Manuscript version: Author’s Accepted Manuscript
The version presented in WRAP is the author’s accepted manuscript and may differ from the published version or Version of Record.

Persistent WRAP URL:
http://wrap.warwick.ac.uk/174634

How to cite:
Please refer to published version for the most recent bibliographic citation information. If a published version is known of, the repository item page linked to above, will contain details on accessing it.

Copyright and reuse:
The Warwick Research Archive Portal (WRAP) makes this work by researchers of the University of Warwick available open access under the following conditions.

Copyright © and all moral rights to the version of the paper presented here belong to the individual author(s) and/or other copyright owners. To the extent reasonable and practicable the material made available in WRAP has been checked for eligibility before being made available.

Copies of full items can be used for personal research or study, educational, or not-for-profit purposes without prior permission or charge. Provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way.

Publisher’s statement:
Please refer to the repository item page, publisher’s statement section, for further information.

For more information, please contact the WRAP Team at: wrap@warwick.ac.uk.
The Role of Space, Density and Migration in Social Dilemmas

Jacques Bara
University of Warwick
Coventry, United Kingdom
jack.bar@warwick.ac.uk

Fernando P. Santos
University of Amsterdam
Amsterdam, The Netherlands
f.p.santos@uva.nl

Paolo Turrini
University of Warwick
Coventry, United Kingdom
p.turrini@warwick.ac.uk

ABSTRACT
Cooperation in multi-agent systems often entails a social dilemma. Cooperators pay a cost to improve public goods whereas defectors free-ride, reaping benefits without incurring any costs or even producing public bads. Much attention has been devoted to understanding cooperation in populations where agents interact with random peers (well-mixed), interact over complex networks, or interact in fixed spatial positions. In spatial settings with mobile agents, however, the effects of cooperation are circumscribed to arbitrary neighbourhoods and the stability of cooperation depends on individuals’ capacity to move between sites and form dense clusters. In this paper we study spatial public goods games in which agents either pollute (defectors) or clean (cooperators) their local area and can migrate to empty sites within range. We ask whether migration promotes cooperation and reduces the negative impacts of defection. Analytically and through agent-based simulations, we show that migration ultimately reduces the pollution felt per-capita in at least two ways: 1) polluters encourage eco-friendly neighbours to migrate away, eventually clustering with other cooperators 2) migration stabilises cooperation in dense population scenarios. Our results reveal a complex interaction between migration and density as key factors to promote cooperation in spatial social dilemmas.

KEYWORDS
Public Goods Games; Emergence of Cooperation; Agent-based Modelling and Simulation; Game Theory

ACM Reference Format:

1 INTRODUCTION
Social dilemmas pervade multi-agent systems [2, 10, 11, 15, 16, 21, 23, 25, 30, 34–36]. The challenges of such dilemmas have been captured, in a stylised way, through the well-known Prisoner’s Dilemma Game and, its large-scale counterpart, the Public Goods Game. In the latter, agents in a group of size $N$ can either cooperate — contributing $c$ of their wealth to a common pot — or defect — free-riding in the contributions of others. This interaction results in a social dilemma, as cooperation is socially desirable yet, from the individual point of view, defection maximises an agent’s payoff. The same dilemma ensues if cooperators are required to pay a cost to prevent producing a public bad or to reduce the chances of a future catastrophe, as it is the case in reducing urban pollution or solving climate change crises [6, 17, 18, 28, 29].

In real-world scenarios, dilemmas of cooperation are often embedded in space – one paradigmatic example is polluting (defect) or cleaning (cooperate), which can be costly yet beneficial for individuals in a local area. Spatial settings can fundamentally impact the emergence and effects of cooperation: first, the effects of cooperation are felt on local neighbourhoods and their long-term consequences depend on how individuals are spatially assorted; second, even considering fixed agents (an assumption that we relax in this paper), cooperation can spread in spatial settings as they cluster together and benefit from a relative advantage compared with defectors [19]; finally, in the specific case of urban pollution, spatial proximity can mediate the diffusion of environmentally friendly technologies, as some empirical examples (chiefly, the imitation of household PV cell adoption between neighbours [27]) illustrate.

Besides introducing a definition of locality, spatial environments suggest that agents can move. While some models of spatial public good games have been analysed in the literature, the role of movement in fostering or hindering cooperation among agents acting in space is still not well understood. Here we develop a new model to understand how agents’ mobility can impact cooperation dynamics.

Our Contribution. In this paper we study the conditions for which cooperation emerges on spatial public good games. Cooperators clean their immediate environment, contributing to increasing the payoffs of other agents, and defectors that pollute their surroundings, worsening their neighbours’ conditions. Agents move around the space based on the perceived levels of pollution, migrating within a radius $M$, to an empty site with the least pollution.

We study how cooperation depends on migration range ($M$) analytically in small cases, finding the regimes in which a social dilemma arises and the migratory patterns that stabilise cooperation in dense cities. We then supplement the theory with simulations of the agent-based model\footnote{All code, data and other supplementary materials (including animations) are openly available at https://github.com/JBara97/Pollution-Game.}: first, with fixed strategies to isolate the effects of migration; second, with a behavioural model of imitation for a fixed size city; third, varying the density of the city we identify a transition from sparse to dense (rural to urban). We find that migration in dense enough areas allows cooperators to cluster together, thus reducing the per-capita pollution and stabilising cooperation in the long term evolution of the coevolutionary process.

