Giant catalytic effect of altruists in Schelling's segregation model

Pablo Jensen,[1, 2, *] Thomas Matreux,[1, 2, 3] Jordan Cambe,[1, 2] Hernan Larralde,[4] and Eric Bertin[5]

[1] *Institut Rhônalpin des Systemes Complexes, IXXI, F-69342 Lyon, France*
[2] *Universite de Lyon, Laboratoire de Physique ENS Lyon and CNRS, 46 Rue d'Italie, F-69342 Lyon, France*
[3] *Physics Department, Ludwig-Maximilians-Universität München, 80799 München, Germany*
[4] *Instituto de Ciencias Físicas, UNAM, Av. Universidad s/n, CP 62210 Cuernavaca Morelos, México*
[5] *LIPHY, Univ. Grenoble Alpes and CNRS, F-38000 Grenoble, France*

We study the effect of introducing altruistic agents in a Schelling-like model of residential segregation. We find that even an infinitesimal proportion of altruists has dramatic catalytic effects on the collective utility of the system. Altruists provide pathways that move the system away from the suboptimal equilibrium it would reach if the system included only egoist agents, allowing it to reach the optimal steady state.

Simple social models may be useful to improve our intuitive, often implicit, conceptualizations of social processes [1–4]. For example, the segregation model proposed by Schelling [5] helps understanding that the collective state reached by agents may well be different from what each of them seeks individually. Specifically, Schelling's model shows that even when all agents share a preference for a mixed city, the macroscopic stationary state may be segregated [6]. In this Letter, we show that introducing a vanishingly small concentration of altruist agents gives rise to a strongly non linear response.

Our model combines two important themes for many disciplines, including physics and economics: The large effects of small perturbations and the influence of altruistic behavior on coordination problems. On the first point, microscopic causes leading to macroscopic effects are well-known in physics. Chaos theory has shown that some dynamical systems are prone to an exponential increase of small perturbations [7], a topic of recurring interest in other fields, such as modeling of ecological competition [8] or pattern formation [9]. More related to our work, there are several examples of large effects arising from small changes in population composition. It has been shown that a small variation in the proportion of uninformed individuals may lead to strong changes in the way collective consensus is achieved by groups manipulated by an opinionated minority [10]. In the minority game [11], introducing a small proportion of fixed agents - i.e. agents that always choose the same option - induces a global change in the population behavior, leading to an increase of the overall gain [12, 13]. In the voter model, a finite density of voters that never change opinion can prevent consensus to be reached [14].

On the second point, altruism is a major topic in evolutionary biology and economics [15–17]. Many models have shown that pair interactions between selfish players lead to stationary states of low utility. They have introduced various types of altruistic behavior to investigate how it may lead to a better equilibrium: altruistic punishment [15, 16], inequity aversion [18], fraternal attitudes [19], agent mobility[20] ... Here, we use a simple definition of altruism (see below) and concentrate on the proportion of altruists needed to reach the social optimum. We show that, unexpectedly, an infinitesimal proportion of altruists can coordinate a large number of egoists and allow the whole system to reach the social optimum.

Description of the model. Our model represents the movement of a population of agents in a "city", which is divided into $Q \gg 1$ non overlapping blocks, also called neighborhoods. Each block is divided into H sites and has the capacity to accommodate H agents (one per site). Initially, a number of agents $N = QH\rho_0$ are distributed randomly over the blocks, leading to an average block density ρ_0 ($\rho_0 = 0.4$ throughout the paper). All agents share the same utility function $u(\rho)$ that depends on the agents density ρ in the neighborhood where they are located. We choose a triangular utility (see Fig. 1): agents experience zero utility if they are alone ($\rho = 0$) or in full blocks ($\rho = 1$), and maximum utility $u = 1$ in half-filled blocks ($\rho = 0.5$). The collective utility U is defined as the sum of all agents' utilities, $U = H \sum_{q=1}^{Q} \rho_q u(\rho_q)$ and the average utility $\tilde{u}$ per agent is $\tilde{u} = U/N$.

Building upon our past work on Schelling's segregation model [6], we now mix two types of agents: "egoists", who act to improve their own, individual, utility, and a fraction p of "altruists", who act to improve the collective utility. Thus, egoists have as objective function the variation of their individual utility Δu, while altruists consider the variation of the overall utility ΔU. The dynamics is the following: at each time step, an agent and a free site in another block are selected at random. The agent accepts to move to this new site only if its objective function *strictly* increases (note that the moving agent is taken into account to compute the density of the new block). Otherwise, it stays in its present block. Then, another agent and another empty site are chosen at random, and the same process is repeated until a stationary state is reached, i.e., until there are no possible moves for any agent.

