amse
école d'économie d'aix-marseille
aix-marseille school of economics

Working Papers / Documents de travail

When Lions meets Krugman: A mean-eld game theory of spatial dynamics

Mohamed Bahlali
Raouf Boucekkine
Quentin Petit

WP 2025 - Nr 17

When Lions meets Krugman: A mean-field game theory of spatial dynamics*

M. Bahlali[†] R. Boucekkine[‡] Q. Petit[§]

October 6, 2025

Abstract

We propose a mean-field game (MFG) set-up to study the dynamics of spatial agglomeration in a continuous space-time framework where trade across locations may follow a broad class of static gravity models. Forward-looking intertemporal utility-maximizing agents work and migrate in a two-dimensional geography and face idiosyncratic shocks. Equilibrium wages and prices depend on their common distribution and adjust statically according to the underlying trade model. We first prove existence and uniqueness of the static trade equilibrium. We then prove existence of dynamic equilibria. In the case of Krugman (1996)'s racetrack economy, we obtain closed-form solutions for small sinusoidal perturbations around the steady state, and we identify the sets of parameters that lead to agglomeration or dispersion. We exploit the MFG structure of the model to explicitly quantify how uncertainty and forward-looking expectations contribute to agglomeration and dispersion. In particular, we show that, regardless of the static trade model, forward-looking expectations always promote agglomeration, but cannot reverse the dominant pattern that would arise under myopic behavior.

*We thank participants at the workshop on the mean-field games in economics the Durham IAS, March 2025, for many useful comments. The project leading to this publication has received funding from the French government under the "France 2030" investment plan managed by the French National Research Agency (reference :ANR-17-EURE-0020) and from Excellence Initiative of Aix-Marseille University - A*MIDEX. Petit acknowledges support from the Finance for Energy Market Research Centre between ENSAE, Université Paris-Dauphine, Ecole Polytechnique and EDF R&D; and the Finance and Sustainable Development Chair between Université Paris-Dauphine, Ecole Polytechnique, EDF R&D and Credit Agricole CIB. The usual disclaimer applies.

[†]Aix-Marseille University, CNRS, AMSE, France. E-mail: mohamed.bahlali@univ-amu.fr

[‡]Aix-Marseille University, CNRS, AMSE, France. E-mail: raouf.boucekkine@univ-amu.fr

[§]EDF & FiME Lab, Paris, France. E-mail: quentin.petit@edf.fr

1 Introduction

Looking back, economic geography has made substantial progress since the seminal core–periphery model of Krugman (1991). The spatial mobility of factors, labor in particular, remains an important line of research. When migration is modeled endogenously, it is usually assumed that workers move in order to increase some measure of utility. Yet, although migration decisions are inherently long-term, models have long assumed that workers are myopic. It is only in 2001 that Baldwin (2001) introduced forward-looking migration in the core-periphery model, although ultimately only producing a numerical exploration. Ottaviano (2001) is the first work yielding some analytical results in a forward-looking model of footloose entrepreneurs. Both models, however, remain within the core–periphery framework. The extension of forward-looking spatial models to a large number of regions came later, in the 2010s. An important contribution is Mossay (2013), who introduces endogenous forward-looking local migration into Krugman's racetrack economy. These theoretical works paved the way for a new generation of quantitative models that study how forward-looking behavior affects spatial dynamics (see Desmet and Parro (2025) for a recent survey on the subject). For instance, Desmet et al. (2018) examine the role of geographic heterogeneity and migration frictions in development; Caliendo et al. (2019) focus on trade shocks and labor mobility; Kleinman et al. (2023) analyze the interaction between forward-looking capital accumulation and migration; and Bilal and Rossi-Hansberg (2023) study how expectations about future extreme climate events affect labor mobility in the United States.

Most of these dynamic models share a two-layer structure. At each point in time, a *static equilibrium* determines local wages and prices given the spatial distribution of production factors (labor, capital). Depending on the context, this static layer corresponds to one of the many trade, regional, or urban models in the literature. On top of this, agents make *forward-looking migration decisions*. Their dynamics are governed by a Bellman equation, where individual incentives depend on future wages, prices, and the evolving spatial distribution of agents. The two layers interact because the *instantaneous utility in the Bellman equation depends on wages and prices computed in the static equilibrium*.

Theoretical results for such dynamic models are still scarce, especially regarding the existence of equilibria. Kleinman et al. (2023) and Bilal and Rossi-Hansberg (2023) linearize their models to obtain closed-form solutions for the transition path following small shocks around the steady state. Caliendo et al. (2019) do not prove existence and uniqueness in their full framework; instead, they refer to results from a simpler dynamic model with exogenous prices (Cameron et al. (2007), see also Allen et al. (2024), Online Appendix A.6, for a similar simplified setting). Yet, proving existence in the general case is important to ensure that the model is well defined. The linearization approach is certainly useful, as it delivers tractable closed-form solutions, but it has some limitations. Spatial shocks are not necessarily small and their effects may amplify over time, especially in economies characterized by strong agglomeration forces. At the same time, recent advances in artificial intelligence applied to optimal control now make it possible to approximate solutions to dynamic models in high-dimensional settings (Carmona and Laurière (2021), Achdou et al. (2022b), Lavigne et al. (2025)). These methods are heuristic, so existence results are important to guarantee that the models being solved are internally consistent. There is therefore a need for theoretical work on dynamic spatial models, in the same way that Allen et al. (2020) provided the key existence and uniqueness results for static gravity-type frameworks.[1]

Certainly, the mathematical structure of dynamic spatial models is more complex, combining a non-linear integral equation for the static equilibrium, a backward Bellman equation characterizing individual optimization, and a forward law of motion for migration flows. When space is continuous, the problem becomes infinite-dimensional. This is why Moll (2025) argues that the standard assumption of rational expectations is conceptually implausible: real-world agents are unlikely to perform the complex forecasts implied by these models.

In this paper, we set aside this epistemological debate and remain within the standard paradigm. We prove existence of dynamic equilibria for a broad class of continuous-space trade models with endogenous forward-looking migrations and idiosyncratic random shocks. We keep the usual two-layer structure of dynamic spatial models: we assume that consumption adjustment is costless and that migration is costly, so that prices adjust instantaneously whereas migration evolves according to a dynamic adjustment process. We show that the individuals' problems can be seen as a **mean-field game** (MFG), and as such, that our model can be solved using the tools of this currently flourishing branch of game theory. The very reason of why these individuals' problems constitute an MFG is that the (instantaneous) indirect utility from consumption depends on the real wage at each location, which in turn depends on the whole distribution of agents across space. This is the so-called mean-field interaction term which makes dynamic spatial models solvable by the MFG methodology. While we use Krugman (1996) trade model as an illustrative example for the static layer, we show that our results extend to the broader family of spatial models with monopolistic competition (e.g., Helpman (1998) with local ownership, Forslid and Ottaviano (2003), Arkolakis et al. (2008), di Giovanni and Levchenko (2013), Redding and Rossi-Hansberg (2017), §3) as well as those with perfect competition (e.g., Armington (1969), Eaton and Kortum (2002), Alvarez and Lucas (2007), Allen and Arkolakis (2014), Caliendo and Parro (2015), Redding (2016) with constant returns to scale).

Our methodology could therefore be applied to the recent quantitative literature on forward-looking migrations, such as Caliendo et al. (2019) or Kleinman et al. (2023). The only difference with these models is that, in our setting, migration is local, whereas in theirs agents "jump" directly from one location to another. Bridging the two would require extending our framework to allow for non-local moves. MFGs with jumps are an active line of research (Bertucci (2020), Dumas and Santambrogio (2024), Dumas and Santambrogio (2025)). It is a highly relevant direction for future research.

MFGs have been introduced in mathematics about 20 years ago (see Huang et al. (2006), Lasry and Lions (2007)). They have been more recently applied in economics, especially in heterogeneous-agent macroeconomic models (Achdou et al. (2022a), Bilal (2023), Alvarez et al. (2023)). They are typically used in the case of large populations with certain characteristics' distributions: while no individual can affect the aggregate variables involved, the individual decisions do depend on some interaction terms typically capturing some key indicators/statistics of the distributions. In our case, real wages play the role of interaction terms. A further complexity coming from economic applications is the adjonction of general equilibrium conditions.[2] In our context, this means embedding the static spatial equilibrium (trade, urban,

[1]See also Allen and Arkolakis (2025) for a recent survey on the subject.

[2]Of course, the overwhelming part of MFG applications in areas like industrial organisation are in partial equilibrium and do not have this additional layer. See for example Aydin et al. (2025).

or regional) into the standard MFG equations, that typically Hamilton-Jacobi-Bellman (HJB) equation for individual dynamic decisions, and Fokker-Planck (FP) equation for the associated evolution of the dynamics of distributions involved.

This is precisely what makes MFG problems in economics mathematically appealing. To establish existence for the dynamic layer corresponding to the HJB–FP system, one must first have a solid mathematical grasp of the static layer. Not only must the static equilibrium be well defined (existence and uniqueness), but one also needs sufficient regularity properties to carry through the existence proof for the MFG. This is exactly what we achieve here. As an intermediate but central result, we prove existence, uniqueness, and regularity of solutions for an entire class of static spatial models in continuous space. This class emcompasses the monopolistic and perfect competition models mentioned above. While analogous results are well established for these models in discrete settings (see, e.g., Alvarez and Lucas (2007), Allen et al. (2020), Kucheryavyy et al. (2023)), they had not yet been obtained in continuous space.

Proving such results in continuous space matters for at least two reasons. First, continuous models naturally capture some specific spatial phenomena that are poorly approximated by discrete models; for instance, the dispersion of pollution at the urban scale is more naturally studied in continuous space. Second, establishing existence and uniqueness in the continuous case ensures that the problem remains well behaved as the number of regions N grows large, which is particularly relevant for applications with high-dimensional data (e.g., Bilal and Rossi-Hansberg (2023) with 3,000 U.S. counties, or Desmet et al. (2018) with a global resolution of $1^\circ \times 1^\circ$, corresponding to 64,800 cells).

Our proof is original, and we believe worth reading, especially for the uniqueness part. It proceeds by *homotopy*. The idea is simple: we first establish uniqueness in the most tractable benchmark case, when there are no trade frictions. In that setting, uniqueness follows immediately. We then gradually extend this result to the general case with arbitrary symmetric trade costs by continuously "deforming" the frictionless model into the general one. In other words, we track uniqueness step by step along a continuous path that connects the simple case to the full model. We use the Leray-Schauder degree theory to formalise this reasoning. We believe that the spirit of this homotopy approach could be applied more broadly to other types of equilibrium models beyond spatial economics.

Finally, we use the racetrack economy (circle) in the last section of the paper for a closer comparison with the huge existing literature on agglomeration/dispersion mechanisms using this spatial setting. Using the Herfindahl–Hirschman Index as a measure of spatial concentration, we show that the evolution of spatial agglomeration can be decomposed into four distinct forces, each corresponding to a specific term in the MFG equations: *idiosyncratic shocks*, captured by the Laplacian term of the Fokker–Planck equation, *myopic adjustment*, linked to the source term of the HJB equation, *uncertainty management*, associated with the Laplacian in the HJB equation, and *forward-looking expectations*, captured by the time derivative of the value function in the HJB equation. These forces are, of course, not aligned: some promote agglomeration, others drive dispersion. We can identify which forces drive agglomeration or dispersion depending on the static model. In particular, we show that, regardless of the static model, rational expectations always contribute to agglomeration, but cannot reverse the dominant pattern that would arise under myopic behavior. Our analysis shows that, in a dynamic and uncertain environment, agglomeration and dispersion cannot be reduced to static trade-offs such as love of variety versus congestion. Instead, they result from the interaction of forward-looking behavior, responses to uncertainty, and the diffusive effect of idiosyncratic shocks. The MFG framework makes these forces explicit and allows their relative contributions to be quantified.

The paper is organized as follows. Section 2 presents the model and studies both the static equilibrium and the MFG structure of the migration individual decisions. The (dynamic) MFG equilibrium is defined precisely and characterized. We prove existence and uniqueness for the static layer, and existence for the dynamic one. Section 3 is more applied: we use a racetrack version of the model to study its local dynamic properties by a standard local perturbation method. We single out four distinct mechanisms locally shaping agglomeration vs dispersion. Section 4 concludes.

2 Model

We consider a spatial economy extending along the torus $\mathbb{T}^2 = \mathbb{R}^2/\mathbb{Z}^2$, populated by a continuum of agents who choose their location over time in order to maximize their welfare. The model is structured in two layers. At each point in time, a static equilibrium determines local wages and prices based on the spatial distribution of agents. This static layer is a two-dimensional extension of the New Economic Geography model from Krugman (1996), without the agricultural sector. We choose this model to fix ideas: in Appendix B.1 and B.2, we show that our results can be extended to a much broader class of static spatial models. On top of this, agents engage in forward-looking migration decisions under uncertainty. Their dynamics can be described by a mean-field game, where individual incentives depend on future wages and the evolving spatial distribution of the population. The interaction between the static and dynamic layers is captured by the fact that the instantaneous utility function in the mean-field game is given by the real wage computed in static equilibrium.

2.1 Static equilibrium

The static layer of our model follows the core structure of the New Economic Geography introduced by Krugman (1996), but is adapted to a continuous two-dimensional space and excludes the agricultural sector. We abstract from agriculture to streamline the analysis: since the agricultural good serves mainly as an exogenous stabilizing force in the standard model, removing it allows us to focus only on the endogenous wage differentials and spatial interactions within the manufacturing sector. For the sake of completeness and clarity, we provide a full presentation of the static equilibrium derivation below, even though it is well known to readers familiar with the New Economic Geography literature. Readers familiar with these models may go directly to subsection 2.1.2.

2.1.1 The Krugman model

For simplicity, we omit explicit time dependence in the notation while describing the static layer. At each location $x \in \mathbb{T}^2$, the density of residents is denoted by $\mu(x)$, where μ is a probability measure on $\mathbb{T}^2$. Each location x produces $n(x) > 0$ differentiated varieties of a single manufactured good. A representative agent residing at x consumes over the full spectrum of varieties and has Dixit–Stiglitz preferences:

$$U(x) = \max_{q(\cdot,x,\cdot)} \left\{ \left(\int_{\mathbb{T}^2} \int_0^{n(y)} q(y,x,i)^{\frac{\sigma-1}{\sigma}} \, di \, dy \right)^{\frac{\sigma}{\sigma-1}} \right\}$$

subject to the budget constraint:

$$\int_{\mathbb{T}^2} \int_0^{n(y)} p(y,x,i) \, q(y,x,i) \, di \, dy = w(x)\mu(x),$$

where $q(y,x,i)$ and $p(y,x,i)$ denote the quantity and price of variety i produced at y and consumed at x, $w(x)$ is the nominal wage at x, and $\sigma > 1$ is the elasticity of substitution.

Solving the consumer's problem yields the demand function:

$$q(y,x,i) = p(y,x,i)^{-\sigma} P(x)^{\sigma-1} w(x)\mu(x), \tag{1}$$

where the local price index is:

$$P(x) = \left(\int_{\mathbb{T}^2} \int_0^{n(y)} p(y,x,i)^{1-\sigma} \, di \, dy \right)^{\frac{1}{1-\sigma}}. \tag{2}$$

The agent's indirect utility at location x is simply the real wage:

$$V(x,w) = \frac{w(x)}{P(x)}.$$

On the production side, we assume that there is a large number of manufacturing firms, each specializing in the production of a single variety. They are in monopolistic competition. Manufacturing firms operate under economies of scale, requiring both a fixed cost and a constant marginal cost. The labor demand for producing a quantity $Q(x,i)$ of variety i at location x follows:

$$L(x,i) = F + cQ(x,i)$$

where F represents the fixed labor requirement, and c is the marginal labor input per unit of output.