In Section 2 we discuss related work after which we formalise our model in Section 3. In Section 4 we present our theoretical analysis which we then complement with computational experiments in Section 5. Finally we discuss the results and conclude in Section 6.
2 RELATED WORK

As in our model, many spatial games [5, 19, 20, 22, 24] consider the underlying space as a discrete periodic 2D lattice (also known as a toroidal lattice). Most previous works in spatial social dilemmas, however, assume that each location contains a fixed agent and dynamics result from strategy evolution. In settings with multiple agents in a site, it is often assumed that games are played within a site and not between sites [1]. As in [13], we consider the existence of empty sites and focus on the movement of agents. Our work adds the analysis of compound agents - the agglomeration of agents into collectives - and studies the effects of movement on their formation and on the agents’ social welfare.

Some related contributions assume space to be continuous - reasonable for sufficiently large distances or population sizes - and focus on dynamical-systems, partial differential equations (PDE) and pattern formation. The work of [33] considers cooperators and defectors as species, that reproduce depending on the average payoff of the species, and that move via diffusion; an extension of which considered directed migration [8]. In these settings cooperators are universally attractive, defectors are repulsive, and the speed with which agents move depend on, for example, whether a defector is moving towards another cooperator.

Regarding the mechanisms for migration, some previous work considered that migration can be caused by success-driven agents [13], which move to a site where, under fictitious play, they would get a higher payoff. Despite some of the criticisms – notably experiments showing that humans are typically conditional cooperators [4, 7, 9] who do not look at payoffs and may show other forms of cooperation that punish misbehaviour [14] – this assumption remains widely used and appealing. Previous agent-based simulations studying cooperation in spatial settings with mobile agents have also found that mobility, by allowing cooperators to cluster together, stabilise pro-social actions [12, 13, 26], but their interaction model ignores the long range effects of undesirable behaviour and, perhaps for the lack of theoretical predictions, the study overlooks the role of spatial density, which we find to have a fundamental impact on the emergence of cooperation.

Relevant, although based on different interaction rules, is also Schelling’s segregation model [32], in which agents of one type prefer to be surrounded by her own type avoiding the other. Sites are limited to at most a single agent, each of whom will move to a random empty site if their current neighbourhood is not within their tolerance level. From such simple mechanisms emerges the complex phenomenon of segregation. Here, we similarly find the emergence of a socially desirable outcome, pollution reduction, via the assortment of agents into compounds. The dynamic network analogue of our work - networked agents playing uniform Prisoner’s Dilemmas who can select partners - finds similar cooperation-cores and defector-peripheries forming from the coevolutionary process, despite the very different game and update rules [3].

3 SPATIAL SOCIAL DILEMMAS

3.1 Agents, Interaction and Movement

We consider a 2D periodic lattice of size \( L \times L \) populated by \( N \leq L^2 \) agents. Each site can only be occupied by up to one agent. A lattice site, regardless of having any inhabitant, is denoted by its 2-vector coordinate \( r \in \{1, \ldots, L\}^2 \). Within this space, agents play a spatial social dilemma, which, for clarity, we frame as a pollution dilemma.

Let the level of pollution at a site \( r \) be denoted \( P(r) \); furthermore, let the per-capita pollution (PCP) - pollution averaged over occupied sites, recalling that these coincide with \( N \) - be denoted \( \hat{P} \). PCP depends not only on their strategies but also on their spatial assortment. Agents have two strategies available to them: defection (D) i.e. pollution and cooperation (C) i.e. cleaning. Each defector at location \( r_d \) pollutes each location \( r \) according to \( P(r) = \frac{r}{r_d^2} \), where \( r = \|r - r_d\| \) is the shortest L2 distance from the defector. Moreover, let such pollution be dubbed a pollution cloud or simply a cloud. This allows us to discuss the size and shape of polluted areas succinctly; the cloud of a single defector is spherical and has a \( r^{-2} \) profile.

As the intensity of pollution decreases at further away sites, it is also natural to assume a threshold below which pollution levels are negligible. To this end let us say the pollution cloud from a single defector can only reach a distance \( R > 1 \) away. Finally, to deal with the singularity at \( r = 0 \), i.e. at the source of pollution, we say the pollution there is simply \( P(0) = 1 \). Explicitly, the pollution at site \( r \) due to a defector living at site \( r_d \) is

\[
P_d(r) = \begin{cases} 
\frac{1}{r_d^2} & \text{for } r \leq 1 \\
0 & \text{otherwise.}
\end{cases}
\]

where \( r = \|r - r_d\| \).

We conceptualise cooperators as agents who clean pollution – for example, the neighbour in lockdown who embraced gardening or picked up litter every week. Formally a cleaner \( c \) at site \( r_c \) removes a fixed quantity \( \phi \) of pollution from all sites within a distance 1. In other words the pollution at site \( r \) due to a single cleaner at \( r_c \) is given by the following.

\[
P_c(r) = \begin{cases} 
-\phi & \text{for } r \leq 1 \\
0 & \text{otherwise.}
\end{cases}
\]

In the case of multiple agents pollution is simply additive, such that if \( P_i(r) \) is the pollution produced/cleansed by agent \( i \) at \( r \), then the total pollution at \( r \) is \( P(r) = \sum_i P_i(r) \). We allow for negative pollution levels as agents look at the relative, not absolute, levels; the behaviour would be equivalent for some constant non-zero background level.

Cooperation and pollution felt are costly; we consider that every agent \( i \) pays \( E_i \), which can result from a combination of perceived pollution and financial assets. Specifically let \( f \) be the fee that an agent pays to clean and \( g \) the gain of an agent that defects. As such an isolated agent – at a distance \( r > R \) from all other agents – will have an expense of \( E_i = f - g \) if she cooperates and \( E_i = 1 - g \) if she defects. Defection is dominant so long as \( f + g > 1 + \phi \), which, as we will see later is also a sufficient condition for the unique Nash Equilibrium strategy to be defection, at least in the two-player game.

