained in equilibrium and 0 otherwise while the expression of *post mortem* transfers is

$$y_{t+2} = \mu^{-1}b_{t+1} + p_{t+2}(x) - \tau_2. \tag{A.32}$$

As a result, the intertemporal budget constraint writes

$$\begin{aligned} c^\star_t + \mu d^\star_{t+1} + \mu g^\star_{t+1} + \mu^2 y^\star_{t+2} - g^\star_t - \mu y^\star_{t+1} = {} & w + \mu^2 p^\star_{t+2}(x) - p^\star_t(x) - \tilde{\kappa}(x) \\ & - \mathbb{1}_C\tau_0(g) - \mu\tau_1(x) - \mu^2\tau_2 \\ & - \mu T + S. \end{aligned} \tag{A.33}$$

The land rent $p^\star_t(x)$ then equals

$$p^\star_t(x) = R_A + \kappa_0(L^o/2) - \kappa_0(x) + \mu^2\left[p^\star_{t+2}(x) - R_A - \tau_2\right] + \mu\left[\kappa_1(L^o/2) - \kappa_1(x) - \tau_1(x)\right] \tag{A.34}$$

By setting

$$\tau_0(x) = \Theta(x) \tag{A.35}$$

$$\tau_1(x) = \kappa_1(L^o/2) - \kappa_1(x) \tag{A.36}$$

$$\tau_2(x) = p_{t+2}(x) - R_A, \tag{A.37}$$

[1] Equivalently, the government could implement a Pigouvian subsidy on voluntary bequests equal to δ/μ financed by a lump-sum tax.

the market land price coincides with shadow land price ($p_t^\star(x) = R_A + \kappa_0(L^o/2) - \kappa_0(x)$). Note that, at the stationary state, the fiscal revenue from $\tau_0(x)$, $\tau_1(x)$, and $\tau_2(x)$ is equal to public expenditures G, as

$$\int_0^{L^o/2} \tau_2(x)\mathrm{d}x + \int_0^{L^o/2} \tau_1(x)\mathrm{d}x + \int_0^{\widehat{x}_\infty^\star} \tau_0(x)\mathrm{d}x = \mathrm{ASLR}_t. \tag{A.38}$$

Given the proceeds of taxation equal total subsidy $T = S$, and that T is set so that the RHS of (A.33) equals $\mathcal{I}_\infty(L^o, R_A)$ given by (see Section 5)

$$\mathcal{I}_\infty(L^o, R_A) = w - (1-\mu^2)R_A + \frac{(1-\mu)R_A}{2} - \frac{(1+\mu)\sum_z \kappa_z(L^o/2)}{2}, \tag{A.39}$$

we get

$$\begin{aligned} S - \mu T =& \frac{1-\mu}{2}R_A \\ &+ \left[\kappa_0(L^o/2) - \frac{\kappa_1(L^o/2)+\kappa_0(L^o/2)}{2}\right] + \mu\left[\kappa_1(L^o/2) - \frac{\kappa_1(L^o/2)+\kappa_0(L^o/2)}{2}\right] \end{aligned}$$

so that

$$T = \frac{R_A + \kappa_0(L^o/2) - \kappa_1(L^o/2)}{2} > 0. \tag{A.40}$$

In this case, the old agents pay

$$\tau_1(x) + T = \frac{R_A + \kappa_0(L^o/2) + \kappa_1(L^o/2)}{2} - \kappa_1(x) \tag{A.41}$$

$$= R_A + \kappa_1(L^o/2) - \kappa_1(x) + \frac{\kappa_0(L^o/2) - \kappa_1(L^o/2) - R_A}{2} \tag{A.42}$$

which includes the shadow land price $R_A + \kappa_1(L^o/2) - \kappa_1(x)$ and a term capturing the transfer between generations to smooth consumption over the lifecycle.

Note that the land rent paid by young agents minus the transfer is equal to

$$p(x) - T = R_A + \kappa_0(L^o/2) - \kappa_0(x) - \frac{R_A + \kappa_0(L^o/2) - \kappa_1(L^o/2)}{2} \tag{A.43}$$

$$= \frac{R_A + \kappa_0(L^o/2) + \kappa_1(L^o/2)}{2} - \kappa_0(x) \tag{A.44}$$

so that $p(x) - T + \kappa_0(x) = \tau_1(x) + T + \kappa_1(x)$, that is the urban costs adjusted by taxation scheme do not vary over the lifecycle.

Steady-state wealth and property tax. We consider a standard uniform tax on residential property τ^p instead of a lump-sum tax when the city size is given by $L = L^o$ and there is no borrowing constraint. Therefore, the budget constraint becomes

$$w + g_t = c_t + s_t + p_t(x)(1+\tau^p) + \kappa_0(x), \tag{A.45}$$

$$\mu^{-1}s_t + y_{t+1} = d_{t+1} + g_{t+1} + b_{t+1} + p_{t+1}(x)\tau^p + \kappa_1(x). \tag{A.46}$$

with $b_{t+1} = \mu y_{t+2} - \mu p_{t+2}$. It is easy to check that the equilibrium land price is then given by

$$p_\infty(x) = R_A + \frac{\tilde{\kappa}(L^o/2) - \tilde{\kappa}(x)}{1 + (1+\mu)\tau^p - \mu^2} \tag{A.47}$$

