The spatial dynamics of population: an agent-based approach

Preliminary version

Davide Fiaschi — University of Pisa

Cristiano Ricci — University of Pisa

February 28, 2023

For the full version of this paper with pictures see the link:

`https://poisson.phc.dm.unipi.it/~cricci/fiaschi_ricci.pdf`

Abstract

This paper considers an economy where agents mobility is driven by differential utilities over space, production and consumption show spatial spillovers, and there exist exogenous and endogenous amenities. The spatial dynamics of population is derived as the mean-field limit of a system of interacting agents as the number of agents becomes infinite. Numerical experiments show that the model can reproduce several stylized effects, as the emergence of cities with different size and shape; the importance of history, with small perturbations in the initial population distribution leading to substantial differences in the long-run dynamics; the phenomenon of metastability, where a long period of stability in the spatial distribution is followed by a sharp transition to a new (meta) stable equilibrium; and, finally, a non-linear out-of-equilibrium dynamics, with regions with a first phase of increasing, followed by a phase of decreasing, population.

JEL Classification Numbers: C23, R12, R15

Keywords: spatial agglomeration, spatial spillovers, Aggregation-Diffusion Equations, metastability, complex landscapes

Contents

1 Introduction

This paper proposed a microfoundation of the spatial dynamics of population based on the idea that individual mobility is driven by differential utilities over space, there exists spatial spillovers in production and consumption, and the space is not uniform but characterized by exogenous and endogenous amenities, i.e. potential markets/economic potentials, infrastructures, facilities mainly determined by population density.

The model is defined in the continuous space, that is the full plane where points are identified by longitude and latitude, and in continuous time. The partial differential equation driving the spatial dynamics of population is derived as the mean-field limit of a system of discrete interacting agents as the number of agents becomes large, belongs to the class of *Aggregation-Diffusion Equations* (ADEs), which over the past 20 years have been employed in several biological applications and stimulated many mathematical works (see Carrillo et al., 2019 for a review). The competing effect of aggregation and diffusion leads to several interesting properties, such as metastability, symmetrization, or, in certain cases, non-uniqueness in the equilibrium solutions (Carrillo et al., 2019).

We conduct numerical experiments for investigating the properties of the model. In particular, we focus on: i) the emergence of cities with different size and shape, i.e. the Zipf's law; ii) the importance of history, i.e. when small perturbations in the initial distribution result in strong difference in the long.run dynamics; iii) the phenomenon of metastability, i.e. long periods of stability in the spatial distribution, followed by sharp a transition to a new (meta) stable equilibrium; and, finally, iv) the nonlinear out-of-equilibrium dynamics, i.e. the spatial income distribution show regions with a first phase of increasing, followed by a phase of decreasing, population.

Krugman (1994) represents one of the pioneering contribution to the study of the observed complex landscapes in the spatial distribution of population and economic activities. Durlauf (1997) contains several key contributions in this respect, among which the one by Steven Durlauf on the social spatial interactions and by Paul Krugman on the self-organization of economies in the space. Durlauf (1994) investigates the importance of local spillovers for the emergence of spatial inequality, while Durlauf et al. (2005) reviews the efforts to introduce complex systems methods into economics and the understanding of empirical phenomena.

Another important strand of literature is related to cities formation (spatial agglomeration), which points out to the spatial externalities in production, in particular on the labor productivity (Glaeser and Mare, 2001) and on the spatial externalities deriving by urban density which " spreads knowledge, which either makes workers more skilled or entrepreneurs more productive" (Glaeser and Resseger, 2010). Davis and Weinstein (2002) test three alternative theories on the determinants of density of economic activity, while Duranton and Puga (2020) discuss benefits and costs of urban density. A general equilibrium spatial model with no dynamics, but very rich for the presence of spatial externalities, trade and labor mobility is presented by Allen and Arkolakis (2014) Arbia (2001) models the geography of economic activities on a continuous space.

Sznitman (1991) is the pioneering contribution on the microfoundations of aggregation-diffusion model, while Morale et al. (2005) contains a microfoundation of the model of aggregation. Anton Bovier (2015) deeply discuss the concept of metastability and its application. Carrillo et al. (2019) contains a review of aggregation-diffusion models. Finally, Mocenni et al. (2010) present a model in spirit very similar to ours, which displays Touring patterns over space.

The contributions of the paper are several. First, it provides a sound microfoundation to some observed geographical patterns of interest, which include the agglomeration of population over time, i.e. the emergence of cities, and their distribution over space (Zipf's law); the evolution over space and time of wage inequality, i.e. the geographical pattern of wage inequality. Second, it individuates as a possible source of spatial agglomeration the non-market interactions derived by spatial externalities in production, i.e. stock of local knowledge, local good and labour markets, in a framework where agents can freely move across locations. It also points out to potential sources of spatial agglomeration related to the evolution of endogenous amenities, i.e. local specific factors changing over time as responding to the evolution of local population density, i.e. infrastructures, schools, universities, facilities. Finally, it introduces in the analysis non-homogeneous spaces for the presence of mountains, rivers, coasts, roads, etc., which deeply affect the observed spatial distribution of population. The analysis also highlights the importance of considering the presence of metastable equilibria in the spatial general equilibrium literature, travelling wave equilibria (KPP) and out-of-equilibrium dynamics as opposed to the search for a balanced path equilibrium and/or steady state equilibrium.

The paper is organized as follows. Section 2 reports the spatial pattern of population in some European countries; Section 3 includes the description of the model; Section 4 contains the numerical explorations of the model; and Section 5 concludes. Appendix gathers technical details and proofs of analysis.

2 Spatial pattern of population in European countries

To make explicit the aim of our analysis consider the evidence reported in Figure 1, which reports the spatial distribution of the intensity of nightlights for three of the main European countries.[1]

Figures/mapFR-eps-converted-to.pdf

Figures/mapDE-eps-converted-to.pdf

Figures/mapIT-eps-converted-to.p

Figure 1: The maps of nightlights in 2021 for France, Germany and Italy (from left to the right) based on cells of about 500m x 500m. Source: VIIRS 2.1 database.

Taken the intensity of nightlights as proxy for the local population (Michalopoulos and Papaioannou, 2018), there exists a clear pattern of spatial aggregation of population, which reflects the localization of the main cities of countries. However, the patterns are very heterogeneous across the three countries: France, with a substantial homogeneity of its territory, presents an extreme degree of spatial aggregation around Paris and only few other clusters. On the contrary, Germany, also with a territory not presenting particular geographical barriers, displays a less concentrated distribution, with the presence of several agglomerations of medium size uniformly distributed over the country. Finally, Italy, with a very heterogeneous territory, show a concentration of population in the costasl areas and in the Pianura Padana. In the latter, we observe that some cities of medium size are located along the ancient Roman *Via*

[1] Nightlights are taken from VIIRS 2.1 database (`https://eogdata.mines.edu/products/vnl/`), which provides the average intensity of nightlight for cells of 500 x 500 meters.

Emilia, a strait road that, starting from Rimini, runs along the Appennini mountains to arrive at Piacenza. Several cities along this road were founded by Romans more than two millenniums ago (e.g. Rimini, Boulogne, Modena, Reggio Emilia, Parma, Piacenza). Several medium Italian cities have experimented in the last 30 years a strong reduction in their population size in favour of some bigger cities also in presence of increasing total population.

3 The model

This section is organized as follows. Section 3.1 illustrates the microfoundations of agents' movement driven by spatial differential utilities; Section 3.2 explains how to allow for an increases in the number of agents over time; Section 3.3 specifies as the local wage is determined; Section provides a detailed explanation of how local exogenous and endogenous amenities are modelled; finally, Section 3.5 reports the main theoretical result of the paper, i.e. the proof that the spatial dynamics of population can be expressed by a partial differential equation belonging to the class of Aggregation-Diffusion Equations.

3.1 The microfoundations of agents' movements

We assume that space and time are continuous variable, with $z \in \Omega \subseteq \mathbb{R}^2$, and Ω is assumed to be open and with boundary of class C^1. Moreover, there exists also an initial non trivial distribution of *technology*, *labour* and of *other exogenous variables across different locations* (e.g. physical capital). At aggregate level no stochastic behaviour is present but microfoundation is crucially based on the presence of a stochastic component.

We focus on the case where the *place of residence* is also the *place of work* of agents, i.e. set aside the issue of *commuting* and agents are participating in local labour markets. We also neglect the accumulation and the spatial reallocation of physical capital, which would require to consider local good markets and trade costs.

Assume that at time $t = 0$ the number of agent is N and that the *aggregate stock of labour* in the economy is equal to one, i.e. each agent has an endowment of labour of $1/N$. Let $X_t^{i,N}$ be the agent i's position in the domain Ω at each time t; agent i is moving over space driven by its local utility, i.e. its movements can be described by:

$$dX_t^{i,N} = \gamma \nabla_x u^i \left(\mathbf{X}_t^N, t\right) \left(X_t^{i,N}\right) dt, \tag{1}$$

where $\mathbf{X}_t^N$ is the vector of positions of all agents, $\gamma > 0$ is a parameter measuring the inverse moving costs agent i faces in moving toward positions where its utility $u_i \left(\mathbf{X}_t^N, t\right)$ is higher.

Agent i's utility is a function of its position, $X_t^{i,N}$, of the positions of all N agents $\mathbf{X}_t^N$, and of time t, i.e. agents' utility can be affected by time varying exogenous factors, as, e.g., technological progress. $\nabla_x u^i$ is the *gradient operator*, i.e., being in a two-dimensional space, $\nabla_x u^i \equiv (\partial_{x_1} u_i, \partial_{x_2} u^i)$ (x_1 and x_2 are the two directions). It indicates the direction of movement which agent i should follow for increasing its utility, given its position $X_t^{i,N}$ and the position of all N agents $\mathbf{X}_t^N$, if in the neighbouring locations of $X_t^{i,N}$ agent i could get a higher utility.

Figure 2 reports an example of the dynamics over space of agent i, when the different circles represent some iso-utility loci (indifference curves) given the spatial distribution of agents$\mathbf{X}_t^N$, the black arrow the gradient $\nabla_x u^i \left(\mathbf{X}_t^N\right) \left(t, X_t^{i,N}\right)$ (according to the direction of gradient in $X_t^{i,N}$ the agent's utility in $X_t^{i,N}$ is slightly above 1 and below 2), and the blue arrow the agent i's movement in period $[t, t+dt]$, i.e. $\gamma \nabla_x u^i \left(\mathbf{X}_t^N\right) \left(t, X_t^{i,N}\right)$. In the panel on the right of Figure 2 the trajectory of agent i is calculated assuming that all the other agents are not moving, i.e. the field of utility is not evolving over time.