The co-existence of cooperators and defectors will determine the pollution levels felt at each site. This will, in turn, impact agents’ movement. We assume that an agent’s movement is a best-response: that is migrate to an empty site within a radius of \( M \geq 0 \), henceforth named the migratory distance, with the lowest level of pollution so long as it has strictly less pollution than her current site, with ties
We consider a deterministic short-distanced imitation model in which an agent $i$ at position $r_i$ imitates the strategy of an agent $j$ at site $r_j$ who has the minimal expense $E_j$ of all agents within a distance $M_i$. Explicitly for $\sigma_i(t)$ the strategy of $i$ at time $t$, agent $i$’s strategy at the next time step is $\sigma_i(t+1) = \sigma_j(t)$ where

$$j = \arg \min_k E_k$$

and where $M_i$ is the *neighbourhood radius*, the furthest distance an agent can be in order to be imitated. In order to reduce complexity we henceforth set $M_i = 1$ since, in principle, an agent $i$ may act in the following sequence: 1) see a best performing agent $j$ a distance $M$ away, 2) migrate to a location next to her and 3) imitate $\sigma_j$. Thus, in effect, it does not matter if $M_i > 1$, since migration allows an agent to reach a far away alter and subsequently imitate them.

### 4 THEORETICAL ANALYSIS

In this section we present two analytic frameworks. First, a game-theoretic analysis on the $N \in \{2, 3\}$ one-shot game (effectively the first part of a single time-step of the whole game) for static agents in Section 4.1. Second, assuming fixed strategies, we analyse migratory patterns arising out of pollution for $N \in \{1, 2, 3\}$ in Section 4.2. Third, we extend the fixed-strategy analysis to the stable structures that arise in larger $N$-agent cities in Section 4.3 and finally the mechanisms by which migration reduces pollution in Section 4.4.

#### 4.1 Static Agents

Two Agents. For an isolated agent $i$, the arrival of agent $j$ complicates her strategic choices not only as she now has to consider $\sigma_j$ but also consider how far away $j$ is. In Table 1 we first show the expense $E_i$ of agent $i$, given her opponent’s strategy $\sigma_j$ if $j$ is a distance $r$ away from her. From these expense values, we arrive at Figure 1 which shows the phase diagram for the financial parameters $f + g$ and the inter-agent distance, highlighting the regions defined by two analytic thresholds: one which separates when defection is dominant and the other which defines when cooperation-cooperation is the social optimum.

<table>
<thead>
<tr>
<th>$(\sigma_j, \sigma_i)$</th>
<th>Clean</th>
<th>Defect</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clean</td>
<td>$f - \phi$</td>
<td>$-g + 1$</td>
</tr>
<tr>
<td>Defect</td>
<td>$f - \phi$</td>
<td>$-g + 1$</td>
</tr>
</tbody>
</table>

When $f + g < 1 + \phi$, we are in the regime of region A in Figure 1, that is the dominant strategy - and therefore the two-player Nash Equilibrium (NE) - is to clean. This coincides with the analysis of the single agent game where an isolated agent’s best choice depends on the difference in $f + g - (1 + \phi)$. If on the other hand $f + g > 1 + \phi$ then the dominant strategy - and the two-player NE - is to defect, as in regions B and C.

The social optimum is the set of strategies that, in this case, minimises the total expense of both agents. In region C, defection is the social optimum since the financial benefits more than outweigh the environmental factors, which notably depends on the inter-agent distance $r$. For $r \leq 1$ the choice for one agent to clean impacts both agents, while for $r > 1$ the choice to defect affects both agents. For $r > R$ the game collapses to two single-agent games each of which has a threshold of $f + g = 1 + \phi$ above which defection is dominant and below which cleaning is dominant.

Below the red dotted line that defines region C, in regions B and A, the SO is to clean. Only in region B, therefore, where the NE is to defect and the SO is to clean, does a social dilemma arise. Given a cloud size of $R$, a social dilemma is guaranteed to occur for two nearby agents if $1 + \phi < f + g < 1 + \phi + R^{-2}$. Notice finally that even when cooperators clean nothing, $\phi = 0$, region B can still exist due to precisely the profile of a pollution cloud given by Equation 1. In fact for any non-zero cloud profile $P_d(r) \neq 0$, region B always exists and is defined by $1 + \phi < f + g < P_d(r) + [\pi r \leq 1]$ where $[\cdot]$ is an indicator function which is 1 when its argument is true and 0 otherwise.

Three Agents. The two agent-case is, though complex, easily tractable and intuitively simple to understand - as shown by Figure 1 - simply by considering the parameters $f, g$ and $\phi$ as well as the inter-agent distance. The three-agent case, on the other hand, is significantly complicated by spatial arrangement. Even if all three agents are restricted to a single contiguous cluster (that is all
agents have at least one immediate neighbour) there are 12 cases to consider as we do in Table 2. In this case both Nash Equilibria and social optimum are likely to be mixed strategies.

Some work has been done on the theory of Nash Equilibria in three-player partially asymmetric zero-sum games [31] however in more general cases calculations of mixed Nash Equilibria become computational rather than analytically tractable. In our case, for example, although a straight line of three agents plays a three-player, partially asymmetric game, it is not zero-sum. When in the alternative assortment (the L-shape) the three agents no longer play a partially asymmetric game.