A higher property tax is capitalized into lower land prices. The capitalization of the property tax into land prices offers support for the idea that property taxes could efficiently finance public expenditures. The budget constraint at each period can be expressed as follows

$$
\begin{aligned}
G &= 2\tau^p \frac{1-\mu^2+(1+\mu)\tau^p}{1-\mu^2+(1+\mu)\tau^p} \int_0^{L^o/2} [p_\infty(x) - R_A]\,\mathrm{d}x + \tau^p L^o R_A \\
&= \frac{2\tau^p}{1-\mu^2+(1+\mu)\tau^p} \int_0^{L^o/2} [\tilde{\kappa}(L^o/2) - \tilde{\kappa}(x)]\,\mathrm{d}x + \tau^p L^o R_A \\
&= \frac{\tau^p \mathbf{D}}{1-\mu^2+(1+\mu)\tau^p} + \tau^p L^o R_A
\end{aligned}
$$

where

$$
\mathbf{D} \equiv 2\int_0^{L^o/2} [\tilde{\kappa}(L^o/2) - \tilde{\kappa}(x)]\,\mathrm{d}x > 0.
$$

The equilibrium property tax is implicitly given by

$$
\frac{G}{L^o} = \tau^p R_A + \frac{\tau^p}{1-\mu^2+(1+\mu)\tau^p}\frac{\mathbf{D}}{L^o} \tag{A.48}
$$

if $\frac{G}{L^o} > R_A + \frac{1}{1-\mu^2+(1+\mu)}\frac{\mathbf{D}}{L^o}$ and since the RHS of (A.48) increases with τ^p. Hence, the left wealth is $g_{t+1} = \gamma\Lambda[w + (\alpha+\gamma)g_t/\gamma - (1-\mu^2+(1+\mu)\tau^p)R_A - \tilde{\kappa}(L^o/2)]$ so that the steady-state wealth can be written as

$$
\begin{aligned}
g_\infty^p &= \frac{\gamma\Lambda}{1-(\alpha+\gamma)\Lambda}\left[w - (1+\mu)\frac{G}{L^o} - \mathcal{K} + (1+\mu)\left(\frac{G}{L^o} - \tau^p R_A\right)\right] \\
&= \frac{\gamma\Lambda}{1-(\alpha+\gamma)\Lambda}\left[w - (1+\mu)\frac{G}{L^o} - \mathcal{K} + \frac{(1+\mu)\tau^p}{1-\mu^2+(1+\mu)\tau^p}\frac{\mathbf{D}}{L^o}\right]
\end{aligned} \tag{A.49}
$$

which is higher than $g_\infty^{\mathcal{N}}$ as $\mathbf{D} > 0$.

Appendix B. Richer models of residential location and wealth transmission

Our urban model predicts that without a borrowing limit, the housing market mechanism intrinsically promotes the convergence of the wealth of different agents. Credit market imperfections modify the distribution of surplus across agents because the borrowing constraint can cap the house price paid by wealthy or lucky agents. We show that our findings are not specific to the model considered thus far. In each extension, we consider that the CRRA utility function used in Section 4 and the interest rate r is assumed to be constant over time. Remember that $\mu \equiv 1/(1+r)$.

B.1. Heterogeneous workers

Assume that at the beginning of life, agents have to decide whether to invest in their own education before entering the labor and housing markets. Education increases their human capital, yielding more efficiency units of labor. If agents do not invest in education, they acquire one efficiency unit of labor (unskilled workers). Given that the wage rate per unit of efficient labor is w, their labor income is w (as in Section 4). Alternatively, the agents may invest in education at fixed cost ϕ. In this case, they supply

$e > 1$ efficiency units of labor (skilled workers) and their labor income achieves $we > w$. Therefore, agents invest in education if both conditions $g_t > w$ and $\phi < w(e-1) \equiv \phi^e$ are met.[2]

First, assume that there are no imperfections in the mortgage market. Under these circumstances, the bid-rent curves are given by $\Psi_t(x, \mathcal{K})$, which does not depend on wage and wealth, so there is no spatial wealth sorting. The dynamics for each dynasty's wealth in this economy are therefore given by the transition rule $g_{t+1} = \gamma\Lambda(w - \tilde{\tau} + (\alpha+\gamma)g_t/\gamma - \mathcal{K})$, if $g_t < \phi$, and $g_{t+1} = \gamma\Lambda(w - \tilde{\tau} + \phi^e + (\alpha + \gamma)g_t/\gamma - \phi - \mathcal{K})$, if $g_t > \phi$. Unsurprisingly, there are two equilibria, and long-run outcomes depend on the initial condition. The long-run wealth of agents with initial wealth lower than ϕ will converge to $g_\infty^{\mathcal{N}}$ given by (A.16), while the long-run wealth of those with greater inherited wealth will converge to $g_\infty^{\mathcal{N}} + \gamma\Lambda(\phi^e - \phi)/(1 - (\alpha+\gamma)\Lambda)$. In this context, long-term wealth inequality is driven by educational investment constraints rather than the mortgage market and depends on the initial condition. If all agents have the same initial wealth, there is no symmetry-breaking, and all agents become either skilled workers or unskilled workers.