Figures/movementAgent-eps-converted-to.pdf

Figures/movementAgent_II-eps-converted-to.pdf

Figure 2: The dynamics of space (x_1,x_2) of the position of agent i according to Eq. (1) with reported also the position of other agents in the plane.

Eq. (1) represents the minimal setup for agents which are pursuing a higher utility moving over space but subject to movement costs (frictions) and are not coordinating their movement choices taking $\mathbf{X}_t^N$ as given. Howitt (2006), LeBaron and Tesfatsion (2008), Colander et al. (2008) and Aoki and Yoshikawa (2011) provide a deep discussion of the implications of this approach in macroeconomics, which is labelled *agent-based model* and it is a complement/alternative approach to the framework with optimizing forward-looking (in the sense of space and time) agents, as in Allen and Arkolakis (2014) and Rossi-Hansberg (2019).

Panel on the right of Figure 2 highlights the *local perspective* of our agents, in the sense that they are choosing the local optimal direction and not the most efficient one to reach the bliss point in the field of utility. In Appendix A we show that this can be approximatively optimal when the field of utility depends on the agents' position, which is evolving over time and the time that elapses between two consecutive decisions is approximately zero; while in Appendix B we discuss as Eq. 1 can be also view as a Nash equilibrium of agents maximizing step by step in discrete time. In both case γ is the inverse of the costs of movement.

According to Eq. (1) at time t agent i is not moving when its utility in the neighbour of its position $X^{i,N}$ is equal to its actual position, i.e. $\nabla_x u^i \left(\mathbf{X}_t^N\right) \left(t, X_t^{i,N}\right) = 0$. The equilibrium is therefore defined as the condition when $\nabla_x u^i \left(\mathbf{X}_t^N\right) \left(t, X_t^{i,N}\right) = 0 \ \forall i$, which admits the possibility of i) *Pareto dominated equilibria* because of the lack of agents' coordination (Howitt, 2006) and ii) equilibria characterized by *complex geographical landscape* (Krugman, 1994). Also important, this framework can produce *non-linear out-of-equilibrium dynamics* as shown in Krugman (1994). All these features find empirical support and represent one of the most important value added of this approach. We will show below that another advantage is the flexibility in taking into account in the analysis of some key features of the geography of population, as the presence of positive spatial spillovers in production (spatial increasing returns) and consumption (endogenous amenities), together with (exogenous amenities) and a non-uniform space (natural barriers).

3.1.1 Local utility function

In an economy without accumulation of physical capital, i.e. no savings, the level of consumption is by definition equal to the level of agents' income; hence:

$$c\left(\mathbf{X}_t^N, t\right)\left(X_t^{i,N}\right) = \frac{w^N\left(\mathbf{X}_t^N, t\right)\left(X_t^{i,N}\right)}{P\left(\mathbf{X}_t^N, t\right)\left(X_t^{i,N}\right)} = w^R\left(\mathbf{X}_t^N, t\right)\left(X_t^{i,N}\right), \tag{2}$$

where $c\left(\mathbf{X}_t^N, t\right)\left(X_t^{i,N}\right)$ is the agent i's consumption, $w\left(\mathbf{X}_t^N, t\right)\left(X_t^{i,N}\right)$ its nominal wage and $P\left(\mathbf{X}_t^N, t\right)\left(X_t^{i,N}\right)$ the level of prices of the position of residence of agent i. All variables are defined with respect the whole distribution of agents in the space, i.e. $\mathbf{X}_t^N$, and evaluated in the position of agent i, $X_t^{i,N}$, at time t. In our simplified setting both nominal wages and prices depends on the local conditions of production, which will be specified below. In an economy where place of residence and place of work are different agent's wage would depend on the place of work, while price on the place of residence (typically the price of land/houses or the level of rents). Moreover, agents' utility also depends on the amenities specific to the position of residence; in particular, these amenities can be exogenous, such as weather conditions, or endogenous, such as the presence of facilities, schools, hospitals, pollution, traffic congestion etc. , which are dependent of local production and population density.

The utility follows the Random Utility Model framework (Train (2009)), i.e.

$$u^i\left(\mathbf{X}_t^N, t\right)(x) := w^N\left(\mathbf{X}_t^N, t\right)(x) + A_{EN}\left(\mathbf{X}_t^N, t\right)(x) + A_{ES}(t, x) + \sigma x \cdot \frac{dB_t^i}{dt}, \tag{3}$$

where utility is linear in consumption, includes two additional terms for endogenous and exogenous amenities, $A_{EN}\left(\mathbf{X}_t^N, t\right)\left(X_t^{i,N}\right)$ and $A_{ES}\left(t, X_t^{i,N}\right)$ respectively. The random component $\sigma x \cdot dB_t^i/dt$ reflects the *idiosyncratic preferences* of agent i for the position x. It is the scalar product between the position x and dB_t^i/dt, which represents the instantaneous variation of a stochastic process with Gaussian independent increments over disjoint time intervals (Brownian motion). The position x is inserted as a multiplicative factor to make the random component of the utility have a variance that depends on the distance between two locations. The parameter σ is a scale factor for the standard deviation of the random component.

From Eqq. (1) and (3) we get:

$$dX_t^{i,N} = \gamma \left[\nabla_x w^R \left(\mathbf{X}_t^N, t \right) \left(X_t^{i,N} \right) + \nabla_x A_{EN} \left(\mathbf{X}_t^N, t \right) \left(X_t^{i,N} \right) + \nabla_x A_{ES} \left(t, X_t^{i,N} \right) \right] dt + \gamma \sigma dB_t^i, \tag{4}$$

where $(B_t^i)_{i=1,\ldots,N}$ are independent Brownian motions. Eq. (4) shows that agent i is moving toward nearby positions where wages and amenities are higher, plus a random component with not systematic preference for specific positions.

3.2 The increase in the number of agents

Let N_t^N be the number of agents that are present at time t, with $N_0^N = N$, that is at time $t = 0$ N agents are present. We allow that the aggregate stock of labour is increasing over time with the total number of agents N_t^N, i.e. at period t the aggregate stock of labour is $N_t^N/N \geq 1$. Assume that in each position there exists a *birth rate* $n(t, x) \geq 0$ specific of position x at time t; a possible death rate has a similar characteristic and it is omitted in the analysis. New agents are introduced by means of a Poisson process with a *time-space non-homogeneous* birth rate $n\left(t, X_t^{i,N}\right)$ for agent i at period t, where the first jump of the process corresponds to the duplication of agent i. The offspring is created in the same position of its "mother", but with an independent noise.

Figures/movementAgent_III-eps-converted-to.pdf

Figure 3: The path of duplication for agent i and its offspring in the space Ω.

Figure 3 reports an example of two duplications for the offspring of agent i. The time span between two duplications are denoted by τ^i for the first duplication of agent i, by $\tau^{i,2}$ for the second duplication of agent i, and so on. In Figure 3 three new agents are generated in a period of length $\max\{\tau^i + \tau^{i,2}, \tau^i + \tau^{N+1}\}$. The length in space that separates two consecutive duplications cannot be easily mapped into the time that between the two duplications, depending on the speed of motion of agent i and the local birth rate of the positions visited by agent i. For the sake of simplicity we don't consider the presence of an exogenous death rate, which however can be treated within the same framework; hence, in presence of a local mortality rate, $n(t,x)$ should be meant as the *local population growth rate*.

3.3 The determination of local wage

Production at position x at period t in an economy with N agents, $y^N(t,x)$, is defined by:

$$y^N(t,x) = A_l^N(t,x) l^N(t,x)^\beta \tag{5}$$

with $\beta \in (0,1)$, where $A_l^N(t,z)$ is a *local-specific technological progress* affecting the *marginal productivity* of labour, and $l^N(t,x)$ is the *stock of labour* available at position x at period t. Therefore, wage is given by:

$$w^N(t,x) \;\; = \;\; A_l^N(t,x) l^N(t,x)^{\beta-1}. \tag{6}$$

A technical difficulty arising in our framework, where space is continuous and the number of agents is finite, is that the participation of each agent to local labour markets, i.e. the stock of labour available at position x at period t, is modelled as the sum of differential contributions of all the agents. In particular, the individual contributions are decided by the (geographical) distance of agents from the locations, and they are conveniently modelled by the *kernel function* $\theta_N(\cdot)$. Hence, the stock of labour at position x at time t is defined by:

$$l^N(t,x) = \sum_{i=1}^{N_t^N} \frac{1}{N} \cdot \theta_N(x - X_t^{i,N}), \tag{7}$$

where:

$$\begin{aligned}
\theta_N(z) &:= \left(\frac{1}{h_N}\right)^2 \cdot \theta\left(\frac{z}{h_N}\right), \ \forall z \in \Omega; \\
h_N &:= N^{-\lambda}; \text{ and} \\
\lambda &\in (0, \overline{\lambda}) \text{ for some } \overline{\lambda} < 1/4.
\end{aligned}$$

Function θ is assumed to be of compact support with radius one, integrates to one, symmetric and non increasing with respect to the distance from the origin. The *bandwidth* h_N is reflecting the influence of the distance on the participation of agents to the labour market at position x, i.e. a higher h_N means that each agent is participating in the labour market of farther locations from its position. Since the kernel function θ_N must integrate to one, Equation (7) is a weighted sum of individual endowment of labour of each agent, where weights are decided by the distance of each agent from the location x. The assumption $\lambda \in (0, \overline{\lambda})$ for some $\overline{\lambda} < 1/4$ is crucial to derive the aggregate dynamics as a Partial Differential Equation (PDE), as discussed in Appendix D. Therefore, if we consider an economy with a larger number of agents, the labour market in each position x is on average participated by agents that are closer to location x.

As regarding the local specific technological progress $A_l^N(t,x)$ in presence of spatial spillovers, following Glaeser and Resseger (2010); Glaeser and Mare (2001) we assume that:

$$A_l^N(t,x) = G(x) \sum_{i=1}^{N_t^N} \frac{1}{N} W_h^P(x - X_t^{i,N}), \tag{8}$$

where $W_h^P(z) := (1/h^2)\, W^P(z/h)$ is a kernel function assumed to be of compact support with radius (bandwith) $h > 0$, integrates to one, symmetric with respect to the origin, and non increasing with respect to the distance from the origin, and $G(x)$ is a non-negative function defined on the space Ω.