Given the isoelastic demand (1), firms set a price $p^0(y,i)$ at a constant markup over marginal cost, and the price received by the producer of variety i at y is:

$$p^0(y,i) = \frac{\sigma}{\sigma-1} c w(y) \tag{3}$$

Taking iceberg trade costs into account, the consumer price of a variety i produced at y and consumed at x is:

$$p(y,x,i) = p^0(y,i)\tau(y,x) = \frac{\sigma}{\sigma-1} c \, w(y) \, \tau(y,x),$$

where $\tau : \mathbb{T}^2 \times \mathbb{T}^2 \to [1,\infty)$ captures the intensity of trade frictions.

Zero-profit condition implies that firms produce a constant quantity:

$$Q_t(x,i) = \frac{F(\sigma-1)}{c}. \tag{4}$$

This is for the offer side. Let us derive the demand of each variety. Because all varieties are produced at the same scale, the density $n(y)$ of manufactured varieties at any location is directly proportional to the local worker density:

$$\mu(y) = \int_0^{n(y)} L(y,i)\,di = \alpha\sigma n(y) \tag{5}$$

Injecting (5) and (3) into (2), we get that

$$P(x) = \frac{c\sigma}{\sigma-1}(F\sigma)^{\frac{1}{\sigma-1}}\left(\int_{\mathbb{T}^2}\mu(y)w(y)^{1-\sigma}\tau(x,y)^{1-\sigma}dy\right)^{\frac{1}{1-\sigma}} \tag{6}$$

Integrating (1) and using (3) and (6), we obtain the demand for variety i produced at x as:

$$Q(x,i) = \int_{\mathbb{T}^2}\frac{(\frac{\sigma}{\sigma-1}c)^{-\sigma}w(x)^{-\sigma}\tau(x,y)^{1-\sigma}}{\frac{c\sigma}{\sigma-1}(F\sigma)\left(\int_{\mathbb{T}^2}\mu(z)w(z)^{1-\sigma}\tau(y,z)^{1-\sigma}dy\right)}w(y)\mu(y)dy \tag{7}$$

Equalizing offer (4) and demand (7), we get the following equilibrium equation on wage, for all x in $\mathbb{T}^2$:

$$w(x)^\sigma = \left(\frac{c}{F(\sigma-1)}\right)\left(\frac{\sigma-1}{c\sigma}\right)^\sigma\int_{\mathbb{T}^2}\left[\frac{\tau(y,x)^{1-\sigma}}{\int_{\mathbb{T}^2}\mu(z)(w(z)\tau(y,z))^{1-\sigma}dz}w(y)\mu(y)\right]dy.$$

Following Fujita et al. (1999), p. 55, it is possible to choose the units of measurement for output and the density of firms in order to have $\left(\frac{c}{F(\sigma-1)}\right)\left(\frac{\sigma-1}{c\sigma}\right)^\sigma = 1$.

2.1.2 Existence and uniqueness

We thus define a static equilibrium wage as follows.

Definition 1. *Fix any population distribution $\mu \in \mathscr{P}(\mathbb{T}^2)$ and trade frictions $\tau : \mathbb{T}^2 \times \mathbb{T}^2 \to [1,\infty)$. We say that $w : \mathbb{T}^2 \to \mathbb{R}_+$ is a static equilibrium wage if it satisfies*

$$w(x)^\sigma = \int_{\mathbb{T}^2}\frac{\tau(y,x)^{1-\sigma}w(y)}{\int_{\mathbb{T}^2}(\tau(y,z)w(z))^{1-\sigma}d\mu(z)}d\mu(y), \quad \forall x \in \mathbb{T}^2, \tag{8}$$

and

$$\int_{\mathbb{T}^2} w(x)d\mu(x) = 1. \tag{9}$$

Condition (9) normalizes the average level of wages in the economy. It serves to close the static model by fixing a common unit for nominal variables. Since equation (8) is homogeneous in w, this constraint prevents equilibrium indeterminacy arising from the fact that any scalar multiple of w would otherwise be a valid solution.

Definition 1 provides a continuous-space formulation of a static Krugman trade equilibrium with exogenous and immobile labor. While existence and uniqueness results are well established in the discrete case (Allen et al. (2020) being the most comprehensive study) such results have not been established in a continuous setting. In the continuous case, the proof is more challenging because we work in an infinite-dimensional space: the unit ball is no longer compact (Riesz's theorem), so additional conditions are needed to obtain a compact set in which a fixed-point theorem can be applied. The following result addresses this gap by proving existence and uniqueness of a static equilibrium wage in continuous space for any population distribution.

Theorem 1. *Fix any population distribution μ of $\mathscr{P}(\mathbb{T}^2)$. If τ is symmetric, of class C^1, and there exists a constant $C > 0$ such that*

$$\frac{1}{C} \leq \tau \leq C \text{ and } \|\tau'\|_\infty \leq C,$$

then there exists a static equilibrium wage map w. Moreover, there exists a constant C' only depending on τ such that

$$\frac{1}{C'} \leq w \leq C' \text{ and } \|w'\|_\infty \leq C'.$$

If, furthermore, supp $\mu = \mathbb{T}^2$, then the equilibrium is unique.

Proof. See Appendix A.2. □

Let us provide a brief sketch of the proof. For existence, it relies on a classical fixed-point argument. The uniqueness is more challenging to obtain. First, notice that without condition (9), uniqueness is hopeless: if w_0 satisfies (8), then so does tw_0 for any $t > 0$. Therefore, we reformulate the problem to work in Θ^0, a linear subspace of the space of continuous functions, which incorporates condition (9), in which uniqueness can be recovered. We introduce a function $G : \Theta^0 \to \Theta^0$ such that any static equilibrium w_0 can be mapped to a zero of G, called u_0. We can show that $DG(u_0)$ is an isomorphism of Θ^0 providing local uniqueness. In terms of global uniqueness, we observe that when $\tau \equiv 1$, there is a unique static equilibrium. We then use a homotopy argument to extend this result to any (symmetric) trade frictions. Fix any τ, and define $\tau_t = (1-t) + t\tau$. If $G : [0,1] \times \Xi^0 \to \Xi^0$ is such that the zeros of $G(t,\cdot)$ are exactly the static equilibria associated with τ_t, then $G(0,\cdot)$ has only one zero. We can show that no other zeros of $G(t,\cdot)$ appear as t grows to 1. We use the Leray-Schauder degree theory to formalize this reasoning.

Beyond existence and uniqueness, Theorem 1 also provides a regularity result: the static equilibrium wage is Lipschitz-continuous and bounded above and below by two positive constants that depend only on the trade cost function. Crucially, these bounds hold independently of the population distribution. To our knowledge, this is the first result proving such regularity for static spatial equilibria in a continuous setting. This property will play an important role in the analysis that follows, as it ensures that the wage function remains smooth enough to guarantee the existence of solutions to the dynamic problem.

2.1.3 Extensions

In Appendix B.1, we show that Theorem 1 extends to other monopolistic competition models (e.g., Helpman (1998) with local ownership, Forslid and Ottaviano (2003), Arkolakis et al. (2008), di Giovanni and

Levchenko (2013), Redding and Rossi-Hansberg (2017), §3) where the equilibrium equation takes the form:

$$w(x)^\sigma = \int_{\mathbb{T}^2} \frac{A(\mu,x)\tau(x,y)^{1-\sigma}}{\int_{\mathbb{T}^2} A(\mu,z)\tau(z,y)^{1-\sigma} w(z)^{1-\sigma}\, d\mu(z)} w(y)\, d\mu(y)$$

where the heterogeneity function $A(\mu,x)$ can be interpreted economically as capturing local productivity or amenities, potentially influenced by spillovers from the population distribution,[3]

In Appendix B.2, we extend Theorem 1 to perfect competition models (e.g., Armington (1969), Eaton and Kortum (2002), Alvarez and Lucas (2007), Allen and Arkolakis (2014), Redding (2016)), where the equilibrium equation takes the form:

$$w(x)^\sigma \mu(x) = \int_{\mathbb{T}^2} \frac{A(\mu,x)\tau(x,y)}{\int_{\mathbb{T}^2} A(\mu,z)\tau(z,y) w(z)^{1-\sigma}\, dz} w(y)\, d\mu(y).$$

2.2 Dynamic equilibrium

So far, the framework has been static. We now extend it to a dynamic setting as follows. We assume that agents are forward-looking and can migrate over time: they form expectations about future wages and spatial distributions, and choose their trajectory in order to maximize expected intertemporal utility.

2.2.1 A mean-field game problem

To formalize this, we work on a filtered probability space $(\Omega, \mathscr{F}, (\mathscr{F}_t)_{t\geq 0}, \mathbb{P})$ and fix a time horizon $T > 0$.[4] Each agent controls their own location, which evolves as a stochastic process. More precisely, the position X of a representative agent follows

$$\begin{cases} dX_t = \alpha_t\, dt + \sqrt{2\nu}\, dB_t, \quad 0 \leq t \leq T, \\ \mathscr{L}(X_0) = \mu_0, \end{cases} \tag{10}$$

where $\nu > 0$ is the diffusivity parameter, $(B_t)_{t\geq 0}$ is a standard Brownian motion, and $(\alpha_t)_{t\geq 0}$ is a control process representing the agent's velocity.

At any time t, given a trajectory of spatial wage profiles $(w(t))_{t\in[0,T]}$, the agent chooses their velocity so as to maximize their expected discounted utility:

$$\max_{(\alpha_t)_{t\geq 0}} \mathbb{E}\left[\int_0^T e^{-\rho t}\left[V(X_t, w(t), \mu(t)) - c(\alpha(t))\right] dt + g(X_T, w(T), \mu(T))\right] \tag{11}$$

[3] For instance, one may take

$$A(\mu,x) = a(x) \int_{\mathbb{T}^2} e^{-|x-y|^2}\, d\mu(y),$$

where the factor $a(x)$ models intrinsic spatial heterogeneity (e.g., natural advantages or baseline productivity), while the convolution term captures agglomeration spillovers from the surrounding population density.

[4] We show in Appendix B that our results generalize to infinite horizon.

where $V(x,w,\mu)$ is the real wage at location x, $c(\alpha) = \frac{c_0}{2}|\alpha|^2$ is a quadratic adjustment cost of migration, with $c_0 > 0$[5], and g is the terminal utility.

The function $V(x,w,\mu)$ describes the real wage available to a worker at location x, given a spatial wage profile w, and is defined by

$$V(x,w,\mu) = \frac{w(x)}{\left(\int_{\mathbb{T}^2} (w(z)\tau(x,z))^{1-\sigma} d\mu(z)\right)^{\frac{1}{1-\sigma}}},$$

where $\tau(x,z)$ is the trade cost function between locations x and z, and $\sigma > 1$ is the elasticity of substitution.

We now define a dynamic equilibrium as follows:

Definition 2. *We say that a flow of distributions $(\mu(t))_{t\in[0,T]} \in C([0,T],\mathscr{P}(\mathbb{T}^2))$ is a dynamic equilibrium if there exists a trajectory of wages $(w(t))_{t\in[0,T]}$, a stochastic process $(X_t)_{t\in[0,T]}$, and an optimal control $(\alpha_t)_{t\in[0,T]}$, such that:*

1. *X satisfies the stochastic differential equation (10);*
2. *The control α maximizes the agent's expected utility (11);*
3. *The law of X_t satisfies $\mathscr{L}(X_t) = \mu(t)$ for all $t \in [0,T]$;*
4. *At each time t, $w(t)$ is the wage profile determined by the static equilibrium (8)-(9) given the distribution $\mu(t)$.*

This dynamic equilibrium can be seen as a **mean-field game** (MFG). Agents choose their locations over time to maximize intertemporal real wages, net of migration costs. Wages, in turn, are determined by the overall distribution of agents via the static equilibrium conditions (8)–(9). Agents form expectations about how this distribution evolves, and at equilibrium, these expectations are correct, in line with the standard rational expectations assumption.

Interpreting the dynamic problem as a mean-field game will prove very fruitful. First, it allows us to reformulate the problem in terms of partial differential equations. Let

$$\begin{aligned} u(t,x) \quad = \quad & \max_{(\alpha_s)_{s\geq t}} \mathbb{E}\left\{ \int_t^T e^{-\rho(s-t)} \left[V(X_s,w(s),\mu(s)) - c_0 \frac{|\alpha_s|^2}{2} \right] ds \right. \\ & \left. + g(X_T,w(T),\mu(T)) \middle| X_t = x \right\} \end{aligned}$$

be the value function of the problem. If u admits enough regularity, it solves a Hamilton–Jacobi–Bellman (HJB) equation reflecting the trade-off between current costs and expected future gains:

$$\begin{cases} -\partial_t u - \nu\Delta u + \rho u = H(x,\nabla u,w(t),\mu(t)), & (t,x) \in [0,T) \times \mathbb{T}^2 \\ u(T,x) = g(x,w(T),\mu(T)), & x \in \mathbb{T}^2 \end{cases}$$

[5] In the theoretical literature, local migration with quadratic adjustment costs has long been the standard (see, e.g., Baldwin and Venables (1994), Mossay (2006), Mossay (2013)). More recent quantitative models, by contrast, describe migration as occurring in discrete jumps. A natural extension of the model would be to allow for non-local moves, where agents can relocate directly across distant locations, in line with recent quantitative approaches.

where, for all $(x,p,w,\mu)\in\mathbb{T}^2\times\mathbb{R}^2\times C(\mathbb{T}^2,\mathbb{R}_+^*)\times\mathscr{P}(\mathbb{T}^2)$, the Hamiltonian has an explicit form:

$$H(x,p,w,\mu)=\max_{\alpha\in\mathbb{R}^2}\left\{p\cdot\alpha-c_0|\alpha|^2/2+V(x,w,\mu)\right\}=\frac{|p|^2}{2c_0}.$$

Therefore, the optimal control is

$$\alpha^*(t,x)=\operatorname*{argmax}_{\alpha\in\mathbb{R}^2}\left\{\nabla u(t,x)\cdot\alpha-c_0\frac{|\alpha|^2}{2}\right\}=\frac{1}{c_0}\nabla u(t,x). \tag{12}$$

This expression defines a closed-loop policy: the optimal migration speed at time t depends only on the agent's current location x and on the anticipated wage trajectory w trough u. Given this optimal policy, HJB writes

$$\begin{cases}-\partial_t u-\nu\Delta u-\frac{|\nabla u|^2}{2c_0}+\rho u=V(x,w(t),\mu(t)), & (t,x)\in[0,T)\times\mathbb{T}^2\\ u(T,x)=g(x,w(T),\mu(T)), & x\in\mathbb{T}^2.\end{cases} \tag{13}$$

Now, the aggregate distribution μ evolves according to a Fokker–Planck (FP) equation, which describes how the mass of agents flows across space under optimal migration:

$$\begin{cases}\partial_t\mu-\nu\Delta\mu+\frac{1}{c_0}\operatorname{div}(\mu\nabla u)=0, & (t,x)\in(0,T]\times\mathbb{T}^2,\\ \mu(0,x)=\mu_0, & x\in\mathbb{T}^2.\end{cases}$$

We now reformulate the notion of dynamic equilibrium accordingly.