Assume now that some unskilled workers face credit constraints when attempting to obtain a mortgage. We seek to build a steady state where there is spatial wealth sorting: cities host unconstrained skilled agents, unskilled agents who are credit constrained, and unskilled agents who are unconstrained. The steady state wealth of unskilled workers who are credit constrained is $g_\infty^{\mathcal{C}}(x)$ (see (A.19)) and the steady state wealth of unskilled workers who are unconstrained is $g_\infty^{\mathcal{N}}$ (see (A.16)). The residential area of the latter class is still $[\widehat{x}_\infty^\star, L]$, where $\widehat{x}_\infty^\star$ is given by (A.17). The residential area formed by agents investing in education is $[0, \widehat{x}_\infty^e]$ where $\widehat{x}_\infty^e$ is such that $g_\infty^{\mathcal{C}}(\widehat{x}_\infty^e) = \phi$ yielding

$$\tilde{\kappa}(\widehat{x}_\infty^e) = \phi\frac{1-(\alpha+\gamma)\Lambda}{\gamma\Lambda}\left[\Phi(\phi) - 1 - \frac{1}{\rho(\lambda)}\right] \quad \text{where} \quad \Phi(\phi) \equiv \frac{(w-\tilde{\tau})\gamma\Lambda}{\phi[1-(\alpha+\gamma)\Lambda]}. \tag{B.1}$$

In addition, if a class of skilled workers exists, their bid rent is $\Psi_\infty^e(x, K_\infty^e)$, where K_∞^e is the integration constant of the bid rent of skilled workers. As $\Psi_\infty^e(\widehat{x}_\infty^e, K_\infty^e) = g_\infty^{\mathcal{C}}(\widehat{x}_\infty^e)/(1-\lambda)$ must hold in equilibrium, $K_\infty^e = \phi[\Phi(\phi) - 1][1-(\alpha+\gamma)\Lambda]/(\gamma\Lambda)$, with $K_\infty^e - \tilde{\kappa}(\widehat{x}_\infty^e) > 0$ and $K_\infty^e < \mathcal{K}$ (so that $\Psi_\infty^e(x, K_\infty^e) < \Psi_\infty(x, \mathcal{K})$). Since the lifetime wealth of skilled workers is $we - \phi - K_\infty^e - \tilde{\tau}$, their wealth converges to

$$g_\infty^e = \frac{\gamma\Lambda(\phi^e - \phi)}{1-(\alpha+\gamma)\Lambda} + \phi. \tag{B.2}$$

Therefore, some workers invest in education if $g_\infty^e > \phi$ and $\widehat{x}_\infty^e \in (0, \widehat{x}_\infty^\star)$. The former condition holds as long as $\phi < \phi^e$. The condition $\widehat{x}_\infty^e > 0$ occurs when

$$\lambda < \lambda^e \equiv 1 - \frac{\gamma\Lambda}{1-(\alpha+\gamma)\Lambda}\frac{1-\mu^2}{\Phi(\phi)-1} \tag{B.3}$$

which decreases with ϕ (notice that $\Phi(\phi) > 1$ as $K_\infty^e > \tilde{\kappa}(\widehat{x}_\infty^e) > 0$). We have $\widehat{x}_\infty^e < \widehat{x}_\infty^\star$ if and only if $\tilde{\kappa}(\widehat{x}_\infty^e) < \tilde{\kappa}(\widehat{x}_\infty^\star)$. Using (26) and (B.1), the latter condition is equivalent to $g_\infty^{\mathcal{N}} < \phi$ which always hold when skilled workers and unskilled workers emerge at the steady state. Furthermore, at the steady state, the emergence of two types of unskilled agents (credit constrained and unconstrained) requires

[2]For simplicity and without loss of generality, agents with wealth $g_t \leq \phi$ cannot borrow to cover the costs of education.

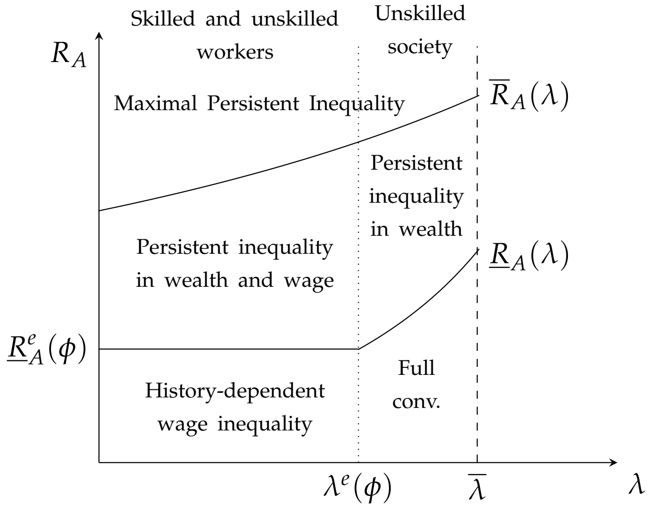

Figure 2: Wealth and wage inequalities in the (λ,R_A)-space

$(1-\lambda)R_A < g^n_\infty < (1-\lambda)\Psi_\infty(\widehat{x}^e_\infty, \mathcal{K})$ or, equivalently,

$$\underline{R}^e_A(\phi) \leq R_A \leq \overline{R}_A(\lambda) \quad \text{with} \quad \underline{R}^e_A(\phi) \equiv \frac{w - \tilde{\tau} - \tilde{\kappa}(L^o/2)}{1-\mu^2} - \frac{\phi}{1-\mu^2}\frac{1-(\alpha+\gamma)\Lambda}{\gamma\Lambda} \tag{B.4}$$

and $\overline{R}_A$ is given by (28). $\overline{R}_A(\lambda) > \underline{R}^e_A(\phi) > \underline{R}_A(\lambda)$ as long as $\widehat{x}^e \in (0, L/2)$ and with $\underline{R}^e_A(\lambda^e) = \underline{R}_A(\lambda^e)$.[3] In the presence of skilled workers, the unskilled workers who are unconstrained reside in remote areas $(\widehat{x}^\star_\infty, L^\star/2)$ while the unskilled workers who are credit constrained live in $[\widehat{x}^e_\infty, \widehat{x}^\star_\infty]$ when $\underline{R}^e_A(\phi) \leq R_A \leq \overline{R}_A(\lambda)$. We summarize our findings in Figure 2 and in the following proposition