$A_l(t,x)$ is the product of the two terms; the first, $G(x)$, represents the *potential use* of location x for production. This potential can be traced out to the possibility to use that specific location for productive purposes. For example, $G(x)$ should be at a minimum value of zero in those positions where a stable production is not possible (or extremely difficult), such as in the centre of rivers, in the sea, etc.. The second term, $\sum_{i=1}^{N_t^N} \frac{1}{N} W_h^P(x - X_t^{i,N})$, is a measure of local density of population, where each agents is considered for its endowment of labour

$1/N$, and its neighbourless to location x is modelled by the kernel function W_h^P. Bandwidth h reflects the extension of spatial spillovers, i.e. higher h means wider spatial spillovers. At the individual level $\sum_{i=1}^{N_t^N} \frac{1}{N} W_h^P(x - X_t^{i,N})$ is a measure of the *positive spatial externalities* in the production of having neighbouring locations with large amount of agents, i.e. agents density positively affects labour productivity (Ciccone and Hall, 1996), therefore

$$y^N(t,x) = G(x) \left[\sum_{i=1}^{N_t^N} \frac{1}{N} W_h^P(x - X_t^{i,N}) \right] l^N(t,x)^{\beta}, \tag{9}$$

$$w^N(t,x) = G(x) \left[\sum_{i=1}^{N_t^N} \frac{1}{N} W_h^P(x - X_t^{i,N}) \right] l^N(t,x)^{\beta-1}. \tag{10}$$

3.4 The determination of local amenities

Local amenities can be both exogenous and endogenous. The exogenous amenities are related to specific characteristics of location x independently of the distribution of agents over space. Typical examples are climate, weather, natural landscape, etc.. All these characteristics can be considered exogenous to the overall population distribution dynamics, and can be modelled as a function $A_{ES}(t,x)$. The shape of A_{ES} reflects differentials over space of individual utility caused by spacial specific characteristics of each location. The extreme case is sea, rivers, and/or high mountains, where function A_{ES} should signal an extremely low utility for people resident in those places.

The endogenous amenities need a more sophisticated modelling, because they are affected by the local population and income. Typical examples include local services, e.g. schools, hospitals, and theatres, local goods such as malls, and any other infrastructures and facilities which are *locally supplied and consumed*. Endogenous amenities are also subject to *congestion*, i.e., given their stock, the individual utility decreases with the number of individuals which are access to them.

To fix the idea suppose that individuals devotes a constant share $\tau \in (0,1)$ of their income to finance the consumption of endogenous amenities; the total amount of expenditure is given by $\tau \cdot y^N(t,x)$. The production of amenities is subject to decreasing marginal productivity;

hence the total amount of supplied amenities at location x at time t is:

$$PA(t,x) = \left[\tau \cdot y^N(t,x)\right]^{\varphi}, \tag{11}$$

where $\varphi \in (0,1)$. An alternative/complementary explanation for $PA(t,x)$ is that local endogenous amenities are financed by a flat tax rate $\tau \in (0,1)$ on income; in this case, $\tau \cdot y^N(t,x)$ would represent the collected revenues to finance the provision of amenities.

Congestion is assumed to be proportional to the number of individuals; hence, the individual utility for the endogenous amenities is assumed to be given by:

$$A_{EN}(t,x) = A_0 \left\{ \left[\tau \cdot y^N(t,x)\right]^{\varphi} - \mu_A l^N(t,x) \right\} = A_0 \left\{ \left[\tau \cdot A_l^N(t,x) l^N(t,x)^{\beta}\right]^{\varphi} - \mu_A l^N(t,x) \right\}, \tag{12}$$

where A_0 is a scale parameter and $\mu_A > 0$ a parameter measuring the intensity of congestion.

Equation (12) shows that endogenous amenities are a source of agglomeration until the stock of labour is under $l^N(t,x)^{TR} = \left[\varphi\beta\tau^{\varphi}\left(A_l^N(t,x)\right)^{\varphi}/\gamma_A\right]^{1/(1-\varphi\beta)}$; after this threshold endogenous amenities works against agglomeration. $l^N(t,x)^{TR}$ is increasing in $A_l^N(t,x)$, τ and β and decreasing in γ_A as expected.

3.5 The spatial dynamics of population in an economy with infinite agents

The main result of the paper is represented by Theorem 3.1, which states that from the movements of agents we can represent the dynamics at aggregate level by a partial differential equation taking the limit to infinity in the number of agents N, as discussed, among others, by Morale et al. (2005) and Flandoli (2016).

Teorema 3.1. *Assume that Eq.* (4) *descrives the agent' movement, Eq.* (7) *the stock of labour in each location, Eq.* (8) *the local technological progress, Eq.* (9) *the local income, Eq.* (10) *the local wage, Eq. 12 the endogenous amenities, and* $n(t,x)$ *the net population rate of the position* x *at time* t*, then the stock of population in position* x *at time* t*,* $l^N(t,x)$ *converges in probability* *as N tends to infinity, in the suitable topology, to the unique solution of*

$$
\begin{aligned}
\partial_t l(t,x) &= n(t,x)l(t,x) + \frac{\gamma^2\sigma^2}{2}\Delta_x l\,(t,x) + \\
&- \gamma\, div_x \left(l(t,x)\nabla_x w\,(t,x)\right) \\
&- \gamma\, div_x \left(l(t,x)\nabla_x A_{EN}(t,x)\right) + \\
&- \gamma\, div_x \left(l(t,x)\nabla_x A_{ES}(t,x)\right), \qquad (13)
\end{aligned}
$$

where

$$
\begin{aligned}
A_l(t,x) &= G(x)\left(W_h^P * l\right)(t,x); \\
y(t,x) &= G(x)(W_h^P * l)(t,x)l(t,x)^{\beta}; \\
w(t,x) &= G(x)(W_h^P * l)(t,x)l(t,x)^{\beta-1};\ \textit{and} \\
A_{EN}(t,x) &= [\tau A_l(t,x)y(t,x)]^{\varphi} - \mu_A l(t,x)
\end{aligned}
$$

are the value in the limit of local technological progress, income, wages and endogenous amenities respectively, where $W_h^P(z) := (1/h^2)\, W^P\,(z/h)$ *with* $h > 0$ *and* $W^P \in C_b^1\,(\Omega)$.

Proof. For a sketch of the proof see Appendix D. □

The first term of Eq. 13 is due to the materialization of new agents at location x. The second term represents the usual diffusion component due to randomic component of agents' movement. A specific mention instead deserves the third term of Eq. 13, which represents the effect of *labor reallocation* across positions, assuming that labour is moving across positions according to differentials in the level of wages $w(t,z)$, γ are two coefficients measuring the speed of reallocation, div_x is the *divergence operator*, i.e. $\text{div}_x f \equiv \partial_{x_1} f_{x_1} + \partial_{x_2} f_{x_2}$, while $\nabla_x f$ is the *gradient operator*, i.e. $\nabla_x f \equiv (\partial_{x_1} f, \partial_{x_2} f)$. Eq. (13) is inspired to Xepapadeas and Yannacopoulos (2016) and contrasts with the pure diffusion proposed by Boucekkine et al. (2009). The fourth and fifth terms Eq. 13 represents the effect of *labor reallocation* across positions due to differential endogenous and exogenous amenities respectively.

4 Numerical explorations

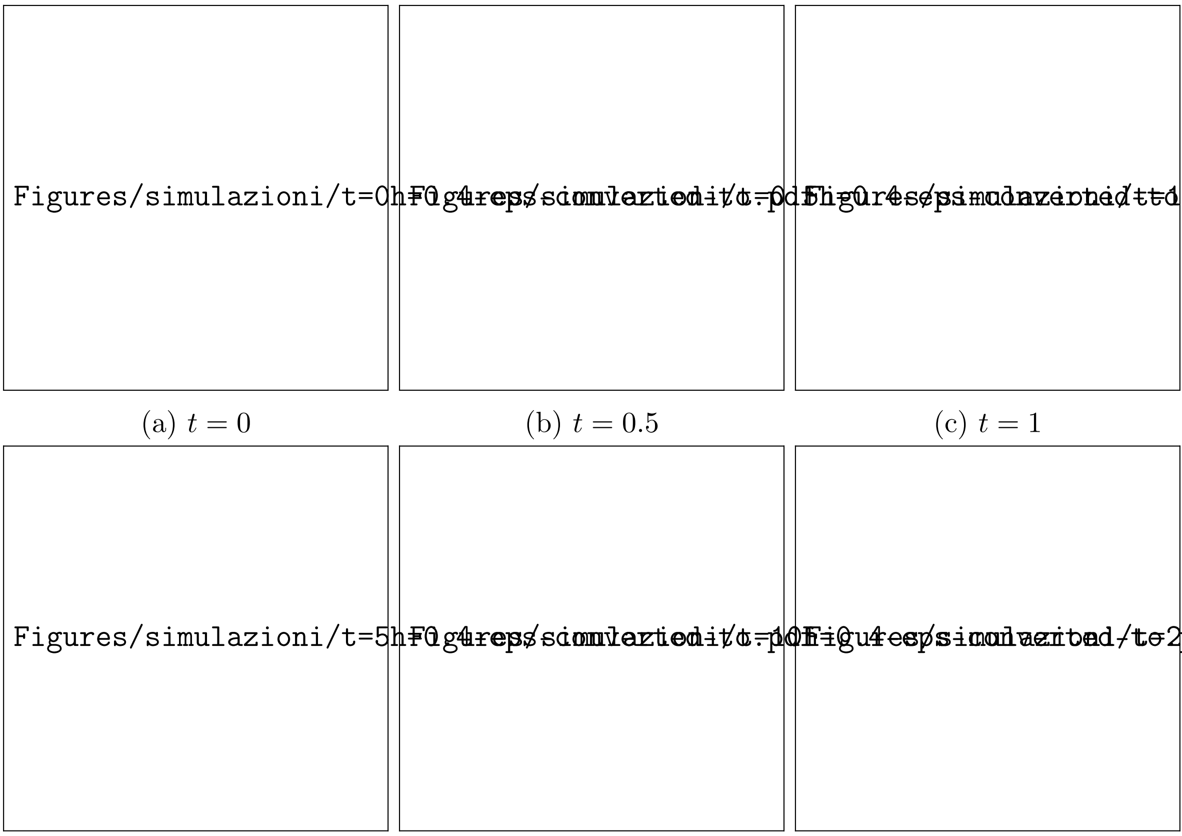
The baseline specification of our numerical explorations of Eq. (13) is in Table 1, while Figure 4 reports the dynamics of spatial distribution of population for different periods, from the initial period (t=0) and for five next periods (t=0.5,...,20).