Definition 3. *We say that a flow of distributions $\mu\in C^{1,2}([0,T]\times\mathbb{T}^2,\mathbb{R})$ is a mean-field game equilibrium if there exists a value function $u\in C^{1,2}([0,T]\times\mathbb{T}^2,\mathbb{R})$ and a trajectory of wages $w\in C([0,T]\times\mathbb{T}^2,\mathbb{R})$ such that the following equations are satisfied in the classical sense:*

$$-\partial_t u-\nu\Delta u-\frac{|\nabla u|^2}{2c_0}+\rho u = V(x,w(t),\mu(t)), \tag{14}$$

$$\partial_t\mu-\nu\Delta\mu+\frac{1}{c_0}\operatorname{div}(\mu\nabla u) = 0, \tag{15}$$

$$\int_{\mathbb{T}^2}\frac{\tau(y,x)^{1-\sigma}w(t,y)}{\int_{\mathbb{T}^2}(w(t,z)\tau(y,z))^{1-\sigma}\mu(t,z)dz}\mu(t,y)dy = w(t,x)^\sigma, \tag{16}$$

completed with $\mu(0)=\mu_0$, $u(T)=g(x,w(T),\mu(T))$, and such that for any t

$$\int\mu(t,x)dx=1 \text{ and } \int w(t,x)d\mu(t,x)=1.$$

Equations (14) and (15) correspond to the standard HJB–FP system in mean-field games. The general equilibrium condition (16) comes from the static layer and endogenously determines the spatial wage profile based on the population distribution.

Remark 1. In Definition 3, the solution of the HJB equation is assumed to be regular. When such a solution exists, Definitions 2 and 3 are equivalent, since the HJB solution coincides with the value function of the optimal control problem (by a verification theorem). Hence, we focus on solving the MFG system (14)–(16) and look for a mean-field equilibrium in the sense of Definition 3.

2.2.2 Existence of equilibria

Most forward-looking migration models, without relying on the MFG formalism, arrive essentially at this same triplet of equations. Quantitative works, such as Caliendo et al. (2019) or Kleinman et al. (2023), are usually formulated in discrete space. The main continuous-space framework with forward-looking migration is Mossay (2013), who develops a version of Krugman's trade model on a circle. To the best of our knowledge, whether in discrete or continuous space, no work has yet demonstrated the existence of a dynamic equilibrium. In the following theorem, we adapt techniques from MFG theory to prove such a result.

Theorem 2. *Assume that τ is symmetric, of class C^1, and there exists a constant $C > 0$ such that*

$$\frac{1}{C} \leq \tau \leq C \text{ and } \|\tau'\|_\infty \leq C,$$

and that g is continuous, continuously differentiable with respect to its first variable, and that

$$\sup_{w,\mu} \|\nabla_x g(\cdot, w, \mu)\|_\infty < +\infty.$$

Then there exists at least one mean-field game equilibrium, in the sense of Definition 3.

Proof. See Appendix A.3. □

The proof of Theorem 2 relies on a fixed-point argument. As mentioned earlier, the particularity of the MFG system (14)-(16) is the additional equilibrium equation (16). To apply Schauder's fixed-point theorem, it is therefore important that the solutions of this equation are regular and unique. We use for this purpose the results obtained in Theorem 1.

2.2.3 Extensions

In Appendix B.1 and B.2, we show that Theorem 2 extends to the case where equation (15) is replaced by other monopolistic or perfect competition models. In Appendix B.3, we show that our result continues to hold in the infinite-horizon setting.

The arguments developed in the proof of Theorem 2 can also be adapted to establish the existence of a steady state equilibrium.

Definition 4. *We say that a distribution $\mu \in C^2(\mathbb{T}^2, \mathbb{R})$ is a stationary mean-field game equilibrium if there exists a value function $u \in C^2(\mathbb{T}^2, \mathbb{R})$ and a wage function $w \in C(\mathbb{T}^2, \mathbb{R})$ such that the following equations* *are satisfied in the classical sense:*

$$-\nu\Delta u - \frac{|\nabla u|^2}{2c_0} + \rho u = V(x,w,\mu), \tag{17}$$

$$-\nu\Delta\mu + \frac{1}{c_0}\operatorname{div}(\mu\nabla u) = 0, \tag{18}$$

$$\int_{\mathbb{T}^2} \frac{\tau(y,x)^{1-\sigma} w(t,y)}{\int_{\mathbb{T}^2}(w(t,z)\tau(y,z))^{1-\sigma}\mu(z)dz}\mu(y)dy = w(x)^{\sigma}, \tag{19}$$

completed with

$$\int \mu(x)dx = 1 \text{ and } \int w(x)d\mu(x) = 1.$$

We get the following existence result.

Theorem 3. *Under the assumptions of Theorem 2, there exists at least one stationary mean-field game equilibrium, in the sense of Definition 4.*

Proof. See Appendix A.4. □

2.2.4 About uniqueness

Uniqueness of the equilibrium does not hold in full generality. It depends crucially on the behavior of the utility term $V(x,w,\mu)$. We briefly discuss this issue in this paragraph, distinguishing two main cases.

First, suppose that agglomeration forces dominate in the static economy. This is typically the case for monopolistic competition models without immobile factors, such as Krugman (1996) considered here, or other variants listed in Appendix B.1. In this case, if space is isotropic, the economy may concentrate at any point. Uniqueness of the steady state therefore fails to hold. This is the sense of the following proposition.

Proposition 1. *Under the assumptions of Theorem 2, for any $x^0 \in \mathbb{T}^2$, the Dirac mass δ_{x^0} characterizes a solution of the deterministic steady state problem, i.e. it characterizes a solution of the MFG system* (17)-(19) *with $\nu = 0$.*

Proof. See Appendix A.5. □

Now, suppose instead that dispersion forces dominate in the static economy. This is typically the case for perfect competition models listed in Appendix B.2. In this situation, agents cannot concentrate in a single location. Intuitively, this would eliminate the multiplicity of equilibria driven by concentration forces, and one might therefore expect the equilibrium to be unique. From the mean-field game perspective, this situation is captured by the *Lasry–Lions monotonicity condition*. Loosely speaking, this condition requires that, on average, the payoff of being at a given location decreases as the density of agents at that location increases.

Proposition 2. *Assume that the Lasry-Lions monotonicity condition holds, i.e.*

$$\int_{\mathbb{T}^d} \left(V(x,w_{\mu_1},\mu_1) - V(x,w_{\mu_2},\mu_2)\right) d(\mu_1-\mu_2)(x) \geq 0, \qquad \forall \mu_1,\mu_2 \in \mathscr{P}(\mathbb{T}^d).$$

Then, under the assumptions of Theorem 2, there is a unique mean-field game equilibrium.

Proof. See Lasry and Lions (2007), Theorem 2.5. □

It is, however, difficult to establish that the Lasry–Lions condition holds for all perfect competition models, even though their well-known dispersive behavior would suggest that it might. One case where the verification is straightforward is in the absence of trade frictions. In this case, uniqueness obtains: with a perfect competition model with no trade frictions and no heterogeneity, the mean-field game equilibrium is unique.

Corollary 1. *Assume that the static equilibrium is given by a perfect competition model from Appendix B.2 with no trade frictions and no heterogeneity, i.e.* $\tau = 1$ *and* $A = 1$*. Then, under the assumptions of Theorem 2, there is a unique mean-field game equilibrium.*

3 The racetrack economy

We now consider a one-dimensional application of the model. The geography, known as the racetrack economy, is represented by the circle $\mathbb{T}^1_R$ of radius $R > 0$. This framework, widely used in economic geography (see, e.g., Krugman (1996) Fujita et al. (1999)), provides an analytically convenient benchmark to study the formation of spatial patterns.

Within this setting, the mean-field game structure proves particularly fruitful to analyze the agglomeration and dispersion forces in a dynamic and uncertain environment. Indeed, HJB and FP equations explicitly capture the forward-looking behavior of agents and their exposure to idiosyncratic shocks. This allows us to formalize and quantify how expectations and uncertainty affect the evolution of population distribution in space.

The remainder of this section introduces a linearized version of the model around a spatially uniform steady state. This simplified framework admits closed-form sinusoidal solutions, which allow us to identify and interpret the main forces shaping the dynamics of spatial agglomeration: idiosyncratic shocks, myopic adjustment, forward-looking expectations, and the management of uncertainty.

3.1 Equilibrium

We begin by defining a dynamic equilibrium in the racetrack economy, following the same definition as in the two-dimensional case.

Definition 5. *In the racetrack economy,* $\mu \in C^{1,2}([0,T] \times \mathbb{T}^1_R)$ *is a dynamic equilibrium if there exists a trajectory of static equilibrium wages* $w \in C([0,T] \times \mathbb{T}^1_R)$ *and* $u \in C^{1,2}([0,T] \times \mathbb{T}^1_R)$ *such that:*

$$\begin{aligned}
-\partial_t u - \nu \Delta u - \frac{|\nabla u|^2}{2c_0} + \rho u &= V(x, w(t), \mu(t)), \\
\partial_t \mu - \nu \Delta \mu + \frac{1}{c_0} \operatorname{div}(\mu \nabla u) &= 0, \\
\int_{\mathbb{T}^2} \frac{\tau(y,x)^{1-\sigma} w(t,y)}{\int_{\mathbb{T}^2} (w(t,z)\tau(y,z))^{1-\sigma} \mu(t,z) dz} \mu(t,y) dy &= w(t,x)^{\sigma},
\end{aligned}$$

completed with $\mu(0,\cdot)=\mu_0,\ u(T,\cdot)=0$, *and such that for any* $t\geq 0$*:*

$$\int_{\mathbb{T}^1_R}\mu(t,x)dx = 2\pi R \text{ and } \int_{\mathbb{T}^1_R} w(t,x)\mu(t,x)dx = 2\pi R.$$

Under the assumptions of Theorem 2, such a dynamic equilibrium exists for any initial condition $\mu_0\in 2\pi R\cdot\mathscr{P}(\mathbb{T}^1_R)$.[6] In particular, the system admits a unique flat-earth solution given by

$$\bar{\mu}=1,\ \bar{u}=-\left[2\rho R\int_0^{\pi} e^{-d(\sigma-1)x}dx\right]^{-1},\ \bar{w}=1,\ \bar{V}=\left[2R\int_0^{\pi} e^{-d(\sigma-1)x}dx\right]^{-1}. \tag{20}$$

To analyze the mechanisms of agglomeration and dispersion, we consider a small deviation from this stationary equilibrium. Specifically, following Krugman (1996), we introduce a sinusoidal perturbation of the initial population distribution:

$$\tilde{\mu}_0(x)=\delta_\mu\cos(kx) \tag{21}$$

with $k\in\frac{1}{2\pi R}\mathbb{N}^*$ and $\delta_\mu>0$ small. This perturbation affects the initial wage equilibrium through the static equation, generating a corresponding perturbation $\tilde{w}[\tilde{\mu}_0]\in L^2(\mathbb{T}^1_R)$ on the nominal wage, and $\tilde{V}[\tilde{\mu}_0]\in L^2(\mathbb{T}^1_R)$ on the real wage. Following Fujita et al. (1999), these perturbations can be characterized analytically using a first-order approximation.

Proposition 3. *For all* $x\in\mathbb{T}^1_R$*, we have, at first-order approximation:*

$$\tilde{V}[\tilde{\mu}_0]=\frac{\delta_V}{\delta_\mu}\tilde{\mu}_0, \tag{22}$$

where

$$\frac{\delta_V}{\delta_\mu}=G^{-1}\frac{1-s^2}{s}\frac{1-Z}{1-Z(1-s)-sZ^2}$$

$$G=\left[R\int_0^{\pi}e^{-d(\sigma-1)z}dz\right]^{1/1-\sigma}$$

$$Z=RG^{\sigma-1}\int_{-\pi}^{\pi}\cos(kz)e^{-d(\sigma-1)|z|}dz.$$

and $s:=(\sigma-1)^{-1}$*. The ratio* (δ_V/δ_μ) *is positive when R is large.*

Proof. See Appendix A.6. □

Proposition 3 states that, at first order, the sinusoidal perturbation (21) of the initial population distribution induces a corresponding sinusoidal perturbation (22) of the initial utility. When the radius R is large, the ratio δ_V/δ_μ is positive, indicating that utility increases at locations where the local population density rises. This property stems from the fact that our static model is a particular case of Krugman (1996), where only the manufacturing sector is active. In our setting, there are no centrifugal forces arising from immobile agriculture, but only from love of variety. Consequently, the "no-black-hole" condition (Fujita et al., 1999,

[6] Following Fujita et al. (1999), in this section we assume that $\int_{\mathbb{T}^1_R}\mu(t,x)dx=2\pi R$ in order to ensure that $\bar{\mu}=1$ is a steady state.

p. 59) does not hold here[7] and, with myopic agents, any small perturbation of population density would evolve toward agglomeration.

Thus, the Krugman model without the agricultural sector leads to agglomeration as migration decisions are not forward-looking. This result is well known. However, our model is dynamic and stochastic: agents form rational expectations and face uncertainty. Additional forces related to these features therefore come into play. We will see that the MFG equations make it possible to identify and quantify these forces. But before doing so, to facilitate the analysis, let us linearize the MFG system around the stationary equilibrium.

Definition 6. *In the racetrack economy, we say that $\tilde{\mu} \in C^{1,2}([0,T] \times \mathbb{T}^1_R)$ is a solution to the linearized perturbation problem if there exists a trajectory of value function perturbations $\tilde{u} \in C^{1,2}([0,T] \times \mathbb{T}^1_R)$ such that:*

$$-\partial_t \tilde{u} - \nu \Delta \tilde{u} + \rho \tilde{u} = \tilde{V}[\tilde{\mu}_t], \tag{23}$$

$$\partial_t \tilde{\mu} - \nu \Delta \tilde{\mu} + \frac{1}{c_0} \Delta \tilde{u} = 0, \tag{24}$$

$$\tilde{V}[\tilde{\mu}_t] = \frac{\delta_V}{\delta_\mu} \tilde{\mu}_t, \tag{25}$$

completed with $\tilde{\mu}(0,x) = \delta_\mu \cos(kx)$, $\tilde{u}(T,x) = 0$ for all $x \in \mathbb{T}^1_R$.

This type of linearization has proven useful in economic applications of MFGs, particularly for analyzing shocks in heterogeneous-agent macroeconomic models (see Bilal (2023), Alvarez et al. (2023)). In our case, it is justified by the fact that the initial perturbation $\tilde{\mu}_0$ is small.

Our linearized system (23)-(25) may admit multiple solutions. Following Fujita et al. (1999), we focus on sinusoidal solutions that oscillate at the same frequency as the initial perturbation. We characterize this class of solutions in closed form in the following proposition.