### Table 2: Three-player game. Expense $E_i$ for each of the three players, given their assortment and their strategies $\sigma_i$.

<table>
<thead>
<tr>
<th>Assortment $(\sigma_1, \sigma_2, \sigma_3)$</th>
<th>$E_1$</th>
<th>$E_2$</th>
<th>$E_3$</th>
</tr>
</thead>
<tbody>
<tr>
<td>(C, C, C)</td>
<td>$-2\phi + f$</td>
<td>$-3\phi + f$</td>
<td>$-2\phi + f$</td>
</tr>
<tr>
<td>(D, C, C)</td>
<td>$1 - \phi - g$</td>
<td>$1 - 2\phi + f$</td>
<td>$0.25 - 2\phi + f$</td>
</tr>
<tr>
<td>(C, D, C)</td>
<td>$1 - \phi + f$</td>
<td>$1 - 2\phi - g$</td>
<td>$1 - \phi + f$</td>
</tr>
<tr>
<td>(D, D, C)</td>
<td>$2 - g$</td>
<td>$2 - \phi - g$</td>
<td>$1.25 - \phi + f$</td>
</tr>
<tr>
<td>(D, C, D)</td>
<td>$1.25 - \phi - g$</td>
<td>$2 - \phi + f$</td>
<td>$1.25 - \phi - g$</td>
</tr>
<tr>
<td>(C, D, D)</td>
<td>$2.25 - g$</td>
<td>$3 - g$</td>
<td>$2.25 - g$</td>
</tr>
<tr>
<td>(C, C, D)</td>
<td>$-2\phi + f$</td>
<td>$-3\phi + f$</td>
<td>$-2\phi + f$</td>
</tr>
<tr>
<td>(D, C, D)</td>
<td>$1 - \phi - g$</td>
<td>$1 - 2\phi + f$</td>
<td>$0.25 - 2\phi + f$</td>
</tr>
<tr>
<td>(C, D, D)</td>
<td>$1 - \phi + f$</td>
<td>$1 - 2\phi - g$</td>
<td>$1 - \phi + f$</td>
</tr>
<tr>
<td>(D, D, D)</td>
<td>$2 - g$</td>
<td>$2 - \phi - g$</td>
<td>$1.5 - \phi + f$</td>
</tr>
<tr>
<td>(D, C, D)</td>
<td>$1.5 - \phi - g$</td>
<td>$2 - \phi + f$</td>
<td>$1.5 - \phi - g$</td>
</tr>
<tr>
<td>(C, D, D)</td>
<td>$2.5 - g$</td>
<td>$3 - g$</td>
<td>$2.5 - g$</td>
</tr>
</tbody>
</table>

#### 4.2 Movement with Few Fixed-Strategy Agents

Here we analyse the migratory patterns that form assuming fixed agent-strategies. Equivalently, the regime of fixed agent-strategies can be thought of as the regime of rapid migration and immensely slow strategic change. Regardless, as strategies are fixed, in this subsection we only deal with the pollution felt by an agent, rather than her full expense. In particular we use the per-capita pollution (PCP) $\hat{P}$ as a measure of city-wide pollution.

**Single Agent.** Consider an isolated agent, entirely by herself with no other agents in the city. If she cooperates then her pollution is simply $-\phi$ and, importantly, she would be at a local minimum – all sites have a pollution of $0$ except hers and her immediate neighbouring sites, which have a pollution of $-\phi$. As such her optimal migratory strategy would be simply to remain stationary.

On the other hand were she a defector then she would experience a pollution of $1$ and thus inhabit a local maximum for pollution. The migratory distance now plays an important factor in her optimal strategy: if $M \leq 1$ she would be unable to see past sites immediately next to her, so the optimal strategy would be to remain. If instead $M > 1$ then she is able to move beyond her immediate neighbourhood to find greener pastures. However she is now cursed to endlessly roam since, wherever she ends up, her immediate neighbourhood will be the peak of pollution hence she will move away. Her overall movement is thus Brownian Motion: for $1 < M < R$ she will move to a random empty site exactly a distance $M$ away and for $M > R$ she will move to a site that is a distance $R < m \leq M$ away. In either case her direction of movement is random.

**Two Defectors.** Consider two defectors, 0 and 1, at sites $r_0$ and $r_1$ respectively. If the inter-agent distance is less than a cloud size, $r \leq R$, then both would be weakly repelled by one another. Concretely, if an agent looks to all sites within a distance $m$, the isopleth with the least amount of pollution is the arc of radius $m$ that is furthest from the opposing defector (see, as an example, Figure 1 in Supplementary Material). As such, assuming agent 0 is fixed in place, if $M < R$ the best locations lie precisely on this isopleth. Otherwise for $M \geq R$ then she would jump to any point outside of the clouds within range. As the two defectors entirely avoid one another, so long as the space is large enough, then each will feel the effect of only themselves – in other words $\hat{P} = 1$.

**Two Cooperators.** Regardless of $r$, the optimal choice for either cooperator would be to remain stationary. If $r = 1$, i.e. the two are next to one another, then they already achieve optimal pollution at $\hat{P} = -2\phi$, else if $r > 1$ then they are still at optimal pollution at $\hat{P} = -\phi$.

There is a singular degenerate case when one of the cooperators will move. Consider a cooperator $c_0$ at $r_{c_0} = (x, y)$ and another $c_1$ at some $r_{c_1}$. If $c_1$ is diagonally next to $c_0$ - that is $r_{c_1} = (x \pm 1, y \pm 1)$ then the site that is coadjacent to both agents has a pollution level of $-2\phi$ despite each only feeling $-\phi$. Either agent will move to this space and stabilise the system with $\hat{P} = -2\phi$.