Limiting cases: pure egoist or altruist populations. In [6], we have computed analytically the stationary states of a *homogeneous* population of egoist or altruist agents.
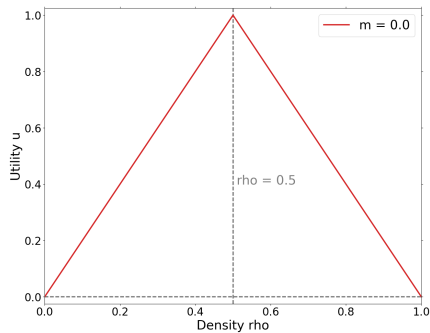
Figure 1: Agent utility function: $u(\rho) = 2\rho$ for $\rho \leq 0.5$ and $u(\rho) = 2(1 - \rho)$ for $\rho > 0.5$.

Altruists always reach the optimal state, given by half filled (or empty) blocks and an average pure altruist utility $\tilde{u}_A \simeq 1$. In contrast, a pure egoist population *collectively* maximizes not U but an effective free energy that we have called the *link L*. The link is given by the sum over all blocks q of a potential l_q: $L = \sum_q l_q$, where $l_q = \sum_{n_q=0}^{N_q} u(n_q/H)$, with $N_q = H\rho_q$ is the total number of agents in block q. In the large H limit,

$$l(\rho_q) \approx H \int_0^{\rho_q} u(\rho) \, d\rho. \tag{1}$$

The link may be interpreted as the cumulative of the individual marginal utilities gained by agents, as they progressively enter the blocks from a reservoir of zero utility. Its key property is that, *for any move*, $\Delta L = \Delta u$. Since egoists move only when their individual Δu is positive, the stationary state is given by maximizing L over all possible densities $\{\rho_q\}$ of the blocks, from which no further $\Delta u > 0$ can be found. Analytical calculations [6] show that this stationary state corresponds to crowded neighborhoods, far above the state of maximum average utility given by $\rho_q = 1/2$. For the case studied in this paper, the stationary density is given by $\rho_E = 1/\sqrt{2}$, leading to a pure egoist utility $\tilde{u}_E = 2(1 - \rho_E) \simeq 0.586 \ll 1$. Numerical simulations have confirmed these results, though the existence of many metastable states around $\rho_E \simeq 0.7$ leads to fluctuations in the simulated final densities.

Mixing populations: qualitative picture. We now investigate how adding a fraction of altruists drives the system away from the frustrated pure egoist case to the optimal configuration observed in the pure altruist case. We find that, instead of a linear response, the system reaches the optimal state even at very low altruist concentrations ($p < 0.01$ in figure 2 a). To help understanding the origin of this strongly non-linear effect, the different panels of Fig. 3 illustrate the evolution of a small system ($H = 225$, $Q = 36$ and $p = 0.04$). Initially, altruists (yellow) and egoists (red) are distributed randomly in the blocks (a), which all have a density $\rho \simeq \rho_0 = 0.4$. Then, blocks with the lowest densities are depleted by both altruists and egoists that prefer districts with higher densities. At
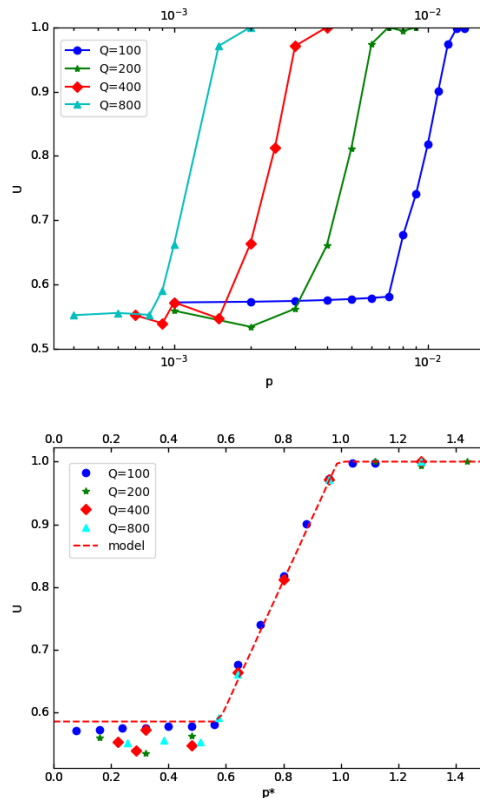

Figure 2: Evolution of the average utility as a function of (a) the altruists' fraction p (note the log scale on the x-axis) and (b) the rescaled fraction $p^* = 2pQ\rho_0$. We take $H = 200$ and vary Q as shown. The fluctuations for low p^* values (before the transition) arise from metastable states in the pure egoist regime.