The initial spatial distribution of population is uniform over the space $\Omega = [0,4] \times [0,4]$. At period 1 four agglomerations (cities) emerge; such a pattern becomes more distinct at period 5 with four distinct cities; however at period 10, they tend to group together, and, finally, at period 20 they merge in one big city, which will remain also for the next periods (not reported in the figure). In a large part of the plane is not present population at period 20, included in regions that experimented in the first periods an increase in population, providing an example of extreme local nonlinear dynamics. Not reported numerical simulations show that the localization of the big city can depend on a small local perturbation in the initial distribution of population, i.e. history matters for the location of big city.

Parameters	Value	Description
$n(t,x)$	0	Local population growth rate
$G(x)$	1	Uniform potential use for production of each location
$A_{ES}(t,x)$	0	No difference in endogenous amenities across different locations
σ	0.05	Variance of random component in the agents' movement
β	0.6	Coefficient of the local production function
γ	0.01	Inverse of movement costs
h	0.4	Extention of spatial spillovers
A_0	6	Scale parameter for the endogenous amenities
τ	0.2	Flat tax rate on local income to finance local amenities
φ	0.5	Coefficient of the local production function of amenities
μ_A	0.2	Measure of congestion in local amenities
$W^P(x)$	$(1-\|\|x\|\|)\,\mathbb{1}_{\|\|x\|\|\leq 1}$	Kernel function used in the simulation

Table 1: The baseline values of model's parameter used in the numerical exploration of Aggregate-Diffusion partial differential equation (13)

Figure 5 instead reports the dynamics of spatial distribution of population when the extension of spatial spillover h is reduced from 4 to 3. While the final equilibrium at period 185 is always a big city, there exists a wide interval from $t = 5$ to $t = 180$ where the spatial pattern with four medium cities appeared as stable. However, in just two periods (from 181 to 182), the spatial pattern radically changes its aspect with the emergence of a big city doomed to persist also in the next periods (not reported in the figure). This phenomenon is called metastability



(a) $t = 0$ (b) $t = 0.5$ (c) $t = 1$

(d) $t = 5$ (e) $t = 10$ (f) $t = 20$

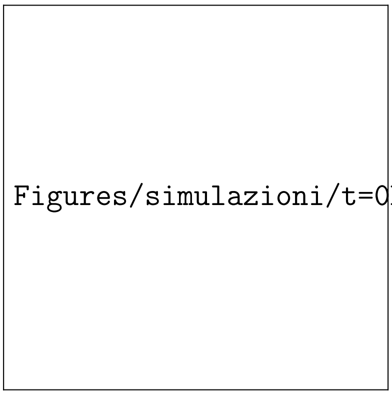
Figure 4: The distribution dynamics of population $l(t, z)$ over space and time for the baseline case with, $h = 0.4$, $\Omega = [0, 4] \times [0, 4]$.

and poses serious doubts on the possibility to use standard econometric techniques for the study of the dynamics of city size distribution.

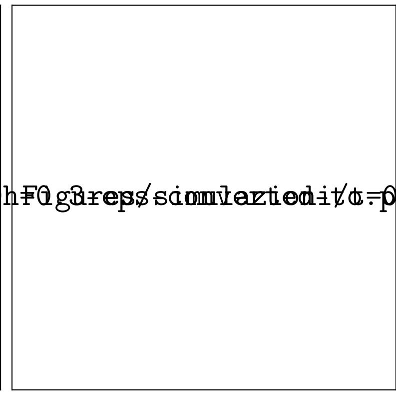
5 Conclusions

We have shown as our model can generate a dynamics of the spatial distribution of population which reproduce some stylized facts regarding the emergence of cities with different size, shape and spatial organization; the importance of history for the location of cities; the phenomenon of metastability, i.e. cities emerging as the outcome of a long and apparently stable process, can abruptedly disappear; and, finally, the nonlinear out-of-equilibrium dynamics of population, i.e. cities which an increasing population in the past can experiment a continuos drop in the future. Our findings cast several doubts on the actual techniques used in the empirical analysis.

The proposed model can be extended along three main lines of research. The first is focused on the commuting, i.e. the possibility that agents are living in a place different from their place of work. This extension is particularly important for the study of the inner structure of cities, with the related phenomena of segregation and unequal spatial distribution of wages. Another line of research is to to study the spatial pattern by the concept of recurrence (Mocenni et al., 2010; Facchini and Mocenni, 2013). Finally, to study the economy in the small sample case, i.e. not in the limit of $N \to \infty$ (Evers and Kolokolnikov, 2016).

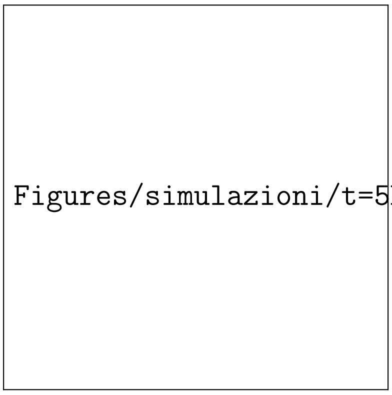


(a) $t = 0$

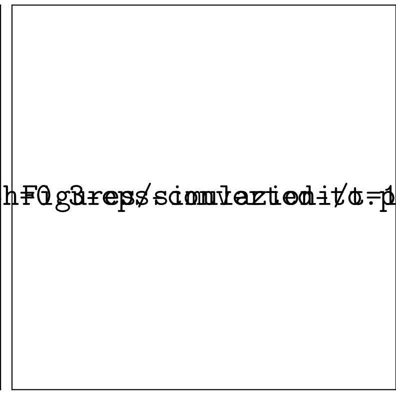


(b) $t = 0.5$

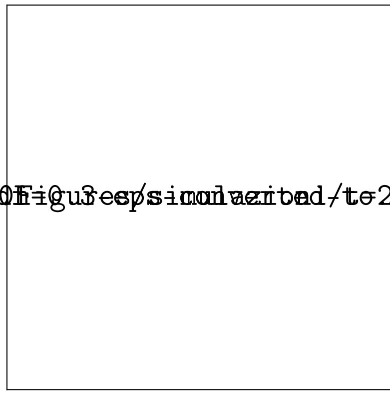
(c) $t = 1$



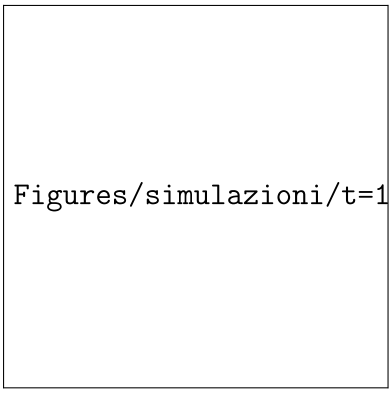
(d) $t = 5$



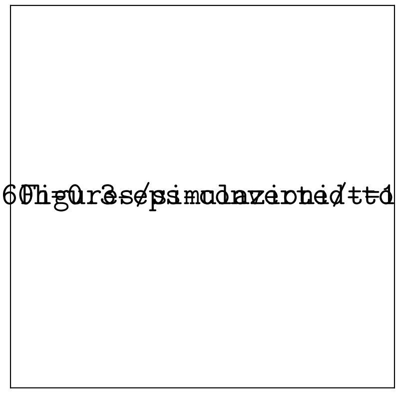
(e) $t = 10$



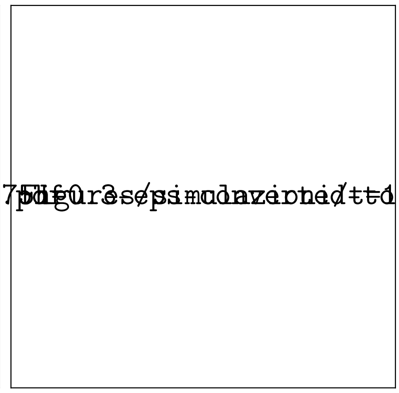
(f) $t = 20$



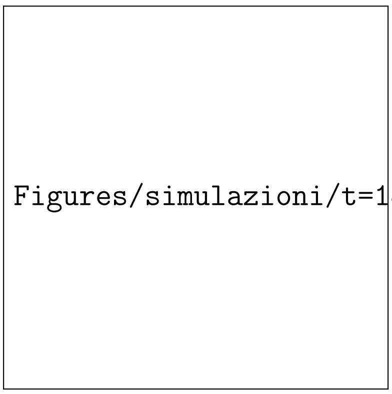
(g) $t = 160$



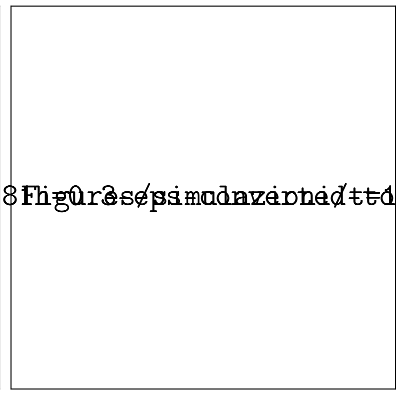
(h) $t = 175$



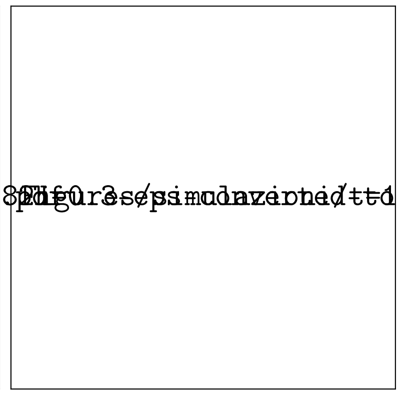
(i) $t = 180$



(j) $t = 181$



(k) $t = 182$



(l) $t = 185$

Figure 5: The distribution dynamics of $l(t, z)$ over space and time for the case $h = 0.3$, $\Omega = [0, 4] \times [0, 4]$.

Appendix

A The microfoundations of Eq. (1) with myopic agents

In this appendix we derive Eq. (1) in a static optimizing framework in continuos time. Assume agent i is optimizing its instantenous utility. At each time t denote the spatial position of agent i by $X_t^{i,N}$, and the position of all other agents (including himself) $\mathbf{X}_t$. For the sake of simplicity, below we assume that the space dimension is one, i.e. $X_t^{i,N} \in \mathbb{R}$. All the arguments made below can be repeated in the same manner, but with a more cumbersone notation, in any dimension $d \geq 1$.