**Cooperator and Defector.** For $r > R$ the problem reduces to the single agent case, that is the cooperator remains stationary while the defector randomly migrates as in Brownian Motion. In the short term the defector will be unaffected by the cooperator and vice-versa, however over a long enough period of time, $t = O(r/M)$, the defector will have drifted close enough to the cooperator to be able to see its effects (i.e. be within distance $r \leq M$ of it). Once this occurs the defector will be attracted towards the cooperator.

If the cooperator is unaffected by the defector, that is $r > R$, then in one step the defector will move (at random) into one of the neighbouring sites of the cooperator, while she remains stationary. From this point the distance between them will be less than a cloud size $r < R$ and a game of cat-and-mouse commences, regardless of scheduling and whether they move simultaneously or sequentially. The cooperator will keep trying to run away from the defector but the defector keeps chasing her down, since they both share the same migratory distance $M$.

### Table 3: Behaviour of a 1-cooperator-1-defector system for different values of $\phi$ and $M/R$.

<table>
<thead>
<tr>
<th>$\phi$</th>
<th>$M/R &lt; 1$</th>
<th>$M/R \geq 1$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\phi &lt; 0.25$</td>
<td>$M$-drift</td>
<td>Teleport</td>
</tr>
<tr>
<td>$\phi = 0.25$</td>
<td>$M$- and 1-drift</td>
<td>Teleport and 1-drift</td>
</tr>
<tr>
<td>$\phi &gt; 0.25$</td>
<td>1-drift</td>
<td>1-drift</td>
</tr>
</tbody>
</table>

The cat-and-mouse will carry on forever, but whether or not the defector can ever occupy an adjacent site to the cooperator depends on the ratio of migration to cloud size $M/R$. If $M/R \geq 1$ then the cooperator can certainly escape the cloud and will thus spend some of its time being stationary.

This long-distance jump away from the defector (though still within a distance $M$), hence dubbed teleporting, causes the cooperator to 1) escape the cloud(s) of the pursuing defector(s) and 2) to
land in an empty site that is a (mostly) random direction from the
defector. Specifically for the latter, by a (mostly) random direction
we mean one that is not necessarily along the direction \( r_\co - r_\def \). On the other hand if \( M/R < 1 \) then she will always move as
the defector is guaranteed to at least keep her inside the cloud.
Moreover given the cooperator \( c \) is stuck inside the cloud, i.e. all
sites within distance \( M \) of her are within the cloud, then the site
with the least amount of pollution will be a distance \( M \) away exactly
along the direction of \( r_\co - r_\def \). Such movement, where there is a
unique site some distance \( m \) away that has minimal pollution, we
dub an \( m \)-drift. In particular, as the defector has the same migratory
distance, this will cause the two to drift endlessly along the direction
of \( r_\co - r_\def \). Table 3 showcases this for different values of \( \phi \) and \( M/R \).

Let us also consider the pollution levels, in particular the per-
capita pollution before and after the two agents have merged. When
sufficiently far away from one another \( r > R \) then \( \hat{P} = (1 - \phi)/2 \)
whereas once they are next to one another then \( \hat{P} = 1 - \phi \). In other
words if the cleaning rate is insufficient (\( \phi < 1 \)) then pollution
is overall increased when the two agents are adjacent. If instead
\( \phi > 1 \) pollution per-capita has significantly decreased. This effect
is important in the large \( N \) regime where, as we will see in Section
5.2, migration significantly decreases pollution over time for \( \phi > 1 \)
and somewhat increases pollution otherwise. Table 4 summarises
the \( N = 2 \) systems in terms of per-capita pollution.

### Table 4: Per-capita pollution \( \hat{P} \) for the \( N = 2 \) system with
different number of deflectors \( D \). The left column gives values
of \( \hat{P} \) if the two agents are sufficiently far away \( r > R \) while the
right column gives values for when they are adjacent \( r = 1 \).

<table>
<thead>
<tr>
<th>( r = 1 )</th>
<th>( r &gt; R )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( D = 0 )</td>
<td>( -2\phi )</td>
</tr>
<tr>
<td>( D = 1 )</td>
<td>( 1 - \phi )</td>
</tr>
<tr>
<td>( D = 2 )</td>
<td>( 2 )</td>
</tr>
</tbody>
</table>

### 4.3 Stable Clusters

In this section we analyse small configurations of agents which are
theoretically stable; we dub **compound agent** or **agent cluster**, as a set
of agents spatially near one another that behaves, effectively, as one
collective. These **compound agents** are critical in how migration can
affect pollution. In particular the stability, formation and clustering
of compound agents allow us to understand and predict macro-scale
phenomenon even for large \( N \) or \( L \). For snapshots of their formation see
Figure 2 or see the GIFs in the Supplementary Material.

**Drifters.** For sparse cities any local areas with very few agents
will exhibit behaviour similar, if not identical, to the two agent case,
due to the short-range nature of interactions. Consider a cooperator
\( c \) with several defectors in hot pursuit. If \( c \) drifts - either because they
cannot escape the defectors’ clouds and so moves to a neighbouring
site or because \( \phi \) is sufficiently high that a neighbouring site to
\( c \) has less pollution than all other sites - then the defectors will
equally drift to the same relative position prior to the move. In
this way the set of \( c \) and her defectors form a compound agent that
collectively drifts, as exemplified by the two-agent case with \( D = 1 \).