Proposition 4. *The linearized system of Definition 6 admits a solution of the form*

$$\tilde{\mu}(t,x) = A(t)\cos(kx), \tag{26}$$

$$\tilde{u}(t,x) = B(t)\cos(kx), \tag{27}$$

if and only if

$$\frac{\delta_V}{\delta_\mu} < \frac{c_0 \rho^2}{4k^2} + c_0 \nu^2 k^2 + c_0 \rho \nu.$$

The amplitudes $A(t)$ and $B(t)$ are linear combinations of exponentials:

$$A(t) = C_1 e^{\lambda_1 t} + C_2 e^{\lambda_2 t},$$
$$B(t) = (C_1/\beta_1) e^{\lambda_1 t} + (C_2/\beta_2) e^{\lambda_2 t},$$

[7] Specifically, since the share of expenditure on manufactured goods m is equal to one, we have $\sigma(1-m) = 0$, whereas the no-black-hole condition requires $\sigma(1-m) > 1$.

with

$$\lambda_1 = \frac{\rho + \sqrt{\rho^2 + 4\left(\nu^2 k^4 + \rho \nu k^2 - \frac{\delta_V k^2}{\delta_\mu c_0}\right)}}{2}, \qquad \lambda_2 = \frac{\rho - \sqrt{\rho^2 + 4\left(\nu^2 k^4 + \rho \nu k^2 - \frac{\delta_V k^2}{\delta_\mu c_0}\right)}}{2},$$

$$\beta_1 = \frac{k^2}{c_0(\nu k^2 + \lambda_1)}, \qquad \beta_2 = \frac{k^2}{c_0(\nu k^2 + \lambda_2)},$$

$$C_1 = \frac{-\delta_\mu \beta_1 e^{\lambda_2 T}}{\beta_2 e^{\lambda_1 T} - \beta_1 e^{\lambda_2 T}}, \quad C_2 = \frac{\delta_\mu \beta_2 e^{\lambda_1 T}}{\beta_2 e^{\lambda_1 T} - \beta_1 e^{\lambda_2 T}}.$$

Proof. See Appendix A.7. □

3.2 Analysis of agglomeration and dispersion forces

The linearized dynamics derived in Proposition 4 provide a tractable way to study the forces that drive agglomeration or dispersion. The amplitudes of $\tilde{u}$ and $\tilde{\mu}$ take the form of a sum of two exponentials. Several cases arise depending on the values taken by the parameters.

If

$$\frac{\delta_V}{\delta_\mu} < c_0 k^2 \nu^2 + c_0 \rho \nu,$$

one exponential, $e^{\lambda_1 t}$, diverges because $\lambda_1 > 0$, while the other, $e^{\lambda_2 t}$, vanishes because $\lambda_2 < 0$. Besides, over an infinite horizon,

$$C_1 \underset{T\to\infty}{\longrightarrow} 0, \quad C_2 \underset{T\to\infty}{\longrightarrow} \delta_\mu.$$

This means that when the time horizon is large, the divergent exponential disappears from the solution. In this case, the perturbation $\tilde{\mu}$ dampens over time, and eventually vanishes. There is dispersion.

Else, if

$$c_0 k^2 \nu^2 + c_0 \rho \nu < \frac{\delta_V}{\delta_\mu} < \frac{c_0 \rho^2}{4k^2} + c_0 k^2 \nu^2 + c_0 \rho \nu,$$

both exponentials $e^{\lambda_1 t}$ and $e^{\lambda_2 t}$ diverge, because $\lambda_1, \lambda_2 > 0$. Over an infinite horizon,

$$C_1 \underset{T\to\infty}{\longrightarrow} 0, \quad C_2 \underset{T\to\infty}{\longrightarrow} \delta_\mu.$$

In this case, the perturbation $\tilde{\mu}$ grows over time, leading to agglomeration.

We summarize these results for large horizon in the following corollary.

Corollary 2. *Assume that* $\frac{\delta_V}{\delta_\mu} < \frac{c_0 \rho^2}{4k^2} + c_0 k^2 \nu^2 + c_0 \rho \nu$. *When the time horizon* T *is large, there exists a solution to the linearized problem such that for all* $0 \le t \le T$, $x \in \mathbb{T}^1_R$,

$$\tilde{\mu}(t,x) = \delta_\mu e^{\lambda t} \cos(kx), \tag{28}$$

with $\lambda = \frac{1}{2}\left[\rho - \sqrt{\rho^2 + 4\left(\nu^2 k^4 + \rho \nu k^2 - \frac{\delta_V k^2}{\delta_\mu c_0}\right)}\right]$. *If*

$$\frac{\delta_V}{\delta_\mu} < c_0 k^2 \nu^2 + c_0 \rho \nu,$$

the initial perturbation vanishes over time: there is dispersion. Else, if

$$c_0 k^2 \nu^2 + c_0 \rho \nu < \frac{\delta_V}{\delta_\mu} < \frac{c_0 \rho^2}{4k^2} + c_0 k^2 \nu^2 + c_0 \rho \nu,$$

the initial perturbation amplifies over time: there is agglomeration.

Proof. See Appendix A.8. □

We now want to go more in detail and quantify the extent to which strategic and dynamic behaviors (rational expectations, consideration of uncertainty) contribute to agglomeration and dispersion effects. To quantify them, we use the Herfindahl–Hirschman Index (HHI), a classical measure of concentration. For any probability density function f defined on $\mathbb{T}^1_R$, we define the HHI of f as

$$H[f] = \int_{\mathbb{T}^1_R} f(x)^2 \, dx.$$

The next proposition illustrates how the HHI captures the degree of spatial concentration:

Proposition 5. *The following statements hold:*

- *The uniform distribution minimizes the HHI over* $\mathscr{P}(\mathbb{T}^1_R)$.
- *Any sequence* $(f_k)_{k\in\mathbb{N}} \subset \mathscr{P}(\mathbb{T}^1_R)$ *that converges (in the sense of distributions) to a Dirac mass* δ_{x_0} *at some point* $x_0 \in \mathbb{T}^1_R$ *satisfies:*

$$\lim_{k\to\infty} H[f_k] = +\infty.$$

Proof. See Appendix A.9. □

The HHI thus provides a continuous measure of spatial concentration: it reaches its minimum for a uniform distribution and diverges as mass becomes increasingly concentrated at a point. We will use this metric to evaluate how our racetrack economy evolves from an initial perturbation and to identify the forces that drive agglomeration and dispersion.

Given the perturbation setup introduced earlier, the HHI at time t takes the form[8]

$$H[\mu](t) = \int_{\mathbb{T}^1_R} (\bar{\mu} + \tilde{\mu}(t,x))^2 \, dx = \int_{\mathbb{T}^1_R} (1 + \tilde{\mu}(t,x))^2 \, dx.$$

Using the fact that $\int_{\mathbb{T}^1_R} \tilde{\mu}(t,x)\, dx = 0$, a second-order expansion yields:

$$H[\mu](t) = \int_{\mathbb{T}^1_R} \tilde{\mu}(t,x)^2 \, dx.$$

[8] For readability, we slightly abuse notation by omitting the normalization constant $(2\pi R)^{-1}$ in front of μ.

To assess whether the system tends toward agglomeration or dispersion, we consider the time derivative of the HHI:

$$\frac{1}{2}H[\mu]'(t) = \int_{\mathbb{T}_R^1} \partial_t \tilde{\mu}(t,x)\, \tilde{\mu}(t,x)\, dx.$$

A positive value indicates increasing spatial concentration (agglomeration), whereas a negative value indicates movement toward uniformity (dispersion). Using the linearized Fokker–Planck equation (24), an integration by parts and equation (12), this derivative decomposes as:

$$\frac{1}{2}H[\mu]'(t) = \underbrace{\nu \int \Delta\tilde{\mu}(t,x)\tilde{\mu}(t,x)\, dx}_{\text{Idiosyncratic shocks}} + \underbrace{\int \alpha^*(t,x)\cdot\nabla\tilde{\mu}(t,x)\, dx}_{\text{Controlled migrations}} \tag{29}$$

The interpretation of equation (29) is as follows. Agglomeration and dispersion dynamics are driven by agent movement, which can be either involuntary due to exogenous shocks or controlled by the agents through migrations.

The involuntary term related to idiosyncratic shocks, $\nu \int \tilde{\mu}(t,x)\Delta\tilde{\mu}(t,x)\, dx$, is always negative[9] and therefore systematically contributes to dispersion. This reflects the fact that idiosyncratic shocks introduce random, uncoordinated movements across space. As agents are hit by independent Brownian shocks, their trajectories diverge over time, leading to a smoothing of the population density. This diffusive mechanism is purely entropic: it pushes the system toward a more uniform distribution, regardless of economic incentives or spatial preferences.

The controlled migration term depends on the interplay between the optimal control α^* and the population distribution $\tilde{\mu}$. If $\int \alpha^*\cdot\nabla\tilde{\mu} > 0$, agents tend to move toward local maxima of $\tilde{\mu}$, leading to agglomeration. Conversely, if $\int \alpha^*\cdot\nabla\tilde{\mu} < 0$, agents tend to move toward local minima of $\tilde{\mu}$, leading to dispersion.

By performing a double integration by parts on the controlled migration term, and then using the linearized HJB equation (23), this term can be decomposed into several interpretable components. This is shown in the following proposition.

Proposition 6. *Let $\tilde{\mu}$ and $\tilde{u}$ be solutions to the linearized problem as in Definition 6. For all $0 \le t \le T$:*

$$\begin{aligned} H[\mu]'(t) = &\underbrace{\nu \int \Delta\tilde{\mu}(t,x)\tilde{\mu}(t,x)\, dx}_{\textit{Idiosyncratic shocks}} \underbrace{- \frac{1}{\rho c_0}\int \tilde{V}[\tilde{\mu}_t](x)\Delta\tilde{\mu}(t,x)\, dx}_{\textit{Myopic adjustment}} \\ &\underbrace{- \frac{\nu}{\rho c_0}\int \Delta\tilde{u}(t,x)\Delta\tilde{\mu}(t,x)\, dx}_{\textit{Uncertainty}} \underbrace{- \frac{1}{\rho c_0}\int \partial_t\tilde{u}(t,x)\Delta\tilde{\mu}(t,x)\, dx}_{\textit{Forward-looking expectations}}. \end{aligned}$$

Proof. See Appendix A.10. □

Proposition 6 identifies four distinct forces that shape the evolution of agglomeration and dispersion. The first term originates from the Fokker–Planck equation and captures the contribution of idiosyncratic shocks. As discussed previously, it is always negative and thus acts as a force of dispersion.

[9] This follows from an integration by parts, using periodic boundary conditions.

The remaining three terms arise from the linearized Hamilton–Jacobi–Bellman equation and reflect how agents adjust their location based on different economic motives.

The myopic adjustment term originates from the instantaneous utility term $\tilde{V}_{\tilde{\mu}_t}(x)$ in the HJB equation and captures myopic migration behavior. Agents react to current spatial differences in utility, relocating toward areas with higher real wages.

The uncertainty term comes from the Laplacian $\nu\Delta\tilde{u}$ in the HJB equation. It reflects the anticipation of random shocks.

Finally, the forward-looking expectation term comes from the time derivative $\partial_t\tilde{u}$ in the HJB equation. It captures the forward-looking component of agents' behavior: individuals form expectations about how the value of each location will evolve, and adjust their migration decisions accordingly.

These forces are not always aligned. Some promote agglomeration, others drive dispersion. To understand their respective roles more precisely, we restrict attention to the class of sinusoidal solutions introduced in Proposition 4 and Corollary 2.

Proposition 7. *Let $\tilde{\mu}$ be the sinusoidal perturbation solution to the linearized problem given in Corollary 2. We have, for all $0 \le t \le T$:*

$$H[\mu]'(t) = e^{-2\lambda t}k^2\delta_\mu^2\left[\underbrace{-\nu}_{\textit{Idiosyncratic shocks}} \quad \underbrace{+(\rho c_0)^{-1}\frac{\delta_V}{\delta_\mu}}_{\textit{Myopic adjustment}} \quad \underbrace{-(\rho c_0)^{-1}\nu k^2\beta_2^{-1}}_{\textit{Uncertainty}} \quad \underbrace{+(\rho c_0)^{-1}\lambda\beta_2^{-1}}_{\textit{Forward-looking expectations}}\right]$$

Proof. See Appendix A.11. □

In this form, the sign of each term is explicit. As expected, the contribution of idiosyncratic shocks is negative, confirming that such random shocks induce dispersion. This reflects the diffusive effect discussed earlier, whereby uncoordinated movements tend to flatten the spatial distribution.

The term related to myopic adjustment has the same sign as the ratio δ_V/δ_μ. According to Proposition 3, this ratio is positive in our model. This means that locations with higher population density also offer higher instantaneous utility. As a result, agents are attracted to already dense regions, reinforcing concentration. This mechanism thus promotes agglomeration, and the corresponding contribution to $H[\mu]'(t)$ is positive.

Conversely, the term related to uncertainty is negative (because $\beta_2 > 0$ when $\delta_V/\delta_\mu > 0$). The intuition is as follows. Noise naturally tends to spread the population by reducing density in crowded areas and increasing it elsewhere. Since δ_V/δ_μ is positive, utility is currently higher in densely populated locations. However, agents anticipate that stochastic fluctuations will gradually redistribute the population, making less crowded areas more attractive in the future. Expecting these zones to gain in utility over time, agents choose to move toward them preemptively. This behavior, driven by uncertainty, reinforces spatial dispersion, and the corresponding contribution to $H[\mu]'(t)$ is negative.

Finally, the term associated with forward-looking expectations is positive (because $\beta_2 > 0$ and $\lambda > 0$ when $\delta_V/\delta_\mu > 0$). The intuition behind this is the following: since the ratio δ_V/δ_μ is positive, being in a densely populated area increases instantaneous utility. Agents therefore anticipate that others will also be drawn to these crowded locations, causing future utility to rise in those places. This expectation creates an incentive for agents to join the crowd immediately, reinforcing agglomeration dynamics.

We thus see that when $\delta_V/\delta_\mu > 0$ (that is, when agglomeration forces dominate in the static framework), forward-looking expectations further reinforce agglomeration. Let us now consider the opposite case: suppose that the static economy is dominated by dispersion forces, that is, $\delta_V/\delta_\mu < 0$. In this case, it is easy to show that $\beta_2 < 0$ and $\lambda < 0$. Hence, the term associated with rational expectations remains positive: rational expectations still act as an agglomeration force. The intuition is as follows. When the ratio δ_V/δ_μ is negative, being in a densely populated area reduces instantaneous utility. Agents therefore anticipate that such crowded areas will eventually empty out, which will increase utility in those locations in the future. This anticipation creates an incentive for agents to join the crowd immediately, once again reinforcing agglomeration.

Thus, whatever the underlying trade model (whether driven by agglomeration or dispersion forces), forward-looking expectations always reinforce agglomeration. This raises a natural question: could rational expectations ever *reverse* the outcome, leading to agglomeration when myopic behavior would have produced dispersion? For instance, if agents expect a crowded city to lose residents and move there in anticipation of others leaving, could this paradoxically cause further agglomeration? The answer is no. One can show that the combined effect of myopic and expectation terms in Proposition 7 always shares the same sign as the myopic term alone. In other words, rational expectations cannot reverse the dominant pattern (agglomeration or dispersion) that prevails under myopic behavior.

All this section shows that, in a dynamic and uncertain environment, agglomeration and dispersion cannot be understood only through static trade-offs such as love of variety versus congestion. They also reflect how agents form expectations, respond to uncertainty, and adjust to the evolving distribution of others. The MFG framework allows these effects to be modeled and quantified explicitly.

4 Conclusion

In this paper, we have presented why and how the rising MFG methodology can be used to overcome the technical difficulties arising in spatial dynamics in continuous space. By providing such a systematic solution strategy, this quickly developing methodology is likely to help handling more intricate spatial models in the future. As shown in this paper, it may also open the door to a deeper and more efficient theoretical approach possibly delivering existence proofs of equilibria in various circumstances.

Several extensions can already be envisaged regarding the state of art in MFG. First, there is room to introduce policy and even interactions between the policy maker (typically referred to as dominating player) and the economic agents (or market players) as evidenced in Bensoussan et al. (2016). Second, one can make the stochastic component more complex if needed for realism. For example, one can add common noise to idiosyncratic noises (see Carmona et al. (2015)). Last but not least, the MFG computational literature is booming and in principle, it should be possible to obtain the transitional dynamics inherent in MFG games with increasingly efficient computational methods (see Bayraktar and Zhang (2023) for instance).