Consider the utility function of agent i that depends on his position in space, and also on the position of all other agents

$$u^i \left(\mathbf{X}_t^N, t\right)(X_t^{i,N}), \tag{14}$$

where the first component of u represents the overall shape of utility determined by the spatial distribution of agents, the second is the overall change of utility due to possible time varying factors (e.g. technological progress, climate change, etc.), while the third component assigns to agent i the utility level corresponding to its position $X_t^{i,N}$.

We now introduce the notation for the partial derivativative of the function u^i. We denote

$$\frac{\partial u^i}{\partial \mathbf{X}_k}(\mathbf{X}, t)(x), \quad \mathbf{X} = (\mathbf{X}_1, \ldots, \mathbf{X}_N) \tag{15}$$

the partial derivative with respect to the k-th component of the vector $\mathbf{X}$. Moreover, we denote

$$\frac{\partial u^i}{\partial t}(\mathbf{X}, t)(x), \quad \frac{\partial u^i}{\partial x}(\mathbf{X}, t)(x) \tag{16}$$

the partial derivatives with respect to the variable t and x respectively.

The partial derivative of u with respect to the variable $X_t^{i,N}$ in the first component, evaluated in $\left(\mathbf{X}_t^N, t\right)(X_t^{i,N})$ is assumed to be equal to zero; this reflects the fact the change in position of agent i cannot change the spatial shape of utility in $X_t^{i,N}$, i.e.

$$\frac{\partial u^i}{\partial \mathbf{X}_i}\left(\mathbf{X}_t^N, t\right)(X_t^{i,N}) = 0. \tag{17}$$

The dynamic over space of agent i is the result of a maximization of utility (14) over space, subject to moving costs. In particular, consider the *instantenous* variation of utility of agent i du_t^i as the result of moving from position $X_t^{i,N}$ to $X_t^{i,N} + dt\, dX_t^{i,N}$ (with dt very small), i.e:

$$du_t^{i,N} = \frac{\partial u^i}{\partial x}\left(\mathbf{X}_t^N, t\right)(X_t^{i,N})dX_t^{i,N} + \sum_{j=1}^{N} \frac{\partial u^i}{\partial \mathbf{X}_j}\left(\mathbf{X}_t^N, t\right)(X_t^{i,N})dX_t^{j,N}. \tag{18}$$

Assume quadratic adjustment cost for moving, at each time t, agent i solves the following *static* optimization problem:

$$\max_{\{dX_t^{i,N}\in\mathbb{R}^2\}} \frac{\partial u^i}{\partial x}\left(\mathbf{X}_t^N, t\right)(X_t^{i,N})dX_t^{i,N} + \sum_{j=1}^{N} \frac{\partial u^i}{\partial \mathbf{X}_j}\left(\mathbf{X}_t^N, t\right)(X_t^{i,N})dX_t^{j,N} - \frac{c_M}{2}\left|\left|dX_t^{i,N}\right|\right|^2, \tag{19}$$

where c_M is a parameter measuring the intensity of the moving cost.[2] In Problem (19) we are assuming that agent i has perfect information about the optimal choice of all the other agents, i.e. $dX_t^{j,N}$ is known. However, at the time of the decision this information is not available to the agent, and this means that we have to replace the term $dX_t^{j,N}$ for $j \neq i$ with the expected behavior of agent j's movement dX_t^{j,N,e_i} formulated by agent i (by definition $dX_t^{i,N,e_i} = dX_t^{i,N}$). Therefore, agent i solves at each time t the following problem:

$$\max_{\{dX_t^{i,N}\in\mathbb{R}^2\}} \frac{\partial u^i}{\partial x}\left(\mathbf{X}_t^N, t\right)(X_t^{i,N})dX_t^{i,N} + \sum_{j=1}^{N} \frac{\partial u^i}{\partial \mathbf{X}_j}\left(\mathbf{X}_t^N, t\right)(X_t^{i,N})dX_t^{j,N,e_i} - \frac{c_M}{2}\left|\left|dX_t^{i,N}\right|\right|^2. \tag{20}$$

The first order condition of Problem 20 is given by (see Eq. (17)):

$$dX_t^{i,N} = \frac{1}{c_M}\left[\frac{\partial u^i}{\partial x}\left(\mathbf{X}_t^N, t\right)(X_t^{i,N}) + \frac{\partial u^i}{\partial \mathbf{X}_i}\left(\mathbf{X}_t^N, t\right)(X_t^{i,N})dX_t^{i,N}\right] = \frac{1}{c_M}\frac{\partial u^i}{\partial x}\left(\mathbf{X}_t^N, t\right)(X_t^{i,N}). \tag{21}$$

Eq. (21) represents the optimal movement of agent i in the limit when its time horizon in the optimization is going to zero, i.e. static optimization. In this regard, agent i is extremely myopic, and this is even more important given the presence of sunk mobility costs. However, it is not myopic with respect to the spatial distribution of agents, since the latter are included in u^i. For $\gamma = 1/c_M$ Eq. (21) is equal to Eq. (1).

[2]In a more general setting where cost of moving are not homogenous over space $c_M = c_M\left(X_t^{i,N}, dX_t^{i,N}\right)$, i.e. the cost of moving should be a function of the position of agent and of the direction of movement.

B The microfoundations of Eq. (1) as a Nash equilibrium

In this appendix we explicitly derive, in a simplified setting, the optimal movement of agent i in *discrete time* when it takes also into account the movement of other agents. Differently from Appendix A, therefore, we explicitly consider the optimizing behaviour of other agents, which will lead to search for solution to the contemporaneous agents' optimization as a Nash equilbrium. For the sake of simplicity, below we assume that the space dimension is one, i.e. $X_t^{i,N} \in \mathbb{R}$. All the arguments made below can be repeated in the same manner, with a more cumbersome notation, in any dimension $d > 1$.

Assume that at time t we have N agents, whose position is labeled by $X_t^{i,N}$. Denote by $\mathbf{X}_t^N$ the vector of all the positions of all the agents, i.e $\mathbf{X}_t^N := (X_t^{1,N}, \dots, X_t^{N,N})$. Assume also that the utility function is:

$$\begin{aligned} U: \ \mathbb{R}^N \times \mathbb{R}^+ &\to \big(\mathbb{R} \to \mathbb{R}\big) \\ (\mathbf{X}, t) &\mapsto \big(x \mapsto U(\mathbf{X}, t)(x)\big), \end{aligned} \tag{22}$$

that is for any $\mathbf{X} := (\mathbf{X}^1, \dots, \mathbf{X}^N) \in \mathbb{R}^N$, representing the position of all the agents, and $t \in \mathbb{R}$, we are given a *field of utility* $U(\mathbf{X}, t) : \mathbb{R} \to \mathbb{R}$ which an any given point x in space associate the utility in that point. Hence, the personal utility of each individual depends on its position, but also on the positions of all other agents.

Assume now that in order to maximize its personal utility in the new position at each time t agent i makes a step $dt\, V_t^{i,N}$ over a time interval of duration dt. It is subject to *quadratic movement costs*. However, since its utility also depends on the positions of the other agents, in this maximization procedure it has to keep into account the *expected behavior* of all the other agents. Therefore, agent i at every time $t = 0, dt, 2dt, \dots$ solves the following optimization problem:

$$\max_{\{V_t^{i,N} \in \mathbb{R}\}} U(\mathbf{X}_t^N + dt\, \mathbf{V}_t^{N,e_i}, t + dt)(X_t^{i,N} + dt\, V_t^{i,N}) - dt \frac{c_M}{2} \left| V_t^{i,N} \right|^2, \tag{23}$$

where $\mathbf{V}_t^{N,e_i} := (V_t^{1,N,e_i}, \dots, V_t^{N,N,e_i})$ and $V_t^{i,N,e_i} := V_t^{i,N}$. Roughly speaking, at each time t agent i has a certain belief of where other agents will go after dt, represented by the vector $\mathbf{V}_t^{N,e_i}$. Therefore, it acts accordingly to that belief, maximizing its own personal utility taking into account its moving costs. c_M transforms the movement cost in utility and it is defined in terms of time interval dt.

The first order condition for Problem (23) is:

$$\begin{aligned}0 &= \sum_{j=1,j\neq i}^{N} \frac{\partial U}{\partial \mathbf{X}_j}(\mathbf{X}_t^N + dt\,\mathbf{V}_t^{N,e_i}, t+dt)(X_t^{i,N} + dt\,V_t^{i,N})\frac{dV^{j,N,e_i}}{dV_t^{i,N}}dt + \\ &+ \frac{\partial U}{\partial \mathbf{X}_i}(\mathbf{X}_t^N + dt\,\mathbf{V}_t^{N,e_i}, t+dt)(X_t^{i,N} + dt\,V_t^{i,N})dt + \\ &+ \frac{\partial U}{\partial x}(\mathbf{X}_t^N + dt\,\mathbf{V}_t^{N,e_i}, t+dt)(X_t^{i,N} + dt\,V_t^{i,N})dt - dtc_M V_t^{i,N}. \end{aligned} \tag{24}$$

If we assume that the expected behavior of agent j, V_t^{j,N,e_i}, does not depend on the optimal behavior of agent i, $V_t^{i,N}$, then all terms in the first summation vanish. Moreover, if Condition (17) holds, any change in the position of agent i has no direct effect on the filed of utility evaluated in its position since the only thing that matters is its relative position with respect to others. Therefore, Eq. (24) reduces to:

$$\frac{\partial U}{\partial x}(\mathbf{X}_t^N + dt\,\mathbf{V}_t^{N,e_i}, t+dt)(X_t^{i,N} + dt\,V_t^{i,N}) - c_M V_t^{i,N} = 0. \tag{25}$$

We now solve Eq. (25) explicitly by assuming a specific shape of utility function as it follows:

$$U(\mathbf{X}, t)(x) = \frac{1}{N}\sum_{j=1}^{N} W(x - \mathbf{X}_j), \tag{26}$$

where the *kernel function* W is defined as

$$W(x) = -\frac{1}{2}x^2 \mathbb{1}_{|x|\leq 1} - \mathbb{1}_{|x|>1}, \tag{27}$$

which is reported in Figure 6 ($\mathbb{1}_{|x|\leq 1}$ means that the value is equal to 1 when $|x| \leq 1$ and 0 otherwise). In particular, $W'(0) = 0$ in agreement with the Condition (17), while the partial derivative of U with respect to $\mathbf{X}^i$ is always zero when evaluated in $\mathbf{X}^i$. The utility function (26) is higher when all the agents are close to each other; in particular, it reaches its maximum value of 0 when all agents are in the same position, i.e. when the level of agglomeration is maximum (see Figure 2).