**Cooperative Clusters.** Consider a cooperator \( c \) that has \( d \leq 4 \)
defectors as immediate neighbours (\( |r_\co - r_\def| = 1 \)). If \( \phi > d \) then
regardless of \( M/R \), the site of minimal pollution for \( c \) will either be a)
the empty site(s) immediately next to her if \( d < 4 \) or b) \( r_\co \) if
\( d = 4 \), i.e. all the sites that she cleans are full. In case a) her best
option would be to 1-drift/tele-drift - a teleport where all candidate
sites are within distance 1 - into one of the empty sites that she
leans, after which the defectors will follow suit. This compound
agent of \( c \) and the \( d \) defectors now moves collectively via 1-drifts
or tele-drifts, as above. On the other hand, in case b), the optimal
strategy for \( c \) is to become stationary as there are no sites within
range that has strictly less pollution than \( r_\co \). The defectors’ optimal
strategy would subsequently be to stay in place. In other words this
compound agent, a cooperator entirely surrounded by defector, is
fully stabilised, in arrangement and movement, due to the defectors.

Consider now several drifters in the lattice. When two drifters,
collide two (or more) constituent cooperators may end up occupy-
ing adjacent sites. In this case both cooperators feel the cleansing
effects of themselves and of each other. As such the two cooper-
ators will, with very high probability, be stabilised by one another
to the extent that no defectors could knock them away from one
another. To see this consider, the case of two drifting clusters one
consisting of a cooperator \( c_1 \) surrounded by \( d_1 < 4 \) defectors the
other consisting of cooperator \( c_2 \) surrounded by \( d_2 < 4 \) defectors.
A sufficient condition for cooperators stability is \( \phi > (d_1 + d_2)/2 \).
For a stricter lower bound, i.e., a necessary condition, one would
need to consider the specific spatial arrangement. For example, for
\( d_1 = d_2 = 3 \) each cooperator has a pollution level of \( P = 4.25 - 2\phi \).
therefore the necessary condition would be \( \phi > 2.125 \).

Depending on \( \phi \) and the spatial configuration, the collision will
genERICALLY be enough to stabilise both cooperators. For a site \( r \) to
be more appealing than either \( c_1 \) or \( c_2 \) ’s current site there needs to be
at least 2 other cooperators, \( c_4 \) and \( c_5 \), neighbouring site \( r \). But
should such a site exist, then there is a reasonable chance either \( c_4 \)
or \( c_5 \) will immediately occupy it. In other words, having such an
attractive site existing in range and for no other agents to move in
first is generally unlikely.

Finally, taking the inverse perspective, we see that for a fixed
\( \phi = a \) then \( d < a \) defectors can stably surround a single cooperator.
Stability comes from the cooperator experiencing negative pollution
at her site \( P(r_\co) = d - a \), and thus having no incentive to teleport

#### Figure 2: Snapshots at times \( T \) of a city of size \( L = 30 \) har-
bouring \( N = 50 \) agents. We see the clustering of cooperators,
surrounded by a plethora of defectors, i.e., the formation of
compound agents.
away, but may still (tele)drift. Defectors similarly have \( P(r_d) < d-a \), as they are further than 1 space away from one another, and are therefore also stable.

### 4.4 Beyond Small Agent Systems

Our theoretical analysis of small fixed-strategy systems suggests two hypotheses on mechanisms for pollution reduction, which we will substantiate for large \((N = L = 50)\) systems using agent-based simulations in Section 5.2.

**Hypothesis 1.** Under fixed strategies, migration reduces PCP.

This does so by a two-fold mechanism. One, that defectors kick otherwise stationary cooperators out of complacency and cause them to move sites. We can clearly see this behaviour already from the two-agent and single-agent cases. Two, as cooperators move and are pushed around by defectors, they will eventually collide with one another thus forming stable clusters. Because of this fact, all the constituents generally experience lower amounts of pollution. These compounds will have a cooperator core and a defector periphery or border (see Figure 2). The defectors on the outskirts feel the effects of one or two cooperators and a minimal amount from their defective neighbours, while the core enjoy the effects of multiple (up to 4) cooperators. The cooperators stabilise one another even with these defectors who attempt to penetrate the core. As a result per-capita pollution drops.

**Hypothesis 2.** Migratory defectors destabilise single cooperators into drifters. The collision of multiple drifters cause stable clusters to form. All agents in a compound experience lower pollution, particularly the cooperators in the core, and hence PCP drops.

## 5 Experimental Analysis

### 5.1 Experimental Setup

Below we present the results from three different experiments, each with slightly different parameter setups. All simulations, however, share the same environmental factors \((R, f, g)\) of pollution with clouds of size \( R = 6 \) and a cleaning rate of \( \phi = 5 \), for a city of fixed size \( L = 50 \) running for \( T = 201 \) timesteps over multiple runs. The code can be found in the Supplementary Materials.

**Fixed Strategies.** First, we fix strategies to isolate the effects of migration on pollution and to validate our hypotheses. To do so we look at a city of \( N = 50 \) agents across a variety of migratory distances \( M \) and initial number of defectors \( D \) (alongside \( C = N - D \) cooperators) across 50 runs. In particular, as we are only concerned by migration, the financial parameters \((f \text{ and } g)\) bare no effect in this case. In doing so we can verify both hypotheses and build an intuition to how migration impacts a city absent strategic evolution.