some point, when the block density increases, the behavior of the two kinds of agents diverge. Altruists "sacrifice" themselves and leave these high density blocks, moving to blocks with lower densities, as this increases the utility of their many (former) neighbors, leading to an increase in global utility. On the other hand, egoists would loose individual utility by doing so, and therefore remain in these high density blocks which continue to feed on the remaining neighborhoods with $\rho < 1/2$. After a few iterations (Fig. 3b-c), selfish agents have gathered into "segregated" neighborhoods. This is the classical segregation observed in the pure egoist case [6], arising from the well studied amplification of density fluctuations. Note that all altruists have left the egoist blocks and gather into few blocks with lower densities (Fig. 3c) and then into a single neighborhood, whose density increases until it becomes attractive for egoist agents who "invade" it (Fig. 3d-e), while altruists leave it for other lower density blocks (Fig. 3e). The density of some of these new blocks then increases, allowing for successive egoist invasions (Fig. 3f-g). These migrations of egoist agents reduce

the density of the overcrowded egoist blocks, increasing the overall utility. Eventually, the system reaches a stationary state in which no agent can move to increase its objective function (Fig. 3h).

Quantitative description. We now give a quantitative explanation of the decrease of egoist block densities and show that an altruist concentration $p \simeq 1/Q$ is sufficient to drive the system towards the optimal state, $\tilde{u} = 1$. To understand altruists' dynamics, it is useful to replace their dynamics by an equivalent egoist dynamics with a utility $u_{\mathrm{altr}}(\rho)$ that differs from the original utility $u(\rho)$. An exact mapping can be done in the following way. As mentioned above, each altruist agent tries to maximize the global utility $U = H \sum_{q=1}^{Q} \rho_q u(\rho_q)$. In contrast, an egoist agent acts to maximize the link function $L = \sum_q \ell(\rho_q)$, with $\ell(\rho_q)$ given in Eq. (1). As a result, an altruist agent exactly behaves as an equivalent egoist agent with a utility function $u_{\mathrm{altr}}(\rho)$ satisfying the relation

$$\rho u(\rho) = \int_0^\rho u_{\mathrm{altr}}(\rho') \, d\rho' \tag{2}$$

since the resulting function to be maximized is the same. Differentiating this last equation, one finds

$$u_{\mathrm{altr}}(\rho) = \frac{\partial \big(\rho \, u(\rho)\big)}{\partial \rho} = \begin{cases} 4\rho, & \text{for } \rho \leq \frac{1}{2} \\ 2(1 - 2\rho), & \text{for } \rho > \frac{1}{2} \end{cases} \tag{3}$$

This effective utility function for altruists is plotted on Fig. 4. Note that this effective utility is *not* the one used to compute average or global utilities, it only helps understanding altruists' moves, since an altruist moves to a new block only if $u_{\mathrm{altr}}(\rho)$ increases. Fig. 4 shows that altruists have a clear preference for blocks with densities just below $1/2$. The large discontinuity at $\rho = 1/2$ arises because at this density the original utility function $u(\rho)$ changes slope and starts to decrease. Then, an altruist moving from a block with $\rho < 1/2$ to a slightly more populated one with $\rho > 1/2$ induces a large decrease of total utility, since all its former neighbors loose utility (as the density of the initial block decreases) and so do its new neighbors, as the density of their block increases.

Fig. 2b suggests that the transition towards the optimal state is continuous and takes place at an altruist concentration $p \simeq 1/Q$ for all values of Q. This Q dependence is important, since in the thermodynamic limit ($Q \to \infty$), the transition would take place at $p \to 0$. We now derive this result in a simple way by computing analytically the evolution of the average utility as a function of the altruist concentration p. Let's start with very low altruist concentrations and assume that the initial dynamics is dominated by egoists, which form the usual Schelling's overcrowded blocks, as observed above (Fig. 3c) and in previous work [6]. Therefore, we take as starting point a city composed of n_E egoist blocks with uniform density $\rho_e = \rho_E > 1/2$, such

that $\rho_E = (1 - p)Q\rho_0/n_E$. Taking a uniform ρ_e value is justified because any density fluctuation for $\rho_e > 1/2$ is rapidly wiped out by the dynamics, as shown by the unique density of egoist blocks in Fig. 3i. Altruists can be initially somewhat scattered over the remaining blocks but, as their effective utility clearly shows (Fig. 4), they rapidly aggregate into a single block, leading to an altruist density $\rho_a = pQ\rho_0$ provided $\rho_a < 1/2$, or equivalently