Proposition 7. *Assume that* $(\alpha+\gamma)\Lambda < 1$, $\lambda < \overline{\lambda}$, *and* $\phi < \phi^e$, the long run city is characterized by one of the following steady-state distribution:

(i) If $\underline{R}^e_A < R_A < \overline{R}_A$ *and* $\lambda < \min\left\{\lambda^e, \overline{\lambda}\right\}$*, then three types of agents emerge in equilibrium: high-wealth agents (skilled workers), middle-wealth agents (unskilled workers who are credit-constrained), and low-wealth agents (unskilled workers who are unconstrained).*

(ii) If $\overline{R}_A < R_A$ *and* $\lambda < \min\left\{\lambda^e, \overline{\lambda}\right\}$*, then two types of agents emerge in equilibrium: skilled workers and unskilled workers who are credit-constrained.*

(iii) If $R_A \leq \min\{\underline{R}^e_A(\phi), \underline{R}_A(\lambda)\}$ *and* $\lambda < \min\left\{\lambda^e(\phi), \bar{\lambda}\right\}$*, then long-run wealth inequality depends on the initial wealth distribution.*

(iv) If $\underline{R}_A < R_A$ *and* $\lambda^e < \lambda < \overline{\lambda}$*, then all agents are unskilled, and persistent inequality occurs.*

(v) If $R_A < \underline{R}_A$ *and* $\lambda^e < \lambda < \overline{\lambda}$*, then full convergence occurs.*

Therefore, mortgage market imperfections have an impact on occupational choice and spatial wealth sorting as well as spatial skill sorting. The equilibrium labor income relies on the interplay between the inherited wealth distribution and access to mortgages. As in Sections 3 and 4, the credit constraint forces the partition of agents within the city according to the distribution of wealth g_t, which also depends on

[3] It is straighforward to check that $\underline{R}^e_A = \underline{R}_A(\lambda) + \frac{1}{1-\mu^2}\frac{\rho}{1+\rho}\tilde{\kappa}(\widehat{x}^e)$ and $\overline{R}_A(\lambda) = \underline{R}_A(\lambda) + \frac{1}{1-\mu^2}\frac{\rho}{1+\rho}\tilde{\kappa}(L^o/2)$ so that $\overline{R}_A(\lambda) > \underline{R}^e_A > \underline{R}_A(\lambda)$.

occupational choice. Agents living in an area close to the city center are wealthier than agents residing in an area farther away. There is perfect skill sorting and perfect wealth sorting. The wealthiest agents (the skilled workers) live in the most attractive locations, while the middle-class agents (lucky unskilled workers) are credit-constrained and can live in better places than the other unskilled workers. Skilled agents also benefit from the presence of constrained agents because they face less competition in the housing market and pay a lower housing price. Hence, the emergence of wealthy classes who are not credit-constrained is not a transitory configuration, and their wealth converges to g_∞^e. Mortgage market imperfections magnify wage inequality and make wealth more unequally distributed than labor income.

In our context where occupational choices are affected by the distribution of wealth and mortgage market imperfections, taxation policy needs to be adjusted to reach the optimal outcome. Indeed, the presence of fixed costs associated with investment in education may generate an under-investment in human capital. The per-capita production is now $w(e\sigma + 1 - \sigma)$ where σ is the share of skilled workers in the economy. Furthermore, an amount of resources must now be allocated to the funding of per-capita education expenditures equal to $\phi\sigma$. It is straightforward to verify that, when $\phi < \phi^e$, the optimal outcome is that all agents are skilled workers. As a result, in the presence of mortgage market imperfections, wealth distribution has a long-lasting effect on aggregate income. To achieve the optimal solution, the tax design characterized by Proposition 5 is required to finance non-education public expenditures, leading the wealth distribution to mirror the labor income distribution. To finance the education of all workers, a lump-sum tax ϕ can be optimally implemented. In this case, all agents become skilled workers and their wealth converges to $g_\infty^{\mathcal{N}} + \gamma\Lambda(\phi^e - \phi)/[1 - (\alpha + \gamma)\Lambda]$ (unskilled workers are better off while no agent is worse off).

B.2. Heterogeneous housing size.

We now consider the case where the supply of housing units varies across locations and the urban wage is identical for agents and is the same for all periods ($w_t = w$). Housing units are available in discrete and fixed sizes $h_t \in \left\{\underline{h}, ..., \overline{h}\right\}$ where $\underline{h}$, and $\overline{h}$ represent the minimum and maximum sizes, respectively. In accordance with empirical evidence, we assume that housing size increases with distance to the CBD. For example, $h_t(x) = \underline{h}$ for $x \in (0, \ell]$ and $h_t(x) = \overline{h}$ for $x \in (\ell, L]$. The utility function is $u\left(c_t, h_t^s, d_{t+1}, g_{t+1}, y_{t+2}\right)$ where h_t^s represent housing services generated by the housing investment (for simplicity, we assume $h_t^s = h_t$) and $y_{t+2} = b_{t+1}/\mu + p_{t+2}(x)h_t(x)$. Agents face the budget constraint. Thus, the household budget constraint is written as

$$w + g_t - p_t(x)h_t(x) = c_t + s_t + \kappa_0(x) + \tau,$$
$$s_t/\mu + y_{t+1} = d_{t+1} + g_{t+1} + b_{t+1} + \kappa_1(x) + \tau.$$

Note that p_t is the price *per unit of housing*. Maximizing $u(.)$ under the budget constraint implies $u_C(.)p_t = u_{h^s}u(.)$ with $u_C \equiv \partial u/\partial c_t$ and $u_{h^s} \equiv \partial u/\partial h_t^s$. The maximum bid rent per unit of housing $\Psi_t(x, g_t, h_t)$ is such that $u'(.) = 0$ or, equivalently,

$$\Psi_t^{'}(x, g_t, h_t) = \frac{-\tilde{\kappa}'(x)}{h_t(x)} + \mu^2 p_{t+2}^{'}(x). \tag{B.5}$$