**Cooperation Stability.** Second, we allow strategies to evolve, again with \( N = 50 \) agents but focusing on initial conditions with a majority of defectors, \( D = 30 \). Financial parameters are now set to \( f = g = 3.5 \) such that \( 1 + \phi < f + g < 2 + 2\phi \), in other words agents next to one another feel a social dilemma (region B in Figure 1) while those further away benefit from defection always (region C). In order to explore the phase space of possible configurations more efficiently we further introduce a mutation rate \( \epsilon = 0.01 \) whereby an agent has probability \( \epsilon \) to flip her strategy and \( 1-\epsilon \) to imitate a nearest neighbour by the mechanism outlined in Section 3.2. To compensate for the noise we ran the simulations over 300 runs.

**Density Dependence.** Third, we investigate the dependence on density by varying the number of agents \( N \in \{5, 10, \ldots, 100\} \) while fixing the city size \( L = 50 \). The number of initial defectors is always fixed \( D = 0.6N \). As we still vary \( M \) to compensate for computation time we reduce runs to 200 per set of parameters.

### 5.2 Fixed Strategies

In absence of genuine units of pollution, such as parts per million (ppm) of CO₂, we use the initial value of per-capita pollution \( \hat{P}(0) \) as a baseline to compare \( \hat{P}(t) \) against and quantify how migration affects PCP; if \( \hat{P}(t) - \hat{P}(0) > 0 \) then individuals are worse off than how they started, whereas if \( \hat{P}(t) - \hat{P}(0) < 0 \) then there is a decrease in pollution. To account for stochastic effects we report a time-average over the final 50 time-steps. For the full evolution of \( \hat{P} \) over time see Figure 2 in the Supplementary Material.

To verify Hypothesis 2, we observe all the compounds at the end of each run. Computationally, a compound is taken to be a set of agents that occupy a contiguous sub-region of the city, for example if agent 0 is adjacent to agent 1 and 1 is adjacent to 2 then \((0, 1, 2)\) is taken to be a compound. From this ensemble of clusters, we get an average number of clusters, the average size of clusters and the fraction of the population that exists inside a cluster per run.
As we see in Figure 3a there are two major factors that lead to an increase in the reduction in pollution (i.e. agents are better off than being housed at random): a decrease in \( D \) and an increase in \( M \). We also find that an increase in \( \phi \) also reduces pollution.

The first trend is fairly self-explanatory: the fewer the defectors the smaller the competition to live next to a cooperator. Equivalently, there are more cooperators and thus more clean sites available for an agent to move into. The increase in \( \phi \) has two effects: first for the same number of cooperators \( C \), a higher \( \phi \) naturally removes more pollution by definition; second is its effect on the stability of clusters (Section 4.3) allowing for larger clusters to form.

This thus leads us naturally to the trends due to \( M \). For a start, when \( \phi \) is very low \((\phi < 1)\) migration benefits defectors far more than it benefits cooperators. The reason for this is that cooperative clusters are very difficult to stabilise for low \( \phi \). By the analysis of Section 4.3, a single cooperator can only stably neighbour \( d < \phi \) defectors, in other words only drifters are even meta-stable, in so far as they teleport together. When two such drifters collide even the double cleaning by two cooperators cannot fully stabilise the compound. As such the likelihood to form bigger clusters which are stable becomes negligible.

On the other hand, for all other \( \phi \), there is a large reduction in \( \bar{P} \) for larger \( M \) (see Figure 3 in Supplementary Material) and thus have compelling evidence for Hypothesis 1. Now, in order to verify the mechanism for this effect, in other words Hypothesis 2, we look to Figure 3b. First, we see that over half of the population over time have gravitated towards each other and formed clusters. Moreover there is a clear correlation between the migratory distance \( M \) and the percentage of agents in clusters, confirming our analysis that migration promotes clustering.

For a large enough \( \phi \) where compounds are stable, it is simply a matter of time that the disorder of random allocations organises itself into ordered and stable clusters. In this way, a higher amount of time, in the agents’ frame of reference, to stabilise.

We find that clusters grow larger and more numerous the more mobile agents are. For a large enough \( \phi \) where compounds are stable, it is simply a matter of time that the disorder of random allocations organises itself into ordered and stable clusters. In this way, a higher

\[
\text{Figure 4: Final cleaning rate after } T = 201 \text{ timesteps against different values of migration } M. \text{ The black line represents the ensemble mean while the vertical line of coloured circles represent the empirical distribution as a histogram over 300 runs. The colour and center of a circle denotes the location of the bin while size denotes the bin count.}
\]

\[
\text{Figure 5: For a range of migratory distance } M (\text{x-axis}) \text{ and number of agents } N (\text{y-axis}), \text{ a heatmap of final mean cleaning rate } C/N \text{ is plotted. For cross-sections see Figure 4 in Supplementary Material.}
\]

\[
M \text{ is akin to a higher speed of movement, or equivalently a shorter amount of time, in the agents’ frame of reference, to stabilise.}
\]

\[
5.3 \text{ Cooperation Stability}
\]

We now focus on agents that are able to migrate and imitate strategies. In Figure 4 we see for cities of \( N = 50 \) agents migration tends to increase the ensemble mean cleaning rate \( C/N \) (for the analogous distribution in per-capita expense see Figure 3 in Supplementary Material). It does so by abruptly transferring some of the mass of the defection peak at \( C/N = 0 \) to the cooperation peak at \( C/N = 1 \).