References

Achdou, Y., Han, J., Lasry, J.-M., Lions, P.-L., and Moll, B. (2022a). Income and wealth distribution in macroeconomics: A continuous-time approach. Review of Economic Studies, 89:45–86.

Achdou, Y., Lasry, J.-M., and Lions, P. L. (2022b). Simulating numerically the krusell-smith model with neural networks.

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

Allen, T. and Arkolakis, C. (2025). Quantitative regional economics. In Donaldson, D. and Redding, S. J., editors, Handbook of Regional and Urban Economics, volume 6 of Handbook of Regional and Urban Economics, pages 1–72. Elsevier.

Allen, T., Arkolakis, C., and Li, X. (2024). On the equilibrium properties of spatial models. American Economic Review: Insights, 6(4):472–89.

Allen, T., Arkolakis, C., and Takahashi, Y. (2020). Universal Gravity. Journal of Political Economy, 128(2):393–433.

Alvarez, F., Lippi, F., and Souganidis, P. (2023). Price setting with strategic complementarities as a mean field game. Econometrica, 91(6):2005–2039.

Alvarez, F. and Lucas, R. E. (2007). General equilibrium analysis of the eaton–kortum model of international trade. Journal of Monetary Economics, 54(6):1726–1768.

Arkolakis, C., Demidova, S., Klenow, P. J., and Rodriguez-Clare, A. (2008). Endogenous variety and the gains from trade. American Economic Review, 98(2):444–50.

Armington, P. (1969). A theory of demand for products distinguished by place of production. IMF Staff Papers, 16:159–178.

Aronson, D. G. (1967). Bounds for the fundamental solution of a parabolic equation. Bulletin of the American Mathematical Society, 73(6):890 – 896.

Aydin, B., Parmaksiz, E., and Sircar, R. (2025). Fare game: A mean field model of stochastic intensity control in dynamic ticket pricing.

Baldwin, R. (2001). Core-periphery model with forward-looking expectations. Regional Science and Urban Economics, 31(1):21–49.

Baldwin, R. and Venables, A. J. (1994). International Migration, Capital Mobility and Transitional Dynamics. Economica, 61(243):285–300.

Bardi, M. and Capuzzo-Dolcetta, I. (2009). Optimal Control and Viscosity Solutions of Hamilton–Jacobi–Bellman Equations. Modern Birkhäuser Classics. Birkhäuser, Boston, MA. eBook reprint of the 1997 edition.

Bayraktar, E. and Zhang, X. (2023). Solvability of infinite horizon mckean–vlasov fbsdes in mean field control problems and games. Applied Mathematics & Optimization, 87(1):13.

Bensoussan, A., Chau, M., and Yam, S. (2016). Mean field games with a dominating player. Applied Mathematics and Optimization, 74:91–128.

Bertucci, C. (2020). Fokker-planck equations of jumping particles and mean field games of impulse control. Annales de l'Institut Henri Poincaré C, Analyse non linéaire, 37(5):1211–1244.

Bilal, A. (2023). Solving heterogeneous agent models with the master equation. NBER Working Paper 31103, National Bureau of Economic Research.

Bilal, A. and Rossi-Hansberg, E. (2023). Anticipating climate change across the united states. Working Paper 31323, National Bureau of Economic Research.

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

Caliendo, L. and Parro, F. (2015). Estimates of the trade and welfare effects of nafta. The Review of Economic Studies, 82(1 (290)):1–44.

Cameron, S., Chaudhuri, S., and McLaren, J. (2007). Trade shocks and labor adjustment: Theory. Working Paper 13463, National Bureau of Economic Research.

Cardaliaguet, P. (2018). A short course on mean field games. Accès le 5 octobre 2025.

Carmona, R., Delarue, F., and Lacker, D. (2015). Mean field games with common noise.

Carmona, R. and Laurière, M. (2021). Convergence analysis of machine learning algorithms for the numerical solution of mean field control and games i: The ergodic case. SIAM Journal on Numerical Analysis, 59(3):1455–1485.

Crandall, M. G., Ishii, H., and Lions, P.-L. (1992). User's guide to viscosity solutions of second order partial differential equations. Bulletin of the American Mathematical Society, 27(1):1–67.

Desmet, K., Nagy, D. K., and Rossi-Hansberg, E. (2018). The geography of development. Journal of Political Economy, 126(3):903–983.

Desmet, K. and Parro, F. (2025). Spatial dynamics. Working Paper 33443, National Bureau of Economic Research.

di Giovanni, J. and Levchenko, A. A. (2013). Firm entry, trade, and welfare in zipf's world. Journal of International Economics, 89(2):283–296.

Du, Y. (2006). Order Structure and Topological Methods in Nonlinear Partial Differential Equations: Maximum principles and applications. Order Structure and Topological Methods in Nonlinear Partial Differential Equations. World Scientific.

Dumas, A. and Santambrogio, F. (2024). Optimal trajectories in l^1 and under l^1 penalizations. Comptes Rendus. Mathématique, 362:657–692.

Dumas, A. and Santambrogio, F. (2025). Deterministic mean field games with jumps and mixed variational structure. ESAIM: COCV, 31:38.

Eaton, J. and Kortum, S. (2002). Technology, geography, and trade. Econometrica, 70(5):1741–1779.

Forslid, R. and Ottaviano, G. I. (2003). An analytically solvable core-periphery model. Journal of Economic Geography, 3(3):229–240.

Fujita, M., Krugman, P., and Venables, A. J. (1999). Many Regions and Continuous Space. In The Spatial Economy: Cities, Regions, and International Trade, pages 79–95. The MIT Press.

Gilbarg, D. and Trudinger, N. S. (1977). Elliptic Partial Differential Equations of Second Order, volume 224 of Grundlehren der mathematischen Wissenschaften. Springer-Verlag Berlin Heidelberg.

Helpman, E. (1998). The Size of Regions. In Pines, D., Sadka, E., and Zilcha, I., editors, Topics in Public Economics, pages 33–54. Cambridge University Press, London.

Huang, M., Malhamé, R. P., and Caines, P. E. (2006). Large population stochastic dynamic games: Closed-loop McKean-Vlasov systems and the Nash certainty equivalence principle. In Proceedings of the 45th IEEE Conference on Decision and Control, pages 2154–2159. IEEE.

Kleinman, B., Liu, E., and Redding, S. J. (2023). Dynamic spatial general equilibrium. Econometrica, 91(2):385–424.

Krugman, P. (1991). Increasing returns and economic geography. Journal of Political Economy, 99:483–499.

Krugman, P. (1996). Models of Spatial Self-Organization. In The Self-Organizing Economy, pages 75–99. Blackwell Publishers, Cambridge, MA.

Kucheryavyy, K., Lyn, G., and Rodríguez-Clare, A. (2023). Grounded by Gravity: A Well-Behaved Trade Model with Industry-Level Economies of Scale. American Economic Journal: Macroeconomics, 15(2):372–412.

Lasry, J.-M. and Lions, P.-L. (2007). Mean field games. Japanese Journal of Mathematics, 2(1):229–260.

Lavigne, P., Petit, Q., and Warin, X. (2025). Growth model with externalities for energetic transition via mfg with common external variable.

Lieberman, G. M. (2005). Second Order Parabolic Differential Equations. World Scientific, New Jersey; Singapore, revised ed. edition.

Moll, B. (2025). The trouble with rational expectations in heterogeneous agent models: A challenge for macroeconomics.

Mossay, P. (2006). Stability of spatial adjustments across local exchange economies. Regional Science and Urban Economics, 36(4):431–449.

Mossay, P. (2013). A theory of rational spatial agglomerations. Regional Science and Urban Economics, 43(2):385–394.

Ottaviano, G. (2001). Monopolistic competition, trade, and endogenous spatial fluctuations. Regional Science and Urban Economics, 31(1):51–77.

Redding, S. J. (2016). Goods trade, factor mobility and welfare. Journal of International Economics, 101:148–167.

Redding, S. J. and Rossi-Hansberg, E. (2017). Quantitative spatial economics. Annual Review of Economics, 9(Volume 9, 2017):21–58.

A Proofs

A.1 Notations

We denote by $C^0(X)$ or $C(X)$ the space of continuous functions on a topological space X and by $C^n(X)$ the space of functions on X with continuous derivatives up to order n. The associated norm is denoted by $\|f\|_{C^n} := \sum_{|\beta|\le n}\|\partial^\beta f\|_\infty$. The notations $C(X,Y)$ stands for the space of continuous maps from X to Y. A map $f : X \to Y$ between metric spaces is said to be Lipschitz if there exists a constant $L > 0$ such that

$$d_Y(f(x), f(y)) \le L d_X(x,y), \qquad \forall x,y \in X.$$

The parabolic Hölder space $C^{\alpha/2,\alpha}$ consists of functions $f = f(t,x)$ such that all derivatives $\partial_t^k \partial_x^\beta f$ with $2k+|\beta| \le \alpha$ are continuous and bounded, and the norm is defined by

$$\|f\|_{C^{\alpha/2,\alpha}} := \sum_{2k+|\beta|\le\alpha} \|\partial_t^k\partial_x^\beta f\|_\infty + \sup_{(t,x)\neq(s,y)} \frac{\left|\partial_t^k\partial_x^\beta f(t,x) - \partial_t^k\partial_x^\beta f(s,y)\right|}{|t-s|^{(\alpha-\lfloor\alpha\rfloor)/2} + |x-y|^{\alpha-\lfloor\alpha\rfloor}},$$

where $\lfloor\alpha\rfloor$ denotes the floor of α, i.e. the greatest integer less than or equal to α. We write $\mathbb{T}^2 := \mathbb{R}^2/\mathbb{Z}^2$ for the two-dimensional flat torus and $\mathbb{T}^1_R := \mathbb{R}/(R\mathbb{Z})$ for the one-dimensional torus of length $R > 0$. The set $\mathscr{P}(\mathbb{T}^2)$ denotes the space of Borel probability measures on $\mathbb{T}^2$. For $p \ge 1$, the Wasserstein space $(\mathscr{P}_p(\mathbb{T}^2), W^p)$ consists of probability measures $\mu \in \mathscr{P}(\mathbb{T}^2)$ with finite p-th moment, endowed with the p-Wasserstein distance

$$W^p(\mu,\nu) := \left(\inf_{\pi\in\Gamma(\mu,\nu)} \int_{\mathbb{T}^2\times\mathbb{T}^2} d(x,y)^p \, d\pi(x,y)\right)^{1/p},$$

where $\Gamma(\mu,\nu)$ is the set of couplings of μ and ν, and d denotes the Euclidean distance on $\mathbb{T}^2$. In particular, $(\mathscr{P}_1, W^1)$ and $(\mathscr{P}_2, W^2)$ denote the spaces equipped with the 1- and 2-Wasserstein distances, respectively.

The symbol $\deg_{LS}$ denotes the Leray–Schauder degree, and Ker stands for the kernel of a linear operator. We denote by $\|\cdot\|_\infty$ the uniform (supremum) norm, $\|f\|_\infty := \sup_{x\in X}|f(x)|$, and by $\|\cdot\|_{\mathscr{L}(B)}$ the operator norm on a Banach space B, defined for a bounded linear operator $T \in \mathscr{L}(E)$ by

$$\|T\|_{\mathscr{L}(E)} := \sup_{\substack{x\in E\\ x\neq 0}} \frac{\|Tx\|_E}{\|x\|_E}.$$

Here, $\mathscr{L}(E)$ denotes the space of bounded linear operators on a Banach space E, and $\mathscr{G}(E)$ denotes the set of invertible (bounded) linear operators on E. A subset $K \subset E$ is called a *cone* if it satisfies $K + K \subset K$ and $\lambda K \subset K$ for all $\lambda \ge 0$. A cone K is said to be *solid* if it has non-empty interior. We say that a linear operator T is *compact* if it maps bounded subsets of E to relatively compact subset, and *strongly positive* if for a solid cone, $T(K) \subset \operatorname{int} K$, where $\operatorname{int} K$ denotes the interior of K.

We denote by $\operatorname{span}\{v_1,\dots,v_n\}$ the linear span of the vectors $v_1,\dots,v_n$. If F is a differentiable operator, its Fréchet derivative is denoted by DF, sometimes also written F'. When considering spectral properties of linear operators, we use the standard terminology: an eigenvalue λ is said to be *simple* if its algebraic multiplicity is one, and we denote by v the associated *eigenvector*.

Finally, for a smooth function $f:\mathbb{R}^d\to\mathbb{R}$, we denote by $\nabla f=(\partial_{x_1}f,\ldots,\partial_{x_d}f)$ the spatial gradient, by $\Delta f=\sum_{i=1}^d\partial_{x_i}^2 f$ the Laplacian, and by $\partial_t f$ the time derivative of f.

A.2 Proof of Theorem 1

Existence

We first establish the existence of a wage function w satisfying (8) and (9). Let us introduce

$$\Xi=\left\{g\in C^1(\mathbb{T}^2)\ :\ \frac{1}{C_1}\le g(x)\le C_1,\ \|g'\|_\infty\le C_2 \text{ and } \int g(x)d\mu(x)=1\right\},$$

for some constants C_1 and C_2 to fix later. It is a convex and compact set of $C(\mathbb{T}^2,\mathbb{R})$ for the uniform topology. Let us now define the application $\Lambda:\Xi\to\Xi$ as follows: for any w in Ξ,

1. associate $f[w](x)=\int_{\mathbb{T}^2}\frac{\tau(x,y)^{1-\sigma}w(y)}{\int_{\mathbb{T}^2}(w(z)\tau(y,z))^{1-\sigma}d\mu(z)}d\mu(y)$, for any $x\in\mathbb{T}^2$,
2. then set $\Lambda[w](x)=f[w](x)^{\frac{1}{\sigma}}/\left(\int_{\mathbb{T}^2}f[w](y)^{\frac{1}{\sigma}}d\mu(y)\right)$.

Step 1. We start by showing that Λ is well-defined. Fix any w in Ξ. Notice that by construction,

$$\int_{\mathbb{T}^2}\Lambda[w](x)d\mu(x)=1.$$

To show estimates on $\Lambda[w]$, we first establish estimates on $f[w]$. Using the fact that τ is bounded by below and above, there exists $C_\tau>0$ such that

$$f[w](x)\le C_\tau\frac{\int_{\mathbb{T}^2}w(y)d\mu(y)}{\int_{\mathbb{T}^2}w(z)^{1-\sigma}d\mu(z)}.$$

From the convexity of $\omega\mapsto\omega^{1-\sigma}$ on $\mathbb{R}_+^*$, Jensen's inequality yields

$$f[w](x)\le C_\tau\frac{\int_{\mathbb{T}^2}w(y)d\mu(y)}{(\int_{\mathbb{T}^2}w(z)d\mu(z))^{1-\sigma}}$$

Using that $\int_{\mathbb{T}^2}w(y)d\mu(y)=1$, we conclude that $f(x)\le C_\tau$.

On the other hand, we have for any z, $w(z)^{1-\sigma}\le C_1^{\sigma-1}$. Therefore,

$$\left(\int_{\mathbb{T}^2}(w(z)\tau(y,z))^{1-\sigma}d\mu(z)\right)^{-1}\ge C_1^{1-\sigma}\left(\int_{\mathbb{T}^2}(\tau(y,z))^{1-\sigma}d\mu(z)\right)^{-1}.$$

Thus, there exists C'_τ such that $f(x)\ge C'_\tau C_1^{1-\sigma}$.