Figures/Wx-eps-converted-to.pdf

Figure 6: The shape of kernel function $W(x)$, the base of individual utility in Eq. (26).

We can now rewrite Eq. (25):

$$V_t^{i,N} = -\frac{1}{c_M}\frac{1}{N}\sum_{j=1}^{N}\left[(X_t^{i,N} + dtV_t^{i,N}) - (X_t^{j,N} + dtV_t^{j,N,e_i})\right]\mathbb{1}_{\left\{\left|(X_t^{i,N}+dtV_t^{i,N})-(X_t^{j,N}+dtV_t^{j,N,e_i})\right|\leq 1\right\}}. \tag{28}$$

Assume now that agents at time t are all within distance one from each other, and that the same holds at time $t + dt$, so that all the indicator functions in the previous sum are equal to one. This is not restrictive if one assume dt to be very small, so that the movement that each agent will perform is very small. Therefore, from Eq. (28) we have

$$V_t^{i,N} = -\frac{1}{c_M}\left[X_t^{i,N} - \frac{1}{N}\sum_{j=1}^{N} X_t^{j,N}\right] - \frac{dt}{c_M}\left[V_t^{i,N} - \frac{1}{N}\sum_{j=1}^{N} V_t^{j,N,e_i}\right], \tag{29}$$

from which

$$V_t^{i,N} = -\frac{1}{c_M + dt - dt/N}\left[X_t^{i,N} - \frac{1}{N}\sum_{j=1}^{N} X_t^{j,N} - \frac{dt}{N}\sum_{j=1,j\neq i}^{N} V_t^{j,N,e_i}\right]. \tag{30}$$

From Eq. (30) we can conclude that:

1. the larger c_M the smaller the optimal movement of agent i $V_t^{i,N}$, which reflects the quadratic moving costs in Eq. (23).

2. By considering a continuous time process, instead of a discrete one, i.e. we let dt going to zero, the expected choices of other agents don't matter in the optimal choice of agent i (as seen from agent i). The intuition is the following: in continuous time the choice of agent i is repeated infinitely many times, and each of this choice, taken singularly (and not as a continuous choice process) has no effect on the dynamic. The same effect holds

for the choice of all other agents.

3. Either if we let dt going to zero, or we assume that agent i has zero-movement expectation on other agents, i.e. $V^{j,N,e_i} = 0$ for $j \neq i$, the optimal strategy to maximize utility (26) is to move towards the barycenter of spatial distribution of agents, which is the mean position of agents. This is coherent with the intuition that the particular shape of utility (26) should tend to aggregate agents over space. Moreover, the barycenter of the position of the agents remains unchanged after each forward step, i.e. the barycenter is an *invariant* for the dynamic.

4. In the continuous time case, i.e. when dt is going to zero, we find that the optimal movement reduces to follow the *gradient of the utility function U*, that is precisely Eq. (1).

So far, we have assumed that agent i has a generic expectation on the movement of other agents. One possible choice is to search for an equilibrium in each time where agent i, for $i = 1, \ldots, N$, solves Problem (23) and expectation are realized, i.e. $V_t^{j,N,e_i} = V_t^{j,N}$. This amount to consider a Nash equilibrium of a non-cooperative game where agent i solve Problem (23), for $i = 1, \ldots, N$. This implies that we have to solve a system of N linear equations derived by Eq. (29), i.e.:

$$V_t^{i,N} = -\frac{1}{c_M}\left[X_t^{i,N} - \frac{1}{N}\sum_{j=1}^{N} X_t^{j,N}\right] - \frac{dt}{c_M}\left[V_t^{i,N} - \frac{1}{N}\sum_{j=1}^{N} V_t^{j,N}\right], \quad i = 1, \ldots, N, \tag{31}$$

and finally

$$V_t^{i,N}(1 + (N-1)b) - b\sum_{j=1, j\neq i}^{N} V_t^{j,N} = -\frac{1}{c_M}\left[X_t^{i,N} - \frac{1}{N}\sum_{j=1}^{N} X_t^{j,N}\right] \quad b = \frac{dt}{c_M N}, \quad i = 1, \ldots, N. \tag{32}$$

This system can be rewritten in matrix-vector form as $A\mathbf{V}_t^N = \tilde{\mathbf{X}}_t^N$, where:

$$A := \begin{bmatrix} 1+(N-1)b & -b & & \cdots & -b \\ -b & 1+(N-1)b & -b & \cdots & -b \\ \vdots & & \ddots & & \vdots \\ \vdots & & & \ddots & -b \\ -b & \cdots & \cdots & -b & 1+(N-1)b \end{bmatrix}, \quad \mathbf{V}_t^N := \begin{bmatrix} V_t^{1,N} \\ V_t^{2,N} \\ \vdots \\ \vdots \\ V_t^{N,N} \end{bmatrix}, \tag{33}$$

$$\tilde{\mathbf{X}}_t^N = \begin{bmatrix} -\frac{1}{c_M}\left(X_t^{1,N} - \frac{1}{N}\sum_{j=1}^N X_t^{j,N}\right) \\ -\frac{1}{c_M}\left(X_t^{2,N} - \frac{1}{N}\sum_{j=1}^N X_t^{j,N}\right) \\ \vdots \\ \vdots \\ -\frac{1}{c_M}\left(X_t^{N,N} - \frac{1}{N}\sum_{j=1}^N X_t^{j,N}\right) \end{bmatrix}. \tag{34}$$

One can verify that the inverse of the matrix A is given by

$$A^{-1} = \frac{1}{Nb+1}\begin{bmatrix} b+1 & b & & \cdots & b \\ b & b+1 & b & \cdots & b \\ \vdots & & \ddots & & \vdots \\ \vdots & & & \ddots & b \\ b & \cdots & \cdots & b & b+1 \end{bmatrix}, \tag{35}$$

so that the solution of System of Eqq. (32) is given by:

$$V_t^{i,N} = -\frac{1}{c_M + dt}\left[X_t^{i,N} - \frac{1}{N}\sum_{j=1}^N X_t^{j,N}\right]. \tag{36}$$

Therefore, the Mash equilibrium is the analogous to the case where agent i has zero-movement belief on the other agents' movement (compare Eq. (36) with Eq. (30)).

C Dynamic spatial aggregation in the sequence of Nash equilibria

In this appendix we will show that in the sequence of Nash equilibria in discrete time described in Appendix B (or analogously in the continuous time limit) increases the spatial agglomeration of agents, when as measure of spatial inequality is used the *spatial Gini Index* defined as

$$G(\mathbf{X}) := \frac{1}{2\overline{\mathbf{X}}}\left(\frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{N}|\mathbf{X}_i - \mathbf{X}_j|\right), \tag{37}$$

where

$$\overline{\mathbf{X}} := \frac{1}{N}\sum_{i=1}^{N}\mathbf{X}^i \tag{38}$$

is the empirical mean of the position of the agents. The spatial Gini Index has the remarkable property to be bounded between 0 and 1, where 0 is the maximum level of agglomeration while 1 is the maximum level of dispersion.

We first consider the level of aggregation at time t, $G(\mathbf{X}_t^N)$ and show that each term of the sum in Eq. (37) gets smaller after a time step. Since the value of the mean position remains unchanged with each time step, the term $\overline{\mathbf{X}}_t^N$ is irrelevant for the calculation. Considering (36) we have:

$$\begin{aligned}
&\left|X_{t+dt}^{i,N} - X_{t+dt}^{j,N}\right| = \left|\left(X_t^{i,N} + dtV_t^{i,N}\right) - \left(X_t^{j,N} + dtV_t^{j,N}\right)\right| = \\
&= \left|\left(X_t^{i,N} - X_t^{j,N}\right) + dt\left(V_t^{i,N} - V_t^{j,N}\right)\right| = \left|\left(X_t^{i,N} - X_t^{j,N}\right) - \frac{dt}{c_M + dt}\left(X_t^{i,N} - X_t^{j,N}\right)\right| = \\
&= \left|\left(1 - \frac{dt}{c_M + dt}\right)\left(X_t^{i,N} - X_t^{j,N}\right)\right| = \left(1 - \frac{dt}{c_M + dt}\right)\left|\left(X_t^{i,N} - X_t^{j,N}\right)\right| \leq \left|\left(X_t^{i,N} - X_t^{j,N}\right)\right|
\end{aligned} \tag{39}$$

since the term $\left(1 - \frac{dt}{c_M+dt}\right)$ is always positive and smaller than one. This shows that the aggregation level after one step is higher, in the sense that the spatial Gini Index is smaller than that at the initial time.

D A sketch of the proof of Theorem 3.1

For the sake of simplicity, we will analyze only the case where $n(t,x) = 0$, i.e. there is no change in the total number of agents. The case $n(t,x) \neq 0$ can be treated by introducing some technicalities related to Poisson processes, but the proof just becomes more involuted without without posing any additional significant conceptual difficulty, see for example Catellier et al. (2021). Moreover, again for the sake of simplicity, in the sketch of the proof we will also assume that endogenous amenities are not present. The technique used below can be also applied, by adding some additional calculations for endogenous amenities, in the general case. Finally, we will also assume the domain Ω is the two dimensional thorus $\mathbb{T}^2 = \mathbb{R}^2/\mathbb{Z}^2$. This simplifies the proof allowing to neglect the boundary effects, but is not restrictive since the same technique of proof can be applied in any domain $\Omega \subseteq \mathbb{R}^2$ with boundary of class C^1.

Assume that at the time $t = 0$ agents are independently distributed at random in the domain Ω, following a common probability density distribution on Ω called $y_0(z)$. For $t > 0$ each agents position evolves by obeying the following Stochastic Differential Equations (SDE):

$$dX_t^{i,N} = \gamma \left[\nabla_x w^R \left(\mathbf{X}_t^N, t \right) \left(X_t^{i,N} \right) + \nabla_x A_{ES} \left(t, X_t^{i,N} \right) \right] dt + \gamma \sigma dB_t^i,$$

or by making each term more explicit

$$\begin{aligned} dX_t^{i,N} &= \gamma G(X_t^{i,N}) \left[\sum_{j=1}^{N} \frac{1}{N} W_h^P (X_t^{i,N} - X_t^{j,N}) \right] l^N(t, X_t^{i,N})^{\beta-1} \\ &+ \gamma \nabla_x A_{ES} \left(t, X_t^{i,N} \right) dt \\ &+ \gamma \sigma dB_t^i. \end{aligned} \tag{40}$$

We now consider the *empirical distribution* of all the agent's positions, that is

$$E_t^N := \frac{1}{N} \sum_{i=1}^{N} \delta_{X_t^{i,N}}, \tag{41}$$

where δ_z is the random variable on $\mathbb{T}^2$ with unitary mass in the point z. Notice that E_t^N is a continuous set of random variables on $\mathbb{T}^2$ depending on time. For any given $N \in \mathbb{N}$ and $t > 0$, this distribution is singular, in the sense that is a distribution over $\mathbb{T}^2$ which does not admit a probability density function, since it has a positive probability only on a finite set (corresponding to the position of the N agents). However, when N tends to infinity the family of random variable E_t^N becomes diffuse and converges (in distribution) to a continuous family of random variables over $\mathbb{T}^2$, labelled by E_t for any $t > 0$. The distribution of E_t is now regular and admits a probability density function for every t, called $l(t,z)$. An explicit expression for $l(t,z)$ for every t is not available. However we can prove that the probability density function $l(t,z)$ is the solution to Eq. (13). This is precisely the content of the following theorem, which is equivalent to Theorem 3.1:

Teorema D.1. *The stock of labour in position x at time t, $l^N(t,x)$ converges in probability as N tends to infinity, in the suitable topology, to the unique solution of system of Eqq.* (13).