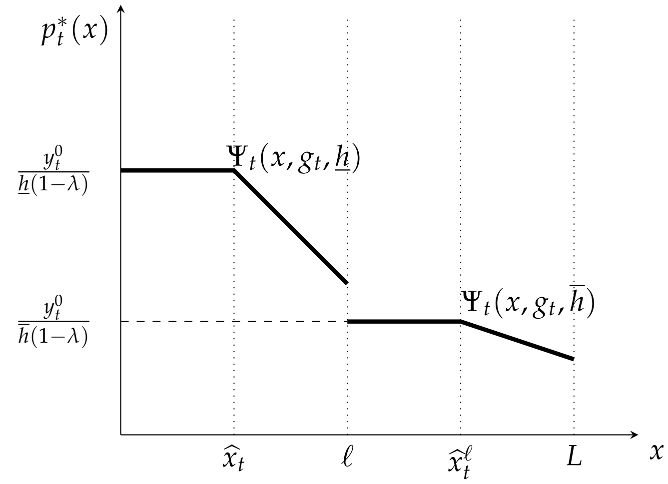

Figure 3: Urban equilibrium with two housing sizes and no wealth heterogeneity.

Contrary to the case with exogenous lot size, the slope of the bid rent depends on housing size. This is a modified version of the Alonso-Muth condition. Agents living far from the CBD are compensated for their long and costly commutes by enjoying larger housing. As a result, the maximum bid rent per unit of housing is

$$\Psi_t(x, g_t, h_t) = K_t - \frac{\tilde{\kappa}(x)}{h_t(x)} + \mu^2 p_{t+2}(x), \tag{B.6}$$

where $K_t = (1-\mu^2)R_A + \tilde{\kappa}(L_t)/\overline{h}$ if no agent is credit constrained (note that agents treat $h_t(x)$ as an exogenous parameter).

The borrowing constraint is now given by $\lambda p_t(x)h_t(x) \geq p_t(x)h_t(x) - g_t$, where $p_t(x)h_t(x)$ represents housing expenditures such that agents can borrow if and only if

$$p_t(x) \leq \frac{g_t}{1-\lambda}\frac{1}{h_t(x)} \equiv \widehat{p}_t(h). \tag{B.7}$$

When there are two sizes of housing, there are now at most two areas ($(0, \widehat{x}_t)$ and $(\ell, \widehat{x}_t^\ell)$) in which the borrowing constraint is binding for each agent (see Figure 3). The threshold locations $\widehat{x}_t \in (0, \ell)$ and $\widehat{x}_t^\ell \in (\ell, L)$ are such that $\Psi_t(\widehat{x}_t, g_t^\star(\widehat{x}_t), \underline{h}) = \widehat{p}_t(\underline{h})$ and $\Psi_t(\widehat{x}_t^\ell, g_t^\star(\widehat{x}_t^\ell), \overline{h}) = \widehat{p}_t(\overline{h})$, respectively. We can generalize this result to any number of housing size classes. The same mechanisms presented in the previous sections are at work. In different places, the equilibrium house prices paid by the credit-constrained agents and the wealthiest agents is lower than the maximum price that an agent would be willing to pay. As in the previous sections, credit constraints can give rise to symmetry-breaking and lead to spatial wealth sorting, which can translate into persistent inequality. A tax on both land rents and lifetime wealth must also be implemented to achieve a better allocation of resources.

B.3. Real estate industry.

Consider a real estate industry that operates under perfect competition and under constant returns to scale to produce an amount $h_t(x)$ of housing units at location x, using housing available at the end of date $t-1$ given by $h_{t-1}(x)$ and labor given as production factor. As in Section 4, the nondurable good is produced under constant returns and perfect competition, using labor only.

The free entry implies that the housing price $p_t(x)$ is equal to the unit cost $\mathcal{C}(\theta_t(x), w)$, which depends positively on $\theta_t(x)$ and on labor price. At the city limit, we have $p_t(L_{0,t}) = \mathbb{C}(R_A, w)$ while the bid rent of g_t-agents is $\Psi_t(x, g_t) = K_t - \tilde{\kappa}(x) + \mu^2\theta_{t+2}(x)$ when $x \in [\hat{x}_t(g_t, K_t), L_{0,t}]$ and $g_t/(1-\lambda)$ when $x \in [0, \hat{x}_t(g_t, K_t)]$. We show that $p_t'(x) = \theta_t'(x)$. As the total cost function is $\mathbb{C}(\theta_t(x), w)h_t(x)$, the demand for housing is

$$h_{t-1}(x) = \mathbb{C}_\theta(.)h_t(x) \tag{B.8}$$

according to the Shephard's lemma. As $p_t(x) = \mathbb{C}(\theta_t(x), w)$ and we assume that $h_{t-1}(x) = h_t(x) = 1$, then $p_t'(x) = \mathbb{C}_\theta(.)\theta_t'(x) = \theta_t'(x)$. Therefore, $\Psi_t'(x, g_t) = -\tilde{\kappa}'(x) + \mu^2 p_{t+2}'(x)$ as in Section 3. Hence, the introduction of a real estate industry does not modify qualitatively the spatial wealth sorting.