Coming back to $\Lambda[w]$, from its definition and the estimates on f, we deduce that

$$\left(\frac{C'_\tau}{C_\tau}\right)^{\frac{1}{\sigma}} C_1^{\frac{1}{\sigma}-1} \leq \Lambda[w](x) \leq \left(\frac{C_\tau}{C'_\tau}\right)^{\frac{1}{\sigma}} C_1^{1-\frac{1}{\sigma}}.$$

Therefore, we fix C_1 large enough to ensure that

$$\frac{1}{C_1} \leq \left(\frac{C'_\tau}{C_\tau}\right)^{\frac{1}{\sigma}} C_1^{\frac{1}{\sigma}-1} \text{ and } \left(\frac{C_\tau}{C'_\tau}\right)^{\frac{1}{\sigma}} C_1^{1-\frac{1}{\sigma}} \leq C_1.$$

In others words, we set C_1 bigger than C_τ/C'_τ. To highlight the fact that C_1 may only depends on τ, we choose $C_1 = C_\tau/C'_\tau$.

It remains to show that $\Lambda[w]$ is of class C^1 and its derivative is uniformly bounded. It is easy to see that $\Lambda[w]$ is differentiable, and

$$\Lambda[w]'(x) = \frac{f[w](x)^{\frac{1}{\sigma}-1}}{\int_{\mathbb{T}^2} f[w](y)^{\frac{1}{\sigma}} d\mu(y)} f[w]'(x),$$

where

$$f[w]'(x) = \int_{\mathbb{T}^2} \frac{\tau(x,y)^{-\sigma} w(y)}{\int_{\mathbb{T}^2} (w(z)\tau(y,z))^{1-\sigma} d\mu(z)} \frac{\partial \tau}{\partial x}(x,y) d\mu(y).$$

From the continuity of $f[w]$ and $f[w]'$, we deduce that $\Lambda[w]$ is C^1. Then, using the bounds on w, τ, $\frac{\partial \tau}{\partial x}$, and $f[w]$, we can deduce that there exists a constant C''_τ which only depends on τ (since all the estimates only depend on τ), such that

$$\|\Lambda[w]'\|_\infty \leq C''_\tau.$$

By setting $C_2 = C''_\tau$, we have proved that $\Lambda[w]$ belongs to Ξ.

Step 2. Let us show that finding a static equilibrium is equivalent of finding a fixed-point of Λ. The implication is trivial. Conversly, let us assume that w is a fixed-point of Λ. Then, the definition of Λ yields

$$C\Lambda[w](x)^\sigma = \int_{\mathbb{T}^2} \frac{\tau(x,y)^{1-\sigma} w(y)}{\int_{\mathbb{T}^2} (w(z)\tau(y,z))^{1-\sigma} d\mu(z)} d\mu(y), \tag{30}$$

where

$$C = \left(\int_{\mathbb{T}^2} f[w](y)^{\frac{1}{\sigma}} d\mu(y)\right)^\sigma.$$

Note that the result follows the fact that $C = 1$. Multiplying equation (30) by $w(x)^{1-\sigma}$ and integrating it against μ, we obtain

$$C\int_{\mathbb{T}^2} \Lambda[w](x)^\sigma w(x)^{1-\sigma} d\mu(x) = \int_{\mathbb{T}^2}\int_{\mathbb{T}^2} \frac{\tau(x,y)^{1-\sigma} w(y)}{\int_{\mathbb{T}^2} (w(z)\tau(y,z))^{1-\sigma} d\mu(z)} d\mu(y) w(x)^{1-\sigma} d\mu(x).$$

Using Tonelli's theorem and rearranging the terms, we obtain

$$C\int_{\mathbb{T}^2}\Lambda[w](x)^{\sigma}w(x)^{1-\sigma}d\mu(x)=\int_{\mathbb{T}^2}\frac{\int_{\mathbb{T}^2}(w(x)\tau(x,y)^{1-\sigma}d\mu(x)}{\int_{\mathbb{T}^2}(w(z)\tau(y,z))^{1-\sigma}d\mu(z)}w(y)d\mu(y).$$

We conclude that $C=1$, by using the symmetry of τ, the fact that $\Lambda[w]=w$, and that $\int wd\mu=1$. Therefore, w satisfies (8) and (9).

Step 3. We show that Λ is continuous for the uniform topology. Let (w_n) be a sequence of Ξ converging to w. The uniform bounds on (w_n) and its convergence to w yield that

$$\int_{\mathbb{T}^2}(w_n(z)\tau(\cdot,z))^{1-\sigma}d\mu(z)\to\int_{\mathbb{T}^2}(w(z)\tau(\cdot,z))^{1-\sigma}d\mu(z)$$

uniformly. From the previous convergence, the uniform bounds on (w_n) and its convergence to w, we deduce that $f[w_n]$ converges to $f[w]$ uniformly. The same kind of arguments allows us to conclude that $\Lambda[w_n]$ converges to $\Lambda[w]$ uniformly.

Step 4. From Step 1 and 3, Λ is a continuous map from Ξ into itself. Moreover, Ξ is a compact, convex and non empty subset of $C(\mathbb{T}^2,\mathbb{R})$ for C_τ chosen large enough. Therefore, we use Schauder's fixed-point theorem to conclude that there exists w in Ξ such that $w=\Lambda[w]$.

Local Uniqueness

The purpose of this paragraph is to use the Inverse Function Theorem to deduce the local uniqueness of a static equilibrium. Let us fix w^0 a static equilibrium, introduce

$$B=\left\{g\in C^0(\mathbb{T}^2,\mathbb{R})\ :\ \frac{1}{C_1}\leq g\leq C_1 \text{ and } \int g(x)d\mu(x)=1\right\}$$

and define $F:B\to C^0(\mathbb{T}^2,\mathbb{R})$ by

$$F(w)=f[w]^{\frac{1}{\sigma}}=\left(\int_{\mathbb{T}^2}\frac{\tau(x,y)^{1-\sigma}w(y)}{\int_{\mathbb{T}^2}(w(z)\tau(y,z))^{1-\sigma}d\mu(z)}d\mu(y)\right)^{\frac{1}{\sigma}}.$$

Step 1. Let us show that 1 is an eigenvalue of $DF(w^0)$, and that if $DF(w^0)(h)=h$, then $h\in\text{span}\{w^0\}$.

First, it is easy to show that F is Gateau-Differentiable and that DF is continuous, therefore F is of class C^1. Moreover,

$$DF(w)(h)=\sigma f[w]^{\frac{1-\sigma}{\sigma}}D_wf[w](h).$$

Therefore,

$$DF(w)(h) = \sigma f[w]^{\frac{1-\sigma}{\sigma}} \Bigg[\int_{\mathbb{T}^2} \frac{\tau(x,y)^{1-\sigma} h(y)}{\int_{\mathbb{T}^2} (w(z)\tau(y,z))^{1-\sigma} d\mu(z)} d\mu(y) + (\sigma-1) \int_{\mathbb{T}^2} \frac{\tau(x,y)^{1-\sigma} w(y) \int_{T^2} (w(z)\tau(y,z))^{1-\sigma} \frac{h(z)}{w(z)} d\mu(z)}{\left(\int_{\mathbb{T}^2} (w(z)\tau(y,z))^{1-\sigma} d\mu(z)\right)^2} d\mu(y) \Bigg].$$

We claim that $DF(w_0)$ is a linear, compact and strongly positive operator of C^0.

Since the linearity is trivial, let us only check that it is compact and strongly positive. Let $\mathscr{B}$ be a bounded set of $(C^0(\mathbb{T}^2), \|.\|_\infty)$. There exists $b > 0$ such that $h \in \mathscr{B}$ implies $\|h\|_\infty \leq b$. The regularity of τ and w, and the bound on h implies that $\|DF(w)(h)\|_\infty$ and $\|DF(w)(h)'\|_\infty$ are bounded by above by a constant that only depends on τ, w and b. Then Arzelà-Ascoli's theorem ensures that $DF(w_0)(\mathscr{B})$ is compact in $(C^0(\mathbb{T}^2), \|.\|_\infty)$.

Now let us take a nonnegative function h, not equals to 0. The kernel functions multiplying h in the integrals are bounded by below by positive constants. Moreover, supp $\mu = \mathbb{T}^2$, then $DF(w_0)(h)$ is strongly positive on $\mathbb{T}^2$.

Since for any $t \in \mathbb{R}$, tw_0 is solution of $w = F(w)$, then $DF(w_0)(w_0) = w_0$. In order to conclude that if $h \in C^0(\mathbb{T}^2, \mathbb{R})$ satisfies $DF(w_0)(h) = h$, then $h \in \text{span}\{w_0\}$, we use the following result:

Theorem 4. *Let X be a Banach space, $K \subset X$ a solid cone, $T : X \to X$ a compact linear operator which is strongly positive. Then*

1. *the spectral radius $r(T)$ is a simple eigenvalue with an eigenvector $v \in \mathring{K}$ and there is no other eigenvalue with a positive eigenvector.*
2. *$|\lambda| < r(T)$ for all eigenvalue $\lambda \neq r(T)$.*

Proof. The proof of this result can be found in Chapter 1 of the book Du (2006). □

Step 2. Let us introduce for $i = 0, 1$,

$$\Theta^i = \left\{ g \in C^0(\mathbb{T}^2, \mathbb{R}) \; : \; \int g d\mu = i \right\}.$$

Let $T : \Theta^0 \to \Theta^1$ be defined by

$$T(u) = u + 1.$$

It is a bijection. Note that $(\Theta^0, \|.\|_\infty)$ is a Banach space. Let us define $G : T^{-1}(B) \subset \Theta^0 \to \Theta^0$ by

$$G(u) = T(u) - F(T(u)).$$

The function G is differentiable in Θ^0 and

$$DG(u)(h) = h - \frac{DF(T(u))(h)}{\int_{\mathbb{T}^2} F(T(u)) d\mu} + \frac{F(T(u))}{\left(\int_{\mathbb{T}^2} F(T(u)) d\mu\right)^2} \int DF(T(u))(h) d\mu$$

Fix u_0 such that $T(u_0) = w_0$. We want to use the Fredholm alternative to show that $DG(u_0)$ is an isomorphism. We already know that $DF(T(u_0))$ is a compact operator. Moreover, let us establish that $\int DF(T(u_0))(h)d\mu = 0$ for any h. Let us notice that for any w,

$$\int F(w)^\sigma w^{1-\sigma} d\mu = \int w d\mu = 1.$$

Differentiating this equation with respect to w in the direction $h \in \Theta^0$, we get

$$\sigma \int F(w)^{\sigma-1} DF(w)(h) w^{1-\sigma} = (\sigma - 1) \int F(w)^\sigma w^{-\sigma} h d\mu.$$

Setting $w = w_0$ and using $F(w_0) = w_0$, we end up with

$$\sigma \int DF(w_0)(h) d\mu = (\sigma - 1) \int h d\mu = 0.$$

It is the wanted equality. With the fact that $\int F(T(u_0))d\mu = 1$, it allows us to write

$$DG(u_0)(h) = h - DF(T(u_0))(h)$$

Note that $DG(u_0)$ is a Fredholm operator on Θ^0. Let us verify that $\mathrm{Ker} DG(u_0)$ is reduced to $\{0\}$. Fix $h \in \mathrm{Ker} DG(u_0)$, then by step 1, $h = tw_0$ but the condition $\int h = 0$ ensures that $t = 0$ meaning that $h = 0$. Therefore, $DG(u_0)$ is an isomorphism in Θ^0 and we can apply the Inverse Function Theorem to deduce the local uniqueness of u_0 in Θ^0.

Global Uniqueness

To deduce the global uniqueness of a solution of $\Lambda(w) = w$ in Θ^1, we will use the fact that when $\tau \equiv 1$ there is a unique static equilibrium. In particular, we will show that the value of the Leray-Schauder degree is 1. Then, using the properties of the degree, we deduce that a change of τ does not modify its value. Therefore, from the degree formula and the previous steps, we deduce that there is at most one static equilibrium.

Fix $\tau \equiv 1$. In this case, it is trivial to see that the unique static equilibrium is $w_0 \equiv 1/\mu(\mathbb{T}^2)$. Therefore, the unique solution of $G_1 = 0$ (the dependency of G with respect to τ has been made clear) is $u_0 = T^{-1}(w_0)$. It allows us to deduce the value of the Leray-Schauder degree as follows:

$$\deg_{LS}(G_1, T^{-1}(B), 0) = \sum_{a \in G_1^{-1}(0)} (-1)^{\sigma(a)},$$

where $\sigma(a)$ is the aglebric multiplicities of the eigenvalues of $DF(T(a))$ contained in $(1, +\infty)$. From the above, the sum is reduced to one element u_0, and from step 1 of the Local Uniqueness paragraph, we know that $DF(T(u_0))$ viewed as an operator of $C(\mathbb{T}^2, \mathbb{R})$ has 1 as simple eigenvalue and that the others have a norm strictly smaller than 1, hence $\sigma(u_0) = 0$, leading to

$$\deg_{LS}(G_1, T^{-1}(B), 0) = 1,$$

where B has been defined in the previous paragraph.

Now, fix any τ satisfying the conditions of the Theorem and introduce $\tau_t = (1-t)+t\tau$. By homotopy invariance of the degree, for any $t \in [0,1]$

$$\deg_{LS}(G_{\tau_t}, T^{-1}(B), 0) = 1.$$

Therefore, using the formula of $\deg_{LS}$ and the properties of DF, we conclude that there is at most one static equilibrium concluding the proof.

A.3 Proof of Theorem 2

Let us start by proving that the wages trajectory admits time regularity when μ is regular in time.

Lemma 1. *If μ belongs to $C([0,T], \mathscr{P}_2(\mathbb{T}^2))$ and is Hölder continuous in time for the 2-Wasserstein distance W_2, i.e. there exists $C>0$ such that for all $s,t \in [0,T]$*

$$W_2(\mu(s), \mu(t)) \leq C\,|t-s|^{\frac{1}{2}},$$

then, denoting $w(t)$ the static equilibrium associated to $\mu(t)$, there exists C' such that for all $s,t \in [0,T]$

$$\|w(t)-w(s)\|_\infty \leq C'\,|t-s|^{\frac{1}{2}}.$$

Proof. The proof is based on the implicit function theorem. We use the notations introduced in the proof of Theorem 1, where the dependency of G with respect to μ has been made clear in the following. Setting $u(t) = T^{-1}(w(t))$, for any $t \in [0,T]$,

$$G(u(t), \mu(t)) = 0.$$

Let us introduce a probabilistic set up $(\Omega, \mathscr{A}, \mathbb{P})$ such that every probability measure μ on $\mathbb{T}^2$ admits a random variable X such that $\mathscr{L}(X) = \mu$. Since $\mathbb{T}^2$ is a compact set, the Wasserstein space $(\mathscr{P}_1(\mathbb{T}^2), W_1)$ and $(\mathscr{P}_2(\mathbb{T}^2), W_2)$ are equivalent. We therefore see $G : (\Theta^0, \|.\|_\infty) \times (\mathscr{P}_2(\mathbb{T}^2), W_2)) \to (\Theta^0, \|.\|_\infty)$. Since $(\mathscr{P}_2(\mathbb{T}^2), W_2))$ is not a Banach space, we lift it into $L^2(\Omega, \mathbb{T}^2)$ i.e. we work with

$$\tilde{G} : (\Theta^0, \|.\|_\infty) \times (L^2(\Omega, \mathbb{T}^2), \|.\|_{L^2}) \to (\Theta^0, \|.\|_\infty)$$

defined for any (u,X) by

$$\tilde{G}(u,X) = G(u, \mathscr{L}(X)).$$

Let us fix (u_0, X_0) such that $\tilde{G}(u_0, X_0) = 0$. Using the same arguments developped in the proof of Theorem 1, we know that $D_u\tilde{G}(u_0, X_0)$ is an isomophism of Θ^0. Since $\tilde{G}$ is C^1, we can apply the implicit function theorem to conclude that in a neighborhood of (u_0, X_0), the zero of $\tilde{G}$ are of the form $(\tilde{U}(X), X)$ where $\tilde{U}$ is C^1 and its derivative satisfies:

$$D_X\tilde{U}(X) = -D_u\tilde{G}(\tilde{U}(X), X)^{-1} D_X\tilde{G}(\tilde{U}(X), X).$$

Coming back to $(\mathscr{P}_2(\mathbb{T}^2), W_2)$ and using Proposition 5.1.8 of Cardaliaguet (2018), we deduce that U, the application satisfying $U(\mathscr{L}(X)) = \tilde{U}(X)$, is C^1 in the intrinsic sense and that for any μ

$$D_\mu U(\mu) = -D_u G(U(\mu),\mu)^{-1} D_\mu G(U(\mu),\mu).$$

Let us introduce

$$\mathscr{K} = \left\{(u,\mu) \in C^0(\mathbb{T}^2) \times \mathscr{P}(\mathbb{T}^2) \ : \ G(u,\mu) = 0\right\}.$$

The continuity of G ensures that $\mathscr{K}$ is closed, and the existence part in the proof of Theorem 1 ensures that $\mathscr{K}$ is compact.