Recall that the stock of population at position x at time t is defined by:

$$l^N(t,x) = \sum_{i=1}^{N} \frac{1}{N} \cdot \theta_N(x - X_t^{i,N}),$$

where $\theta_N(x) = N^{2\lambda}\theta(N^\lambda x)$, $\theta : \mathbb{T}^2 \to \mathbb{R}$ is a function that is non negative, C^∞ and integrate to one. We will sometimes refer to the function $l^N(t,x)$ as the *mollified empirical measure.*

Remark. The parameter λ can be taken in the intervall $[0,1]$. However, some choices of λ are not possible in our framework, either because they don't make sense as modeling assumption, or because they are currently out of reach for the present mathematical theory. In particular:

- $\lambda = 0$: this case correspond to the classical *mean-field case*, where the radius of interaction is fixed. This choice leads to a non local macroscopic dynamics, not suitable for our purposes;

- $\lambda \in (0,1)$: this corresponds to the theory of *moderate interactions* originally developed in Oelschläger (1985) with some additional restriction on the range of λ. The theory has been extensively generalized, e.g. in Méléard and Roelly-Coppoletta (1987) allowing the parameter λ to be any number in $(0,1)$. Our result fall in this category, altough without reaching the full generality of the case $\overline{\lambda} = 1$, therefore we assume $\lambda \in (0, \overline{\lambda})$ with $\overline{\lambda} < 1$;

- $\lambda = 1$: this is the case of *local interactions* that is still not well understood, with very few results in the literature: Uchiyama (2000); Varadhan (1991).

D.1 Preliminaries

In this section we provide some preliminaries needed for the proof of the theorem.

Definition D.2 (Weak solutions of (13))**.** We say a function l of suitable regularity is a weak solution of system of equations (13) if for every test function $\varphi \in C^\infty(\mathbb{T}^2;\mathbb{R})$ it holds

$$\begin{aligned} \langle l(t), \varphi \rangle &= \langle l(0), \varphi \rangle + \frac{\gamma^2 \sigma^2}{2} \int_0^s \langle l(t), \Delta_x \varphi \rangle \, ds \\ &+ \gamma \int_0^t \left\langle l(s), \nabla_x \varphi \cdot \nabla_x \left(G \, (W_h^P * l)(s) \, l(s)^{\beta-1} \right) \right\rangle \, ds \\ &+ \gamma \int_0^t \langle l(s), \nabla_x \varphi \cdot \nabla_x A_{ES}(s) \rangle \, ds. \end{aligned} \tag{42}$$

Introduce the *empirical measure* of the agents position

$$S_t^N(dx) = \frac{1}{N} \sum_{i=1}^N \delta_{X_t^{i,N}}(dx),$$

that is random element of the space $C([0,T];\mathbf{P}_1(\mathbb{T}^2))$ where

$$\mathbf{P}_1(\mathbb{T}^2) := \left\{ \mu \text{ probability measure on } (\mathbb{T}^2, \mathcal{B}(\mathbb{T}^2)) \,\middle|\, \int_{\mathbb{T}^2} |x| \; \mu(dx) < \infty \right\}$$

is the space of all probability measure on the Borel sets of $\mathbb{T}^2$, with finite first moment. We endow this space with the Wasserstein-1 metric, that can be defined equivalently as

$$\mathcal{W}_1(\mu,\nu) := \sup_{[\varphi]_{Lip} \le 1} \left| \int_{\mathbb{T}^2} \varphi \, d\mu - \int_{\mathbb{T}^2} \varphi \, d\nu \right|$$

where $[\varphi]_{Lip}$ is the usual Lipschitz seminorm. Endowed with this metric the space $\mathbf{P}_1(\mathbb{T}^2)$ becomes a complete separable metric space, whose convergence implies the weak convergence of probability measures.

D.2 A candidate for the limit

We recall that for any test function φ it holds

$$\left\langle S_t^N, \varphi \right\rangle = \int_{\mathbb{T}^2} \varphi(x) S_t^N(dx) = \frac{1}{N} \sum_{i=1}^N \varphi(X_t^{i,N}).$$

Hence, by taking the family of function parametrized by x $\varphi(\cdot) = \theta_N(x - \cdot)$ we have

$$\left\langle S_t^N, \theta_N(x - \cdot) \right\rangle = \frac{1}{N} \sum_{i=1}^{N} \theta_N(x - X_t^{i,N}) = l^N(t, x)$$

Lemma D.3 (Itô Formula). *For any test function φ it holds*

$$\begin{aligned} d\varphi(X_t^{i,N}) &= \frac{\gamma^2\sigma^2}{2} \Delta_x \varphi(X_t^{i,N})\, dt \\ &+ \nabla\varphi(X_t^{i,N}) \cdot \nabla w^N(t, X_t^{i,N})\, dt \\ &+ \nabla\varphi(X_t^{i,N}) \cdot \nabla A_{ES}(t, X_t^{i,N})\, dt \\ &+ \gamma\sigma \nabla\varphi(X_t^{i,N}) \cdot dB_t^i, \end{aligned} \tag{43}$$

and

$$\begin{aligned} d\left\langle S_t^N, \varphi \right\rangle &= \frac{\gamma^2\sigma^2}{2} \left\langle S_t^N, \Delta_x \varphi \right\rangle dt \\ &+ \left\langle S_t^N, \nabla_x\varphi \cdot \nabla_x w^N(t, \cdot) \right\rangle dt \\ &+ \left\langle S_t^N, \nabla_x\varphi \cdot \nabla_x A_{ES}(t, \cdot) \right\rangle dt \\ &+ \frac{\gamma\sigma}{N} \sum_{i=1}^{N} \nabla_x \varphi(X_t^{i,N}) \cdot dB_t^i, \end{aligned} \tag{44}$$

and

$$\begin{aligned} dl^N(t, x) &= \frac{\gamma^2\sigma^2}{2} \Delta_x l^N(t, x)\, dt \\ &- \gamma \frac{1}{N} \sum_{i=1}^{N} \nabla\theta_N(x - X_t^{i,N}) \cdot \nabla w^N(t, X_t^{i,N})\, dt \\ &- \gamma\, div_x \left(l^N(t, x) \nabla_x A_{ES}(t, x) \right) dt \\ &+ \frac{\gamma\sigma}{N} \sum_{i=1}^{N} \nabla_x \theta_N(x - X_t^{i,N}) \cdot dB_t^i. \end{aligned} \tag{45}$$

Proof. The first part follows easily by applying Itô formula to the function $\varphi(X_t^{i,N})$ and using Eq (40). The second part simply follows from the first part using the linearity of the sum, while the last part is obtained by taking the family of functions parametrized by x $\theta_N(x - \cdot)$ and applying Itô formula, then using again the linearity of the sum. □

The expression for the drift in the second line on the right-hand side of equation (45) can be expressed as

$$
\begin{aligned}
-\frac{1}{N}\sum_{i=1}^{N}&\nabla_x\theta_N(x-X_t^{i,N})\cdot\nabla_x w^N(t,X_t^{i,N}) \\
&= -\int_{\mathbb{T}^2}\nabla_x\theta_N(x-x')\cdot\nabla_x w^N(t,x')\,S_t^N(dx') \\
&= -\mathrm{div}_x\int_{\mathbb{T}^2}\theta_N(x-x')\nabla_x w^N(t,x')\,S_t^N(dx').
\end{aligned}
$$

This is not the right expression to obtaine the desired limit in Eq. (13). The proper one, in order to obtain the correct expression when taking the limit in N, is

$$
-\mathrm{div}_x\int_{\mathbb{T}^2}\theta_N(x-x')\nabla_x w^N(t,x)\,S_t^N(dx') = -\mathrm{div}_x(l^N(t,x)\nabla_x w^N(t,x)).
$$

Hence, we rewrite Eq. (45) as the proper expression, plus a remainder that we will have to show that vanishes in the limit

$$
\begin{aligned}
dl^N(t,x) &= \frac{\gamma^2\sigma^2}{2}\Delta_x l^N(t)\,dt \\
&- \gamma\,\mathrm{div}_x(l^N(t,x)\nabla_x w^N(t,x))\,dt \\
&- \gamma\,\mathrm{div}_x\left(l^N(t)\nabla_x A_{ES}(t,x)\right)\,dt \\
&+ R_t^N(x)\,dt + dM_t^{N,\theta_N}(x)
\end{aligned}
\tag{46}
$$

where

$$
M_t^{N,\theta_N}(x) = \frac{\gamma\sigma}{N}\sum_{i=1}^{N}\nabla_x\theta_N(x-X_t^{i,N})\cdot B_t^i
$$

is a martingale, and

$$
R_t^N(x) = \mathrm{div}_x\int_{\mathbb{T}^2}\theta_N(x-x')\left[\nabla_x w^N(x)-\nabla_x w^N(x')\right]\,S_t^N(dx').
$$

Eq. 46 provides us with a candidate for the limit behaviour of system of Eq. (40).

D.3 Steps for the proof

The rigorous proof of Theorem D.1 needs some intermediate lemmas that one has to prove first. The general strategy is that of *compactness*, and is made of three main steps.

(*Compactness*) Firstly one has to prove that the sequence $(S_t^N)_{N\in\mathbb{N}}$ lies uniformly in some compact subset of $\mathbf{P}_1(\mathbb{T}^2)$[3] This yelds the existence of limits of every subsequence. In our case this follows from an uniform estimate in $L^1(\Omega)$ for each of the agents position. We remark that the convergence is only in distribution since is the one induced by the metric used on $\mathbf{P}_1(\mathbb{T}^2)$. However, is a well known fact that convergence in ditribution of random variables to a deterministic limit implies convergence in probabilty.