As shown by Figure 4, this abrupt transition occurs at \( M = R = 6 \) when suddenly a significant cooperative peak appears. From Section 4.2 the significance of \( M = R \) is clear: agents are now able to escape the pollution clouds of neighbouring defectors. After this threshold, however, migration becomes a double-edged sword. Cooperators can escape defectors yet defectors can chase cooperators. In particular, although stable cooperative compounds can form, further away defectors can infiltrate these compounds and potentially destroy them as proposed in Hypothesis 2. As such we see a broadening of both peaks such that not all runs end in pure cooperation nor pure defection.

\[
5.4 \text{ Density Dependence}
\]

Here we look at how the city density impacts migration and pollution. In particular we will see that for dense cities migration improves cleaning rate while in sparse cities migration facilitates more defection. For that we need a formal treatment of density.

Consider a city of size \( L \) which contains \( N \) agents, a fraction \( \delta \) of which are defectors such that \( D = \delta N \). For pollution clouds of area \( A \), the total amount of cloud area is given by \( DA \). A necessary condition for the entire city to be covered in pollution clouds, assuming no overlaps, is for \( L^2/(DA) \geq 1 \), or similarly for the number of agents to exceed some critical value \( N \geq N_* \), given below.

\[
N_* = \frac{L^2}{\delta A}
\]
In particular we use $N_c$ to define the high-density and low-density regimes as $N \geq N_c$ and $N < N_c$ respectively. As we will see from Figure 6b, this not only has a priori physical meaning but coincides with an important threshold that emerges out of the dynamics. Note that for perfectly circular clouds, the area of an individual cloud is $A = \pi R^2$ while for discretised circles $A = A_d = \pi R^2$ - there is in general no explicit form for $A_d$ given arbitrary $R$, however this can be found algorithmically. As such we can bound $N_c \in [N_c^-, N_c^+]$ where $N_c^- = L^2/(\pi R^2)$ and $N_c^+ = L^2/A_d$. For our specific physical parameters ($L = 50$, $R = 6$) we find the lower bound on $N_c$ as $N_c^- = 36.84$ (4s.f.) and the upper bound as $N_c^+ = 42.96$ (4s.f.).

Looking at Figures 5 and 6a one can see that the two density regimes do behave very differently. In low-density cities ($N < N_c$) migration reduces the cleaning rate and often leads to mass defection (black and dark regions for $N \leq 40$). In the sparse city, agents largely teleport around and, due to the comparative size of a cloud versus a cleaning area, are more likely to encounter a defector than a cleaner. In this way, cleaners are very unlikely to cluster together and thus cleaning cannot be sustained and stabilised.

In contrast, in high-density cities ($N \geq N_c$) migration improves cleaning rate by giving agents more manoeuvrability to first escape defectors and second, as a cooperator, quickly cluster with other cooperators. By speedily forming compound agents with stable cooperator cores cleaning becomes a viable strategy. Not all cities will reach 100% cooperation or defection - although these are certainly meta-stable states up to some mutations - and in those with somewhat intermediate values of cleaning rate, a compound of cleaners surrounded by defectors is a common motif.

Finally, Figure 6b shows that the value of this threshold, by which to define high- and low- density regimes, does coincide with $N_c$. In the latter we fit a sigmoid, $f_c(N) = \theta(1 + \exp(-k(N - N_0)))^{-1} + b$, to the final cleaning rate for lengthy migration $M = 10$ by ordinary least squares and infer the parameters as $(\theta, N_0, k, b) = (0.5344, 40.98, 0.08862, -0.02870)$. In particular notice that the midpoint of the sigmoid, $N_0$, lies within the bounds of our theoretical $N_c$ of [36.84, 42.96] confirming that there is a qualitative difference between the high- and low-density cities.

6 CONCLUSION

We provided a theoretical and computational analysis of migration in spatial public good games, stylised as urban pollution, showing that migration enables the formation of stable compounds - contiguous sets of agents that behave cooperatively - which in turn reduces the per-capita pollution over time. As agents imitate the strategies of their neighbours, migration leads to a form of strategic coexistence and polarisation. Coexistence is due to the formation of stable compounds, a core of cooperators who stabilise one another surrounded by a periphery of defectors. Polarisation on the other hand occurs because compounds are either destabilised into mass defection or the cores fully engulf the periphery such that all defectors flip into cooperators.

Contrarily discussions in prior literature, migration does not, alone, promote cooperation. We identified a density threshold below which migration favours defectors and above which it favours cooperators. For sparse cities, defectors are able to effectively follow cooperators preventing stable cooperative clusters to emerge. For dense cities, migration allows cooperators to quickly find one another, cluster and stabilise, thus resisting the allure of defectors.

An interesting and important line of research is to understand the policy implications of our results. From a system designer point of view our findings can suggest new urban interventions - such as the introduction of ‘parks’ as areas of fixed negative pollution, which may serve as nucleation points for cooperators (and defectors) to cluster around. Moreover in reality movement is almost always costly, and thus presents an important avenue for future work. Finally, finding suitable datasets to test the model against real-world application is an important research direction.

Figure 6: Final cleaning rate for a range of different migratory distances $M$ (left figure) and different number of agents $N$ (right figure). a) Two trend lines for a low-density ($N = 5$; orange squares) and high-density ($N = 90$; blue circles) city. b) Cleaning rate as a function of density for $M = 10$; solid black line represents the sigmoid $f_c(N)$ found by ordinary least squares, the dashed purple line the inferred parameter $N_0$ and the shaded purple area the theoretical bounds on $N_c$, namely $[N_c^-, N_c^+]$. 

(a) Dependence on migration for low- and high-density cities. 
(b) For high migration, the dependence on density is sigmoidal.
This work was supported by Engineering and Physical Sciences Research Council grant EP/S022244/1 for the MathSys CDT.

REFERENCES