Moreover, the set $\mathscr{G}(E)$ of invertible operators is open in $\mathscr{L}(E)$, the space of bounded linear operators, and the map

$$\mathscr{G}(E) \ni A \mapsto A^{-1} \in \mathscr{L}(E)$$

is continuous. The regularity of G and the previous observation ensure that the map

$$\mathscr{K} \ni (u,\mu) \mapsto D_u G(u,\mu)^{-1} \in \mathscr{L}(E)$$

is continuous.

It allows us to deduce that its direct image is compact. Therefore, there exists $C > 0$ such that

$$\max_{(u,\mu)\in\mathscr{K}} \|D_u G(u,\mu)^{-1}\|_{\mathscr{L}(C^0(\mathbb{T}^2))} \leq C < +\infty.$$

We deduce that

$$\|D_\mu U(\mu)\|_\infty \leq \max_{(u,\mu)\in\mathscr{K}} \|D_u G(u,\mu)^{-1}\|_{\mathscr{L}(C^0(\mathbb{T}^2))} \max_{(u,\mu)\in\mathscr{K}} \|D_\mu G(u,\mu)\|_\infty =: R < +\infty.$$

Therefore, U is Lipschitz on $\mathscr{K}$.

Noticing that

$$\|w(t) - w(s)\|_\infty = \|u(t) - u(s)\|_\infty = \|U(\mu(t)) - U(\mu(s))\|_\infty \leq R W_2(\mu(t),\mu(s)) \leq RC\,|t-s|^{\frac{1}{2}}.$$

□

Let us now state the existence of a dynamic equilibrium using the Schauder fixed-point theorem.

Let us introduce

$$\Omega = \left\{\mu \in C^0([0,T],\mathscr{P}_2(\mathbb{T}^2)) \ : \ \sup_{t\neq s} \frac{W_2(\mu(t),\mu(s))}{|t-s|^{\frac{1}{2}}} \leq C\right\}.$$

It is a non-empty, closed and convex subset of $C^0([0,T],\mathscr{P}_2(\mathbb{T}^2))$. Moreover, from Arzela-Ascoli's theorem, Ω is compact.

Let us now define the map Ψ in the following way: from $\mu \in \Omega$

1. associate w such that for any t, $w(t)$ is the static equilibrium associated with $\mu(t)$,
2. then associate u the solution of the HJB equation (14),
3. then define $\Psi[\mu]$ the solution of the Fokker-Planck equation (15).

In the paragraphs below we establish that Ψ is well-defined. We also develop arguments to show that the solution of the different equations admits strong regularity properties : this will help to show the continuity of Ψ and then apply the Schauder fixed-point theorem.

Let us fix μ an element of Ω.

Step 1. Invoking Lemma 1, w admits a $1/2-$Hölder estimate in time. On the other hand, thanks to Theorem 1, we already know that w admits uniform and Lipschitz estimates in space which only depend on τ.
The regularity of w and μ ensures that the function $\tilde{V}(t,x) = V(x, w(t), \mu(t))$ belongs to $C^{1/2,1}$ and admits estimates in it, depending only on the constant C and the map τ.

Step 2. Concerning the HJB equation, by inverting the time as follows: $\tau = T - t$, (14) becomes a parabolic equation. Then we can use the results on quasilinear parabolic-equations to state existence and uniqueness of the solution as well as estimates.
Indeed, using Theorem 12.16 from Lieberman (2005), we obtain existence of classical solutions $u \in C^{1,2}$. Uniqueness is ensured by comparison principle (Theorem 9.1 or 9.7 from Lieberman (2005)). The comparison principle also provides a uniform bound for u only depending of $\|\tilde{V}\|_\infty$ and $\|g(\cdot, w(T), \mu(T))\|_\infty$. Using the Ishii-Lions' method (see Crandall et al. (1992), for an introduction of the method), we deduce a uniform bound for ∇u only depending on $\|\nabla g(\cdot, w(T), \mu(T))\|_\infty$, $\|\nabla \tilde{V}\|_\infty$ and ρ. Now using Theorem 12.10, we deduce that there exists a constant C' only depending of ρ, $\|g(\cdot, w(T), \mu(T))\|_{C^1}$ and $\|\tilde{V}\|_{C^{1/2,1}}$ such that

$$\|u\|_{C^{1/2,1}} \leq C'.$$

Step 3. The previous regularity of u is sufficient to guarantee the existence and uniqueness of a classical solution (use for instance Theorem 9.7 and 12.22 from Lieberman (2005)). Moreover, classical arguments (see Cardaliaguet (2018)) ensure that when $\|\nabla u\|_\infty < +\infty$, then the solution of the Fokker-Planck equation μ is Hölder continuous in time, i.e. there exists a constant C'', only depending on $\|\nabla u\|_\infty$, such that for any t, s

$$W_2(\mu(t), \mu(s)) \leq C'' |t - s|^{\frac{1}{2}}.$$

Conclusion. By fixing $C = C''$, the above discussion ensures that $\Psi[\Omega] \subset \Omega$.
To use the Schauder fixed-point theorem, it remains to show the continuity of Ψ. It is established by showing the continuity of the maps composing Ψ. In the following, we develop the argument for the first one, i.e. the one that associates w a static equilibrium trajectory from μ, and let the reader apply the same reasoning for the next ones. The argument is based on compactness and uniqueness properties.
Let μ_n converging to μ in $C([0,T], \mathscr{P}_2(\mathbb{T}^2))$. Let us denote by w_n (resp. w) the static equilibrium trajectory associated to μ_n (resp. μ). Step 1 and Arzela-Ascoli's theorem ensures that w_n belongs to a compact set

of $C([0,T]\times\mathbb{T}^2,\mathbb{R}_+)$. Therefore, there exists $\tilde{w}$ and a subsequence of w_n (also denoted by w_n for ease of notations) such that w_n converges to $\tilde{w}$ uniformly. For any n and t, $G(T^{-1}(w_n(t)),\mu_n(t))=0$. The continuity of T^{-1} and G ensure that $G(T^{-1}(\tilde{w}(t)),\mu(t))=0$. The uniqueness of the static equilibrium established in Theorem 1 leads to $\tilde{w}=w$. Therefore, every converging subsequence of w_n converges to w. Meaning that the whole sequence converges uniformly to w.
To prove the continuity of the others maps composing Ψ, we use the same argument where the compactness comes from the estimates established above and the Arzela-Ascoli theorem, while uniqueness comes from the uniqueness of the solution of the HJB and Fokker-Planck equation.
We conclude that Ψ is continuous and by Schauder's fixed-point theorem that it admits a fixed-point corresponding to an equilibrium in the sens of Definition 3.

A.4 Proof of Theorem 3

The proof of this result is simpler in the stationary case than in the time-dependent case. Let us point out which paragraphs are simplified.
First, no time regularity is needed, therefore Lemma 1 should only be stated as below simplifying its proof:

Lemma 2. *If (μ_n) converges to μ in $\mathscr{P}_2(\mathbb{T}^2)$, then the sequence of associated static equilibria converges uniformly to the static equilibrium associated to μ.*

Second, in the time-dependent case, the HJB (when inverting the time) and the FP equation are parabolic-type equations; in the stationary case, they become elliptic. Existence and uniqueness of classical solutions of the equations are ensured, as well as uniform estimates for u and its gradient (Gilbarg and Trudinger (1977)). These results and the fact that the subset of elements of $C([0,T],\mathscr{P}_2(T^2))$ satisfying Holder estimate is compact are enough to reproduce steps 1, 2 and 3 of the proof of Theorem 2.

A.5 Proof of Proposition 1

Consider the Dirac measure $\mu=\delta_{x^0}$. The normalization $\int w\,d\mu=1$ implies that $w(x^0)=1$. The static equation

$$w(x)^\sigma=\int_{\mathbb{T}^2}\frac{\tau(y,x)^{1-\sigma}w(y)}{\int_{\mathbb{T}^2}(w(z)\tau(y,z))^{1-\sigma}d\mu(z)}d\mu(y)$$

collapses to the evaluation at $y=x^0$ and $z=x^0$. The denominator becomes $\left(w(x^0)\tau(x^0,x^0)\right)^{1-\sigma}$, yielding

$$w(x)=\left(\frac{\tau(x^0,x^0)}{\tau(x^0,x)}\right)^{\frac{\sigma-1}{\sigma}}.$$

By definition,

$$\begin{aligned} V(x,w,\mu) &= \frac{w(x)}{\left(\int_{\mathbb{T}^2}(w(z)\tau(x,z))^{1-\sigma}\,d\mu(z)\right)^{\frac{1}{1-\sigma}}} \\ &= \frac{w(x)}{w(x^0)\tau(x,x^0)} \\ &= \frac{1}{\tau(x,x^0)}\left(\frac{\tau(x^0,x^0)}{\tau(x^0,x)}\right)^{\frac{\sigma-1}{\sigma}}. \end{aligned}$$

In particular $V(\cdot,w,\mu)$ admits a maximum at x^0 and $V(x^0)=1/\tau(x^0,x^0)$.

The stationary HJB on the torus is

$$-\frac{|\nabla u|^2}{2c_0}+\rho\,u=V(x,\mu,w) \qquad \text{on } \mathbb{T}^2. \tag{31}$$

Set $F(x,r,p):=-\frac{|p|^2}{2c_0}+\rho r-V(x,\mu,w)$. Then F is continuous and *proper* (strictly increasing in r). By the standard comparison principle for first–order discounted Hamilton–Jacobi equations (see, e.g., Bardi and Capuzzo-Dolcetta (2009)), any bounded upper semicontinuous subsolution u and bounded lower semicontinuous supersolution v of (31) satisfy $u\le v$ on $\mathbb{T}^2$. In particular, (31) has *at most one* bounded continuous viscosity solution.

Existence via Perron. Constants give barriers: let $M:=\max_{\mathbb{T}^2}V(\cdot,\mu,w)$ and $m:=\min_{\mathbb{T}^2}V(\cdot,\mu,w)$. Note that $M=V(x^0)$. Then $u_-(x):=m/\rho$ is a subsolution and $u_+(x):=M/\rho$ is a supersolution (test with constants in the viscosity sense). Perron's method (see Bardi and Capuzzo-Dolcetta (2009)) in combination with the comparison principle yields a unique bounded continuous viscosity solution u with

$$\frac{m}{\rho}\ \le\ u(x)\ \le\ \frac{M}{\rho} \qquad \forall x\in\mathbb{T}^2. \tag{32}$$

Moreover, by the coercivity of H the solution is Lipschitz and semiconcave on $\mathbb{T}^2$ (cf. Bardi and Capuzzo-Dolcetta (2009)). Evaluating the solution at x^0 gives

$$u(x^0)=\frac{V(x^0)}{\rho},$$

ensuring that u admits a maximum at x_0.

Differentiability of u at x^0. Because u is semiconcave, it is differentiable a.e. and its Clarke generalized gradient at any point is the closed convex hull of limits of $\nabla u(x_n)$ along differentiability points $x_n\to x$. Pick any sequence $x_n\to x^0$ where u is differentiable. At each such point the equation holds classically:

$$\frac{|\nabla u(x_n)|^2}{2c_0}=\rho\,u(x_n)-V(x_n).$$

By continuity of u and V and the identity $u(x^0)=V(x^0)/\rho$, the right-hand side tends to 0, hence $|\nabla u(x_n)|\to$

0. Therefore every reachable gradient at x^0 is 0, the Clarke gradient reduces to $\{0\}$, and u is differentiable at x^0 with

$$\nabla u(x^0) = 0.$$

It ensures that $\mu = \delta_{x^0}$ is a solution in the distributional sense of the Fokker–Planck equation.

A.6 Proof of Proposition 3

Up to multiplicative constants, the proof closely follows Fujita et al. (1999), sections 6.4 and 6.5. Let

$$\begin{aligned} Y_0(x) &= \mu_0(x) w_0(x) \\ G_0(x) &= \left[R \int_{-\pi}^{\pi} w_0(y)^{1-\sigma} e^{-d(\sigma-1)|x-y|} \mu_0(y) dy \right]^{1/(1-\sigma)} \\ w_0(x)^{\sigma} &= \left[R \int_{-\pi}^{\pi} Y_0(y) G_0(y)^{\sigma-1} e^{-d(\sigma-1)|x-y|} dy \right] \\ V_0(x) &= w_0(x)/G_0(x). \end{aligned} \tag{33}$$

At the flat-earth equilibrium, we have $\bar{\mu} = 1$, $\bar{w} = 1$, $\bar{Y} = 1$, and $\bar{G} = \left[R \int_0^{\pi} e^{-d(\sigma-1)x} dx \right]^{1/1-\sigma}$. Totally differentiating (33) around the flat earth, we obtain:

$$\begin{aligned} \tilde{Y}_0(x) &= \tilde{\mu}_0(x) + \tilde{w}_0(x) \\ \frac{\tilde{G}_0(x)}{\bar{G}} &= \bar{G}^{\sigma-1} R \int_{-\pi}^{\pi} \tilde{w}_0(x+y) e^{-d(\sigma-1)|y|} dy \\ &\quad - \frac{\bar{G}^{\sigma-1}}{\sigma-1} R \int_{-\pi}^{\pi} \tilde{\mu}_0(x+y) e^{-d(\sigma-1)|y|} dy \\ \tilde{w}_0(x) &= \frac{\bar{G}^{\sigma-1}}{\sigma} R \int_{-\pi}^{\pi} \tilde{Y}_0(x+y) e^{-d(\sigma-1)|y|} dy \\ &\quad + \frac{\sigma-1}{\sigma} \bar{G}^{\sigma-1} R \int_{-\pi}^{\pi} (\tilde{G}_0(x+y)/\bar{G}) e^{-d(\sigma-1)|y|} dy \\ \tilde{V}_0(x) &= \left[\tilde{w}_0(x) - \tilde{G}_0(x)/\bar{G} \right] \bar{G}^{-1}. \end{aligned} \tag{34}$$

Consider a sinusoidal solution to the system (34) of the form:

$$\begin{aligned} \tilde{Y}_0(x) &= \delta_Y \cos(kx), \\ \tilde{G}_0(x)/\bar{G} &= \delta_G \cos(kx), \\ \tilde{w}_0(x) &= \delta_w \cos(kx), \\ \tilde{V}_0(x) &= \delta_V \cos(kx). \end{aligned} \tag{35}$$

where the constants δ_Y, δ_G, δ_w and δ_V have to be determined. Introducing

$$Z = R \bar{G}^{\sigma-1} \int_{-\pi}^{\pi} \cos(kx) e^{-d(\sigma-1)|x|} dx,$$

the system of equations (34) becomes:

$$\begin{aligned}\delta_Y &= \delta_\mu + \delta_w, \\ \delta_G &= \frac{1}{1-\sigma} Z\delta_\mu + Z\delta_w, \\ \sigma\delta_w &= Z\delta_Y + (\sigma-1)Z\delta_G, \\ \delta_V &= \bar{G}^{-1}(\delta_w - \delta_G).\end{aligned} \tag{36}$$

which gives:

$$\frac{\delta_V}{\delta_\mu} = \bar{G}^{-1}\frac{1-s^2}{s}\frac{1-Z}{1-Z(1-s)-sZ^2} \tag{37}$$

where $\rho = (\sigma-1)^{-1}$. When R is large, Z belongs to $(0,1)$ (Fujita et al. (1999), section 6.5), which guarantees that the ratio given by equation (37) is positive.