(*Characterization*) Secondly one has to provide a characterization of limit points as solutions to Eq. (13). This usually follows from the convergence of each of the terms in Eq. (45) to the corresponding one in the limit equation.

(*Uniqueness*) Finally one has to prove a uniqueness results for Eq. (13). This final steps is necessary to pass from the convergence of every subsequence, as argued in the first step, to the convergence of the full sequence. The rationale is the following: since every subsequence is converging to solutions of Eq. (13), if the solution is unique then every subsequence is converging to the same limt. One then has only to realize that in every metric space, if every subsequence converges to the same limit, the same holds for the whole sequence.

In order to complete the three previous steps we highlight some key points of the procedure:

- One proves that there exists a constant C_T that depends on T but not on N, such that

$$\mathbf{E}\left[\left|\left| \nabla_x l^N(t) \right|\right|_{L^2(\mathbb{T}^2)} \right] \le C_T. \tag{47}$$

  Here we will also need to assume additional hypothesis on the production function. In particular, we will assume that to be equal to l^β when l is greater than some small threshold value $\underline{l}$, and instead has the shape of a parabola for values of l between zero and $\underline{l}$. One possible expression is given by

$$f(l) = \begin{cases} l^\beta + (-\frac{1}{2}\beta^2 + \frac{3}{2}\beta - 1)\underline{l}^\beta & \text{if } l > \underline{l}, \\ \frac{1}{2}\beta(\beta-1)\underline{l}^{\beta-2} l^2 + \beta(2-\beta)\underline{l}^{\beta-1} l & \text{if } 0 \le l \le \underline{l}. \end{cases}$$

  In this way we keep the usual assumption on the production function, that is positive first derivative with negative second derivative, but the first derivative is now globally

[3]The notion of *compactness* here is simply the classical one of compacts sets in metric spaces.

bounded for any values of l. This first step is necessary to complete the first step about *Compactness* of the sequence $(S_t^N)_{N\in\mathbb{N}}$ in $\mathbf{P}^1(\mathbb{T}^2)$.

- Assume that $S_t^N \overset{N\to\infty}{\to} \mu_t$ in $\mathbf{P}_1(\mathbb{T}^2)$, where μ_t is a probability measure with density $l(t,x)$ (or analogously for a subsequence $S_t^{N_k}$ from the first step). Then we have to check that the expression for wages converges to the analogous limit expression of wages

$$\begin{aligned} &w^N(t,x) \\ &= G(x)\frac{1}{N}\sum_{i=1}^N W_h^P(x - X_t^{i,N})\left(\frac{1}{N}\sum_{i=1}^N \theta_N(x - X_t^{i,N})\right)^{\beta-1} \overset{N\to\infty}{\longrightarrow} G(x)(W_h^P * l)(t,x)l(t,x)^{\beta-1} \\ &\qquad = w(t,x). \end{aligned}$$

This part is necessary to provide a characterization of limit points (*Characterization*). We also need to show that the martingale term $M_t^{N,\theta_N}(x)$ goes to zero in probability as N goes to infinity. This however is readily done using that we have an explicit expression for the quadratic variation

$$\left[M^{N,\theta_N}\right]_t = \frac{\gamma^2\sigma^2}{N^2}\sum_{i=1}^N \left|\nabla_l \theta_N(x - X_t^{i,N})\right|^2$$

since the Brownian motions B_t^i are independent, and use classical martingale inequalities. In this step we will need to assume $\lambda < 1/4$ in order to prove such inequality, as well as use Eq. (47). Finally, we need to show that the remainder $R_t^N(x)$ goes to zero as well, as N goes to infinity. This again follows from the estimate in Eq. (47), using the fact that the difference $\left[\nabla_x w^N(x) - \nabla_x w^N(x')\right]$ is non zero only when x and x' are close, since it is multiplied by $\theta_N(x - x')$, whose support is shrinking with N.

- To show uniqueness of solutions of Eq. (13) one follows the classical strategy of assuming the existence of two different solutions $l^1(t,x)$ and $l^2(t,x)$ with the same initial condition at time $t = 0$. Then we consider the difference of the two and study the evolution in the L^2-norm in $\mathbb{T}^2$

$$\left|\left|l^1(t) - l^2(t)\right|\right|_{L^2(\mathbb{T}^2)}^2 = \int_{TT^2} \left|l^1(t,x) - l^2(t,x)\right|^2 dx$$

and show that it satisfies

$$\left|\left|l^1(t) - l^2(t)\right|\right|_{L^2(\mathbb{T}^2)}^2 \leq C_T \int_0^t \left|\left|l^1(s) - l^2(s)\right|\right|_{L^2(\mathbb{T}^2)}^2 \, ds$$

for some constant C_T. Then, classical Gronwall Lemma implies that if $||l^1(0) - l^2(0)||_{L^2(\mathbb{T}^2)} = 0$ then $||l^1(t) - l^2(t)||_{L^2(\mathbb{T}^2)} = 0$ for all $t > 0$, so that solutions are unique. This provides a proof of (*Uniqueness*).

References

Allen, T. and C. Arkolakis (2014). Trade and the topography of the spatial economy. *The Quarterly Journal of Economics 129*(3), 1085–1140.

Anton Bovier, F. d. H. a. (2015). Metastability: A potential-theoretic approach. *Grundlehren der mathematischen Wissenschaften №351*.

Aoki, M. and H. Yoshikawa (2011). *Reconstructing macroeconomics: a perspective from statistical physics and combinatorial stochastic processes*. Cambridge University Press.

Arbia, G. (2001). Modelling the geography of economic activities on a continuous space. *Papers in Regional Science 80*(4), 411–424.

Boucekkine, R., C. Camacho, and B. Zou (2009). Bridging the gap between growth theory and the new economic geography: The spatial ramsey model. *Macroeconomic Dynamics 13*(1), 20–45.

Carrillo, J. A., K. Craig, and Y. Yao (2019). Aggregation-diffusion equations: dynamics, asymptotics, and singular limits. In *Active Particles, Volume 2*, pp. 65–108. Springer.

Carrillo, J. A., S. Hittmeir, B. Volzone, and Y. Yao (2019). Nonlinear aggregation-diffusion equations: radial symmetry and long time asymptotics. *Inventiones mathematicae 218*(3), 889–977.

Catellier, R., Y. D'angelo, and C. Ricci (2021). A mean-field approach to self-interacting networks, convergence and regularity. *Mathematical Models and Methods in Applied Sciences 31*(13), 2597–2641.

Ciccone, A. and R. E. Hall (1996). Productivity and the density of economic activity. *The American Economic Review*, 54–70.

Colander, D., P. Howitt, A. Kirman, A. Leijonhufvud, and P. Mehrling (2008). Beyond dsge models: toward an empirically based macroeconomics. *American Economic Review 98*(2), 236–40.

Davis, D. R. and D. E. Weinstein (2002). Bones, bombs, and break points: the geography of economic activity. *American economic review 92*(5), 1269–1289.

Duranton, G. and D. Puga (2020). The economics of urban density. *Journal of Economic Perspectives 34*(3), 3–26.

Durlauf, S. N. (1994). Spillovers, stratification, and inequality. *European Economic Review 38*(3-4), 836–845.

Durlauf, S. N., P. A. Johnson, and J. R. W. Temple (2005). Growth econometrics. In P. Aghion and S. N. Durlauf (Eds.), *Handbook of Economic Growth*, Volume 1A, pp. 555—677. Elsevier.

Durlauf, W. B. A. S. N. (1997). *The Economy As A Complex Evolving System II*. Santa Fe Institute Studies in the Sciences of Complexity Lecture Notes. Westview Press.

Evers, J. H. and T. Kolokolnikov (2016). Metastable states for an aggregation model with noise. *SIAM Journal on Applied Dynamical Systems 15*(4), 2213–2226.

Facchini, A. and C. Mocenni (2013). Recurrence methods for the identification of morphogenetic patterns. *Plos one 8*(9), e73686.

Flandoli, F. (2016). Elements of mathematical oncology.

Glaeser, E. L. and D. C. Mare (2001). Cities and skills. *Journal of labor economics 19*(2), 316–342.

Glaeser, E. L. and M. G. Resseger (2010). The complementarity between cities and skills. *Journal of Regional Science 50*(1), 221–244.

Howitt, P. (2006). Coordination issues in long-run growth. *Handbook of computational economics 2*, 1605–1624.

Krugman, P. (1994). Complex landscapes in economic geography. *The American Economic Review 84*(2), 412–416.

LeBaron, B. and L. Tesfatsion (2008). Modeling macroeconomies as open-ended dynamic systems of interacting agents. *American Economic Review 98*(2), 246–50.

Méléard, S. and S. Roelly-Coppoletta (1987). A propagation of chaos result for a system of particles with moderate interaction. *Stochastic processes and their applications 26*, 317–332.

Michalopoulos, S. and E. Papaioannou (2018). Spatial patterns of development: A meso approach. *Annual Review of Economics 10*, 383–410.

Mocenni, C., A. Facchini, and A. Vicino (2010). Identifying the dynamics of complex spatio-temporal systems by spatial recurrence properties. *Proceedings of the National Academy of Sciences 107*(18), 8097–8102.

Morale, D., V. Capasso, and K. Oelschläger (2005). An interacting particle system modelling aggregation behavior: from individuals to populations. *Journal of mathematical biology 50*(1), 49–66.

Oelschläger, K. (1985). A law of large numbers for moderately interacting diffusion processes. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete 69*(2), 279–322.

Rossi-Hansberg, E. (2019). Geography of growth and development. In *Oxford Research Encyclopedia of Economics and Finance*.

Sznitman, A.-S. (1991). Topics in propagation of chaos. In *Ecole d'été de probabilités de Saint-Flour XIX—1989*, pp. 165–251. Springer.

Train, K. E. (2009). *Discrete choice methods with simulation*. Cambridge university press.

Uchiyama, K. (2000). Pressure in classical statistical mechanics and interacting brownian particles in multi-dimensions. In *Annales Henri Poincaré*, Volume 1, pp. 1159–1202. Springer.

Varadhan, S. (1991). Scaling limits for interacting diffusions. *Communications in mathematical physics 135*(2), 313–353.

Xepapadeas, A. and A. Yannacopoulos (2016). Spatial growth with exogenous saving rates. *Journal of Mathematical Economics 67*, 125–137.