In addition, the introduction of a real estate industry does not change the wealth dynamics, given by equation (24). Therefore, the introduction of a real estate industry does not modify qualitatively the long run equilibrium. The long run housing value $\theta_\infty(x)$ is now such that $\mathbb{C}(\theta_\infty(x), w) = g_\infty^{\mathcal{C}}/(1-\lambda)$ when the borrowing constraint is binding and

$$\mathbb{C}(\theta_\infty(x), w) = \mathcal{K}_c - \tilde{\kappa}(x) + \mu^2\theta_\infty(x), \tag{B.9}$$

when the borrowing constraint is not binding, with $\mathcal{K}_c \equiv \mathbb{C}(R_A, w) - \mu^2 R_A + \tilde{\kappa}(L^o/2)$.

The shadow differential land rent at a distance x from the CBD corresponds to the resource saving from having an additional unit of land there. The shadow value of location at the city limit is $\mathbb{C}(R_A, w)) + \kappa(L^o/2)$. Moving an agent $z =$ young, old from the city limit to location $x < L^o/2$ would result in the instantaneous transport cost savings $\kappa_z(L^o/2) - \kappa_z(x)$ (notice that the use in labor to produce housing does not vary across space as $h_t(x) = h_{t-1}(x) = 1$). Hence, the shadow value of living at location x is $\mathbb{C}(s(x), x) + \kappa_z(x)$ where $s(x)$ is the shadow rent on land at a distance x from the CBD. Therefore, the shadow differential land rent at a distance x from the CBD is $s(x) - R_A$ where $s(x)$ is implicitely given

$$\mathbb{C}(s(x), w) - \mathbb{C}(R_A, w) = \kappa_0(L^o/2) - \kappa_0(x). \tag{B.10}$$

The optimal tax on inherited land asset $\tau_{t+2}^d(x)$ is still $\theta_{t+2}(x) - R_A$ leading to $y_{t+2} = (1+r)b_{t+1} + R_A$ and the optimal location-specific tax by old agents is still $\kappa_1(L^o/2) - \kappa_1(x)$ so that the housing price when the borrowing constraint is not binding is $p_t(x) = \mathbb{K} - \kappa_0(x)$ with $\mathbb{K} = \mathbb{C}(R_A, w) + \kappa_0(L^o/2)$ and $p_t(x) = \mathbb{C}(\theta(x), w)$, yielding

$$\mathbb{C}(\theta_t(x), w) - \mathbb{C}(R_A, w) = \kappa_0(L^o/2) - \kappa_0(x). \tag{B.11}$$

Using (B.10) and (B.11), it follows that $\theta_t(x) = s(x)$. thus, the introduction of a real estate industry does not alter qualitatively the optimal tax design.

B.4. Payment-to-income constraint

One simple way to write the payment-to-income constraint is such that the mortgage payment π is less than a fraction $\widetilde{\lambda}$ of income w

$$\pi \leq \widetilde{\lambda} w \tag{B.12}$$

with the mortgage payment being such that the discounted sum of mortgage payments over the mortgage length n equals the amount borrowed

$$p_t(x) - g_t = \pi \frac{1-(1+r)^{-n}}{r} = \pi \cdot \frac{\mu(1-\mu^n)}{1-\mu}. \tag{B.13}$$

In our set-up, we assume $n = 1$. Hence, the land rent paid by credit-constrained agents is now given by $p_t(x) = g_t + \lambda w$ with $\lambda \equiv \mu\widetilde{\lambda}$ and $\widehat{x}_t$ is now such that $\Psi_t(\widehat{x}_t, g_t) = g_t + \lambda w_t$ where $\widehat{x}_t$ is a decreasing function of g_t. Therefore Propositions 1, 2, and 3 remain valid. We still have symmetry breaking and wealth sorting.

Assuming the urban wage is the same for all periods, first-order conditions are given by (20), the wealth dynamics is

$$g_{t+1} = \gamma\Lambda \begin{cases} w[1-(1-\mu^2)\lambda] + \frac{\alpha}{\gamma} g_t + \mu^2 g_{t+2}[x^\star(g_t)] - \tilde{\kappa}_t[x^\star(g_t)] - \tilde{\tau} & \text{if borrowing-constrained,} \\ w + \frac{\alpha+\gamma}{\gamma} g_t - \mathcal{K} - \tilde{\tau} & \text{otherwise.} \end{cases} \tag{B.14}$$

As a result, the long-run wealth of credit-constrained households is

$$g_\infty^{\mathcal{C}}(x) = \frac{\gamma\Lambda\left\{w[1-(1-\mu^2)\lambda] + \tilde{\kappa}_t(x) - \tilde{\tau}\right\}}{1-\Lambda(\alpha+\mu^2\gamma)}, \tag{B.15}$$

that is a regular saddle point.

We can easily show that $\mathrm{d}g_\infty^{\mathcal{C}}(x)/\mathrm{d}\lambda < 0$, which means that a tighter PTI constraint (lower λ) leads to a higher long-run wealth for rich agents. Hence, as in the case of the LTV constraint, the PTI constraint generates inequality.

B.5 Borrowing costs

We now consider that, if $p_t(x) > g_t$, agents borrow and pay an interest rate $\zeta \geq r$, while we abstract from the presence of a downpayment requirement.[4] The budget constraint is now given by

$$\begin{aligned} w + g_t - p_t(x) &= c_t + s_t + \kappa_0(x) + \tau + \mathbb{1}_{\{p>g\}}\zeta\left[p_t(x) - g_t\right], \\ (1+r)s_t + y_{t+1} &= d_{t+1} + g_{t+1} + b_{t+1} + \kappa_1(x) + \tau. \end{aligned}$$

where $\mathbb{1}_{\{p_t>g\}}$ is a dummy variable that indicates if the young agent borrows. Under this configuration, the lifetime wealth of borrowers W_t^B and lenders W_t^S are expressed as follows

$$W_t^B = \widetilde{w} - \Upsilon_t(x) - (1+\zeta)[p_t(x) - g_t] + \mu y_{t+1} \quad \text{and} \quad W_t^S = \widetilde{w} - \Upsilon_t(x) - [p_t(x) - g_t] + \mu y_{t+1}.$$