A.7 Proof of Proposition 4

Define

$$X(t) = \begin{pmatrix} A(t) \\ B(t) \end{pmatrix}, \quad M = \begin{pmatrix} -\nu k^2 & \frac{k^2}{c_0} \\ -\frac{\delta_V}{\delta_\mu} & \nu k^2 + \rho \end{pmatrix}.$$

Then $X'(t) = MX(t)$, whose general solution is

$$X(t) = e^{Mt}X(0).$$

Since $A(0) = \delta_\mu$, write

$$X(0) = \begin{pmatrix} \delta_\mu \\ B(0) \end{pmatrix}, \quad e^{MT} = (e_{ij})_{1\le i,j\le 2}$$

Hence

$$X(T) = \begin{pmatrix} e_{11}\,\delta_\mu + e_{12}\,B(0) \\ e_{21}\,\delta_\mu + e_{22}\,B(0) \end{pmatrix},$$

and the terminal condition $B(T) = 0$ yields

$$e_{21}\,\delta_\mu + e_{22}\,B(0) = 0.$$

Thus, if

$$e_{22} \neq 0,$$

then there is a unique solution

$$B(0) = -\frac{e_{21}}{e_{22}}\delta_\mu,$$

and hence a unique trajectory

$$X(t) = e^{Mt} \begin{pmatrix} \delta_\mu \\ B(0) \end{pmatrix},$$

solving the mixed boundary-value problem. A simple calculation of e^{MT} by diagonalizing M gives:

$$e_{22} = \frac{\beta_2^{-1} e^{\lambda_2 T} - \beta_1^{-1} e^{\lambda_1 T}}{\beta_2^{-1} - \beta_1^{-1}},$$

which is nonzero for almost any combination of parameters ν, k, c_0, δ_V, δ_μ, ρ, T.

A.8 Proof of Corollary 2

When T tends to infinity,

$$C_1 \underset{T\to\infty}{\longrightarrow} 0, \quad C_2 \underset{T\to\infty}{\longrightarrow} \delta_\mu.$$

Thus when the time horizon is large, the exponential $e^{\lambda_1 t}$ disappears from the solution, and $A(t)$ goes to $\delta_\mu e^{\lambda t}$, with $\lambda = \frac{1}{2}\left[\rho - \sqrt{\rho^2 + 4\left(\nu^2 k^4 + \rho \nu k^2 - \frac{\delta_V k^2}{\delta_\mu c_0}\right)}\right]$.

A.9 Proof of Proposition 5

We identify the one-dimensional torus $\mathbb{T}_R^1$ with the interval $[0, L)$ with $L = 2\pi R$ and endpoints identified. Let $f \in \mathscr{P}(\mathbb{T}_R^1)$. We firt show that:

$$H[f] \geq \frac{1}{L},$$

with equality if and only if f is the uniform distribution $f_{\text{unif}}(x) := \frac{1}{L}$.

Apply the Cauchy–Schwarz inequality to the functions f and 1:

$$\left(\int_{\mathbb{T}_R^1} f(x) \cdot 1\, dx\right)^2 \leq \left(\int_{\mathbb{T}_R^1} f(x)^2\, dx\right)\left(\int_{\mathbb{T}_R^1} 1^2\, dx\right).$$

Since $\int_{\mathbb{T}_R^1} f = 1$ and $\int_{\mathbb{T}_R^1} 1 = L$, we get:

$$1 \leq H[f] \cdot R \quad \Rightarrow \quad H[f] \geq \frac{1}{L}.$$

Equality in Cauchy–Schwarz occurs if and only if f is constant almost everywhere. As f is a probability density, the constant must be $\frac{1}{L}$. Hence, equality holds exactly for the uniform distribution.

Now, let $(f_k)_{k\in\mathbb{N}} \subset \mathscr{P}(\mathbb{T}_R^1)^{\mathbb{N}}$ converge to δ_{x_0} in the sense of distributions. That is, for every continuous test function φ on $\mathbb{T}$:

$$\lim_{k\to\infty} \int_{\mathbb{T}_R^1} f_k(x)\varphi(x)\, dx = \varphi(x_0).$$

We want to show that $\lim_{k\to\infty} H[f_k] = +\infty$.

Let $M > 0$ be arbitrary. We will show that for all sufficiently large k, $H[f_k] > M$.

1. Choose an open neighborhood U of x_0 such that $|U| < \frac{1}{4M}$, where $|U|$ denotes the Lebesgue measure of U.

2. Since $f_k \to \delta_{x_0}$ weakly and U is an open set containing x_0, we have:

$$\lim_{k\to\infty} \int_U f_k(x)\,dx = \delta_{x_0}(U) = 1.$$

3. In particular, for $\varepsilon = \frac{1}{2}$, there exists $K \in \mathbb{N}$ such that for all $k > K$:

$$\int_U f_k(x)\,dx > \frac{1}{2}.$$

4. Apply the Cauchy–Schwarz inequality on U:

$$\left(\int_U f_k(x)\,dx\right)^2 \le \left(\int_U f_k(x)^2\,dx\right) \cdot |U|.$$

5. Using the lower bound $\int_U f_k > \frac{1}{2}$, we get:

$$\frac{1}{4} \le \left(\int_U f_k\right)^2 \le \left(\int_U f_k^2\right) \cdot |U| \quad \Rightarrow \quad \int_U f_k^2 \ge \frac{1}{4|U|}.$$

6. Since $H[f_k] = \int_{\mathbb{T}} f_k^2 \ge \int_U f_k^2$, we conclude:

$$H[f_k] > \frac{1}{4|U|} > M,$$

because $|U| < \frac{1}{4M} \Rightarrow \frac{1}{4|U|} > M$.

Since M was arbitrary, $\lim_{k\to\infty} H[f_k] = +\infty$.

A.10 Proof of Proposition 6

A double integration by parts, using periodic boundary conditions, gives:

$$\int \Delta\tilde{u}(t,x)\tilde{\mu}(t,x)\,dx = \int \tilde{u}(t,x)\Delta\tilde{\mu}(t,x)\,dx$$

Now, by the linearized HJB equation (23), $\tilde{u} = \frac{1}{\rho}(\tilde{V} + \nu\Delta\tilde{u} + \partial_t\tilde{u})$. We conclude by using equation (29).

A.11 Proof of Proposition 7

By Corollary 2, $\tilde{\mu}(t,x) = \delta_\mu e^{\lambda t}\cos(kx)$ and $\tilde{u}(t,x) = \beta_2^{-1}\delta_\mu e^{\lambda t}\cos(kx)$. By Proposition 3, $\tilde{V}[\tilde{\mu}_t](x) = \frac{\delta_V}{\delta_\mu}\tilde{\mu}(t,x)$. We conclude by using Proposition 6.

B Extensions

B.1 Other monopolistic competition models

We consider alternative models of monopolistic competition (see, e.g., Helpman (1998) with local ownership, Forslid and Ottaviano (2003), Arkolakis et al. (2008), di Giovanni and Levchenko (2013), Redding and Rossi-Hansberg (2017), §3), in which the equilibrium wage equation takes the form

$$w(x)^{\sigma} = \int_{\mathbb{T}^2} \frac{A(\mu,x), \tau(x,y)^{1-\sigma}}{\int_{\mathbb{T}^2} A(\mu,z), \tau(z,y)^{1-\sigma} w(z)^{1-\sigma}, d\mu(z)}, w(y), d\mu(y).$$

Relative to the baseline specification, this introduces a spatial heterogeneity component, potentially dependent on the distribution μ. We assume that the heterogeneity function A is C^1, Lipschitz continuous in both arguments, and uniformly bounded above and below by positive constants.

This extension does not alter the core of the existence proof. One only needs to adjust the fixed-point operator and the bounding constants. In the proof of Theorem 1, this amounts to replacing the operator f with

$$f[w](x) = \int_{\mathbb{T}^2} \frac{A(\mu,x), \tau(x,y)^{1-\sigma}, w(y)}{\int_{\mathbb{T}^2} A(\mu,z), (w(z)\tau(y,z))^{1-\sigma}, d\mu(z)}, d\mu(y),$$

and modifying the bounding constants accordingly.

Similarly, in the proof of Theorem 2, the regularity assumptions on A ensure that the derivative $D_{\mu}G(u,\mu)$ remains bounded in the argument of Lemma 1 leading to the Lipschitz property of the application U defined in the proof of the lemma.

B.2 Perfect competition models

We consider here perfect-competition models (e.g., Armington (1969), Eaton and Kortum (2002), Alvarez and Lucas (2007), Allen and Arkolakis (2014), Redding (2016)), in which the equilibrium wage equation takes the form

$$\mu(x)w(x)^{\sigma} \ = \ \int_{\mathbb{T}^2} \frac{A(\mu,x)\,\tau(x,y)}{\int_{\mathbb{T}^2} A(\mu,z)\,\tau(z,y)\,w(z)^{1-\sigma}\,dz}\, w(y)\,d\mu(y).$$

We assume that the heterogeneity function A is C^1, Lipschitz in both arguments, and uniformly bounded above and below by positive constants.

In the proof of Theorem 1, the equilibrium is static, hence the distribution μ is fixed. We therefore assume $\mu \in \mathscr{P}(\mathbb{T}^2) \cap C^1(\mathbb{T}^2)$ and that it is uniformly bounded away from zero and infinity. Under these hypotheses, the structure of the existence proof is unchanged: one simply adapts the fixed-point operator and the bounding constants. Specifically, replace the operator f by

$$f[w](x) \ = \ \frac{1}{\mu(x)} \int_{\mathbb{T}^2} \frac{A(\mu,x)\,\tau(x,y)\,w(y)}{\int_{\mathbb{T}^2} A(\mu,z)\,\tau(z,y)\,w(z)^{1-\sigma}\,dz}\, d\mu(y),$$

and adjust the associated upper and lower bounds accordingly.

For the proof of Theorem 2 (the dynamic problem), a bit more work is needed. Since Theorem 1 assumed μ is uniformly bounded above and below, we now require a priori bounds on μ. These can be obtained from the Fokker–Planck equation provided the agents' velocity field is bounded. To this end, we slightly constrain the dynamic problem by imposing an L^∞ bound on admissible controls. The individual control problem becomes

$$\max_{\alpha\in S} \mathbb{E}\left[\int_0^T e^{-\rho t}\big(V(X_t,w(t)) - c(\alpha_t)\big)\,dt\right],$$

where, for instance,

$$S = \Big\{(\alpha_t)_{t\in[0,T]} \text{ progressively measurable } : \ |\alpha_t| \le \bar{\alpha} \text{ a.e. on } [0,T]\Big\}.$$

With this uniform bound on the drift in the state dynamics (together with $\nu > 0$ diffusion on $\mathbb{T}^2$), standard parabolic estimates yield time-uniform upper and lower bounds on the solution μ to the Fokker–Planck equation (Aronson (1967)). This delivers the required a priori control on μ and allows the dynamic existence argument to proceed as in the baseline case, with the fixed-point maps and constants adapted to the perfect-competition kernel above.

B.3 Infinite horizon problem

Another natural extension is to consider the infinite time horizon problem. The definition of an equilibrium in this case is:

Definition 7. *We say that a flow of distributions $\mu \in C^{1,2}(\mathbb{R}_+\times\mathbb{T}^2,\mathbb{R})$ is a mean-field game equilibrium if there exists a value function $u \in C^{1,2}(\mathbb{R}_+\times\mathbb{T}^2,\mathbb{R})$ and a trajectory of wages $w \in C(\mathbb{R}_+\times\mathbb{T}^2,\mathbb{R})$ such that the following equations are satisfied in the classical sense:*

$$-\partial_t u - \nu\Delta u - \frac{|\nabla u|^2}{2c_0} + \rho u = V(x,w(t),\mu(t)), \tag{38}$$

$$\partial_t \mu - \nu\Delta\mu + \frac{1}{c_0}\operatorname{div}(\mu\nabla u) = 0, \tag{39}$$

$$\int_{\mathbb{T}^2} \frac{\tau(y,x)^{1-\sigma} w(t,y)}{\int_{\mathbb{T}^2}(w(t,z)\tau(y,z))^{1-\sigma}\mu(t,z)dz}\mu(t,y)dy = w(t,x)^\sigma, \tag{40}$$

completed with $\mu(0) = \mu_0$ and such that for any t

$$\int \mu(t,x)dx = 1 \ \textit{and} \ \int w(t,x)d\mu(t,x) = 1.$$

From Theorem 2 it is possible to deduce

Theorem 5. *There exists at least one mean-field game equilibrium in the sense of Definition 7.*

Proof. The proof is based on the Arzelà-Ascoli diagonalisation argument.
Let us fix $T_n = n$ for any $n \in \mathbb{N}^*$. For any n, let us note by (u_n,μ_n,w_n) a mean-field game equilibrium

associated to the game when the terminal cost is 0 and the time horizon is T_n. We prolong (u_n, μ_n, w_n) for any $t > T_n$ by

$$u_n(t) = u_n(T_n), \quad \mu_n(t) = \mu_n(T_n), \quad w_n(t) = w_n(T_n).$$

Fix any $N \in \mathbb{N}^*$. From the estimates established in the proof of Theorem 2 and Arzélà-Ascoli's theorem, we deduce that the sequence (u_n, μ_n, w_n) is compact when restraining the time interval to $[0, T_N]$. Let us consider a subsequence of (u_n, μ_n, w_n), denoted in the same way for ease of notation, that converges to a triplet (u^N, μ^N, w^N). Repeating the process when the time horizon is $2N$, $3N$, ect. we deduce that there is a subsequence, also noted (u_n, μ_n, w_n), and a triplet (u, μ, w) such that

- u_n converges to u in $C^{0,1}$ on every compact.
- μ_n converges to μ in $C^0(\mathbb{R}_+, \mathscr{P}_2(\mathbb{T}^2))$ on every compact.
- w_n converges to w in C^0 on every compact.

Therefore, (u, μ, w) is a weak solution of (38)-(40): the Fokker-Planck equation is satisfied in the distributional sense, the HJB equation is satisfied in the viscosity sense, and the static equation holds pointwise in Ω. However, the estimates developed in the proof of Theorem 2 allow us to use a bootstrap argument to increase the regularity of u and μ leading to the result. □