[4]Note that $\zeta > r$ reflects credit market imperfections, based, for example, on exogenous default premium or market power in the banking industry.

with $\widetilde{w} \equiv w - \widetilde{\tau}$ and $\Upsilon_t(x) \equiv \widetilde{\kappa}_t(x) - \mu^2 p_{t+2}(x)$. Let us denote by $\Psi_t^B\left(x, K_t^B\right)$ and $\Psi_t^S\left(x, K_t^S\right)$ the bid rent of borrowers and savers, respectively, that solves the equilibrium condition $\mathrm{d}u(.)/\mathrm{d}x = 0$, so that

$$\Psi_t^B(x, K_t^B) = \frac{1}{1+\zeta}\left[K_t^B - \Upsilon_t(x)\right] \quad \text{and} \quad \Psi_t^S(x, K_t^S) = K_t^S - \Upsilon_t(x), \tag{B.16}$$

where $\Psi_t^B\left(x, K_t^B\right)$ is flatter than $\Psi_t^S\left(x, K_t^S\right)$, while K_t^B and K_t^S depend on the residential location of borrowers and savers. The rent formation endogenously determines whether an agent needs to borrow, possibly facing a borrowing limit. For any agent endowed with g_t the bid-rent function can be written as follows

$$\psi(x, g_t) = \begin{cases} \Psi_t^B\left(x, K_t^B\right) & \text{if } g_t < \Psi_t^B\left(x, K_t^B\right), \\ \Psi_t^S\left(x, K_t^S\right) & \text{if } \Psi_t^S\left(x, K_t^S\right) < g_t. \end{cases} \tag{B.17}$$

As in Section 3, land competition forces agents to partition across space. The capacity of any agent to pay to reside at location x depends on whether she needs to borrow or not. The wealthier the agent, the smaller the set of locations where she needs to borrow. If some agents are rich enough, they never need to borrow to live in the city and their bid rent is $\Psi_t^S\left(x, K_t^S\right)$. The corresponding lifetime wealth for a g_t-wealth agent is

$$W_t(x, g_t) = \begin{cases} \widetilde{w} + (1+\zeta)g_t - K_t^B + \mu y_{t+1} & \text{if borrower,} \\ \widetilde{w} + g_t - K_t^S + \mu y_{t+1} & \text{otherwise.} \end{cases} \tag{B.18}$$

Let us assume that all agents receive the same inheritance g_t^0 (and, thus, $\mu y_{t+1}^0 = \alpha g_t^0/\gamma$) that is insufficient to pay the house price in the CBD without borrowing (i.e., $g_t^0 < \Psi_t(0, K_t^B)$), while it is sufficient to pay the price in the periphery (i.e., $g_t^0 > \Psi_t(L_{0,t}^\star, K_t^S)$). Then the urban equilibrium features borrowers in the area $[0, \widehat{x}_t]$ and savers in the area $[\widehat{x}_t, L_{0,t}^\star]$, with $\widehat{x}_t$ and K_t^B such that $g_t^0 = \Psi_t^S\left(\widehat{x}, K_t^S\right) = \Psi_t^B\left(\widehat{x}, K_t^B\right)$ while K_t^S is determined by the condition $\Psi_t^S\left(L_{0,t}, K_t^S\right) = R_A$. The resulting equilibrium price function is:

$$p_t^\star(x) = \begin{cases} g_t^0 + \frac{1}{1+\zeta}[(R_A - g_t^0) + \Upsilon_t(L_{0,t}^\star) - \Upsilon_t(x)] & \text{for } 0 \le x \le \widehat{x}_t \\ R_A + \Upsilon_t(L_{0,t}) - \Upsilon_t(x) & \text{for } \widehat{x}_t \le x \le L_{0,t}^\star \\ R_A & \text{for } x \ge L_{0,t}. \end{cases} \tag{B.19}$$

where $\Upsilon_t(L_{0,t}^\star) = \widetilde{\kappa}_t(L_{0,t}^\star) - \mu^2 R_A$. By plugging the price paid by agents in the permanent wealth function, it is straightforward to show that the permanent wealth of both borrowers and savers is $W_t(x, g_t) = \widetilde{w} + (1 + \alpha/\gamma)g_t^0 - \mathcal{K}$ (where $\mathcal{K} = (1-\mu^2)R_A + \widetilde{\kappa}(L_{0,t}^\star)$). Therefore, the borrowing cost does not give rise to *symmetry breaking*. The reason is that, unlike the borrowing constraint, borrowing costs are capitalized by the bid rent. In this case, some agents locate in the best locations where they need to borrow because prices are high while the rest of the population does not borrow as they locate in the cheap city locations. Even though their status and location is different, neither their utility nor wealth inequality is impacted. Hence, we can characterize a steady state in which all agents' wealth converge to the same level

$$g_\infty = \frac{\gamma\Lambda(\widetilde{w} - \mathcal{K})}{1 - (\alpha+\gamma)\Lambda}.$$

The bid-rent functions in the long run are defined as:

$$\Psi^B_\infty(x, \mathcal{K}^B) = \frac{\mathcal{K}^B - \tilde{\kappa}(x)}{1 + \zeta - \mu^2} \quad \text{and} \quad \Psi^S_\infty(x, \mathcal{K}) = \frac{\mathcal{K} - \tilde{\kappa}(x)}{1 - \mu^2}. \tag{B.20}$$

We can define the distance $\widehat{x}_\infty$ dividing borrowers and savers in the city and the constant $\mathcal{K}^B$. As, $\Psi^S_\infty(\hat{x}_\infty, \mathcal{K}) = g_\infty = \Psi^B_\infty(\hat{x}_\infty, \mathcal{K}^B)$, we have

$$\tilde{\kappa}(\hat{x}_\infty) = \mathcal{K} - (1 - \mu^2)g_\infty \quad \text{and} \quad \mathcal{K}^B = \mathcal{K} + \zeta g_\infty.$$

In the long run all agents borrow if $g_\infty < R_A$. Conversely, all agents living in a city are savers if $g_\infty > \Psi^S_\infty(0, \mathcal{K})$ which is equivalent to $R_A < g_\infty + \tilde{\kappa}(L/2)/(1 - \mu^2) \equiv R^S_A$. When $R_A \in [g_\infty, R^S_A]$, agents are split into borrowers and savers even though they have the same wealth level. Agents do not borrow in locations far from the center where prices are low and pay $\Psi^S_\infty(\widehat{x}_\infty, \mathcal{K})$. Agents living close to the center in the area $[0, \widehat{x}_\infty]$ need to borrow as prices are high and pay $\Psi^B_\infty(\widehat{x}_\infty, \mathcal{K}^B)$.

It is also worth stressing that the slope of bid-rent functions with ζ is lower than the slope of shadow rent making the aggregate differential land rents lower than aggregate differential shadow rents. In this case, an additional tax is required to finance efficiently public expenditures.

B.6. Bubbly dynamics

Thus far, we have studied the rent dynamics by assuming away housing bubbles. If we do no more impose the transversality condition, a wider class of solutions may exist. We now explore rent dynamics with bubbles where house prices are now expressed as $\widetilde{p}_t(x) = p_t(x) + a_t$ with $p_t(x)$ being the fundamental solution of Equation (A.13) and a_t being a bubble. Under this configuration, we have

$$p_t(x) + a_t = \mu^2 \left[p_{t+1}(x) + \mathbb{E}_t(a_{t+1})\right] + \mathcal{K} - \tilde{\kappa}(x). \tag{B.21}$$

where $\mathbb{E}_t$ stands for the expectation operator. Any a_t such that $a_t = \mu^2\mathbb{E}_t(a_{t+1})$ is a solution of Equation (A.13). As $\mu^2 < 1$, a_t explodes in expected value. All agents expect the sale price to increase in the future. The bubble a_t does not modify the wealth dynamics of unconstrained agents because the rent internalizes the future sale price and thus the bubble. However, the bubble has some key consequences for the borrowing constraint. As it is recognized that rapidly rising house prices increase pressure to relax borrowing constraints (Acolin et al., 2016), we integrate into the downpayment requirement the expectation of the future increase in house prices as follows: $\widetilde{p}_t(x) - g_t < \lambda\widetilde{p}_{t+1}(x)$, which is equivalent to

$$p_t(x) - \lambda p_{t+1}(x) - (\lambda/\mu^2 - 1)a_t < g_t$$

with p_t given by (A.13) and where we have used (B.21) and $a_t = \mu^2\mathbb{E}_t(a_{t+1})$. The dynamics of $\widehat{x}^\star_t$ depend on the trend followed by the bubble a_t. The bubble generates two opposite effects on the borrowing constraint. On the one hand, expectations of an increase in the pledgeable sale price make the borrowing constraint less stringent. On the other hand, rising sale prices will increase the need to borrow and tighten the borrowing constraint. The overall effect depends on the credit market imperfections and on the type of bubble.

Assume that the bubble follows a deterministic increasing trend, given by $a_t = (\mu^2)^{-t} a_0$ with a_0 being an arbitrary initial condition. In this case, if credit market frictions are weak (resp., strong), *i.e.*, $\lambda/\mu^2 > 1$ (resp.,$<$) the borrowing constraint becomes less tight (resp., tighter). In the long run, the whole population ends up satisfying the borrowing constraint (resp., being borrowing-constrained) holding the same long-run wealth level, (resp., remaining unequal in the long run).

Assume now that the bubble follows stochastic dynamics given by

$$a_{t+1} = \begin{cases} \frac{1}{\mu^2 \pi} a_t + \xi_{t+1} \text{ with probability } \pi, \\ \xi_{t+1} \text{ with probability } 1 - \pi. \end{cases}$$

with $\mathbb{E}_t [\xi_{t+1}] = 0$. It amounts to assuming that if the bubble bursts with some probability $1 - \pi$ or continues to grow with probability π, the dynamics of the city would change. Consider $\lambda/\mu^2 > 1$, as in the deterministic case: As long as the bubble inflates, the borrowing constraint would become less stringent, and credit would be easier. After a bursting of the bubble, the borrowing constraint can suddenly be binding for a part of the population, thereby generating wealth inequality.

References

Acolin, Arthur, Jesse Bricker, Paul Calem, and Susan Wachter. 2016. "Borrowing Constraints and Homeownership." *American Economic Review*, 106(5): 625–629.

Behrens, Kristian, and FrÃ©dÃ©ric Robert-Nicoud. 2015. "Agglomeration Theory with Heterogeneous Agents." In *Handbook of Regional and Urban Economics*. Vol. 5, 171–245. Elsevier.

Chiappori, Pierre-AndrÃ©. 2017. *Matching with Transfers : The Economics of Love and Marriage.* Princeton University Press.

Michel, Philippe, and Pierre Pestieau. 2004. "Fiscal Policy in an Overlapping Generations Model with Bequest-as-Consumption." *Journal of Public Economic Theory*, 6(3): 397–407.