$$p < p_{\mathrm{high}} \equiv \frac{1}{2\rho_0 Q} \, . \tag{4}$$

The driving force for the transition are the relative values of agents' utilities in egoist and altruist blocks, respectively $u_e = u_E = 2 - 2\rho_E$ and $u_a = 2\rho_a$ since $\rho_a < 1/2$ and $\rho_E > 1/2$. For very low p values, ρ_a is small, leading to $u_e > u_a$ and the system remains in the usual frustrated Schelling egoist state $\tilde{u}(p) \simeq \tilde{u}_E$ which is essentially constant. When p reaches a value p_{low} such that $u(\rho_a + 1/H) > u(\rho_E)$, a first egoist can improve its utility by moving into the altruist block, whose density becomes $\rho_a + 1/H$ (Fig. 3d-f). This gives :

$$p_{\mathrm{low}} \equiv \frac{1 - \rho_E - 1/H}{\rho_0 Q} \, , \tag{5}$$

The density of the invaded block rapidly increases (Fig. 3e) and eventually reaches $1/2$. At this point, altruists' effective utility becomes negative, pushing them to leave for other lower density blocks (Fig. 3f). As previously, altruists gather in another single block of identical density $\rho_a = pQ\rho_0$. The invasion has led to a slight decrease of the density of egoist blocks to $\rho_e < \rho_E$, and therefore to a slight increase of egoists' utility, $u_e = u(\rho_e) > u(\rho_E)$. Successive invasions of the block partially filled by altruists are possible until ρ_e decreases down to the value ρ_e^* such that $u(\rho_e^*) = u(\rho_a + 1/H)$. This leads to $\rho_e^* = 1 - pQ\rho_0 - 1/H$ ($\rho_e^* > 1/2$ as long as $p < p_{\mathrm{high}}$). The equality of utilities implies $\tilde{u}(p) = u(\rho_a) = 2pQ\rho_0 + 2/H$. When $p = p_{\mathrm{high}}$, the final (lowest) egoist density reaches the optimal value $\rho_e^* = 1/2$ and no further improvement in average utility is possible: $\tilde{u}(p) = 1$ (to simplify the discussion, we ignore here corrections of order $1/H$ that depend on the parity of H). This description remains valid for larger altruist concentrations, the only difference being that, at the end, the additional altruists form stable blocks with densities $\rho_a = 1/2$.

In summary, the evolution of the average utility $\tilde{u}$ follows:

$$\begin{cases} \tilde{u}(p) = 2 - 2\rho_E & \text{for } p \leq p_{\mathrm{low}} \\ \tilde{u}(p) = 2pQ\rho_0 + 2/H & \text{for } p_{\mathrm{low}} \leq p \leq p_{\mathrm{high}} \\ \tilde{u}(p) = 1 & \text{for } p \geq p_{\mathrm{high}} \end{cases} \tag{6}$$

Our analysis predicts that plotting $\tilde{u}$ as a function of the rescaled altruist proportion $p^* = p/p_{\mathrm{high}} = 2pQ\rho_0$ should lead to a universal transition starting at $p^* = 2 - 2\rho_E \simeq 0.586$ and ending at $p^* = 1$. Simulations perfectly confirm our calculations (Fig. 2b).

Discussion. Our model illustrates the complexity of the dynamics produced by two types of agents, even when they follow simple rules. Introducing altruists into a population dominated by egoists increases the average utility much more rapidly than expected from a linear projection. The interplay between the different behaviors leads to complex "catalytic" phenomena. By catalytic, we mean that altruists are not "consumed" once they coordinate egoists, and can continue to help egoists finding the optimal configuration indefinitely. The global utility increase *per* altruist at small p can be computed easily: $\delta U_{\mathrm{altr}} \equiv (U(p) - U(p = 0))/N_A \simeq (1 - 0.56)\rho_0 QH/(p\rho_0 QH) = 0.44/p$. When $p = 1/Q$, $\delta U_{\mathrm{altr}} \simeq 0.44Q$. Each altruist induces a utility change proportional to the system size, which becomes infinite for infinite systems. The fact that such large effects arise from *endogenous* dynamics is a key point suggested by recent models of social systems [2, 10, 13, 14]. Indeed, most economic models focus on equilibrium and explain crisis by exogenous factors. In contrast, physics models can interpret large opinion swings in a population as a branching process that generates avalanches or abrupt optimal strategy switches as intrinsic first order transitions [2].

Interestingly, while the stationary state of a system composed of a single type of agents (either egoists or altruists) can be mapped to an equilibrium state, this is no longer the case when including two types of agents, except if some restrictive conditions are met [6]. In a thermodynamic analogy, the utility function can be mapped (in the zero temperature limit considered here) to a chemical potential, as shown in [21], when a single type of agents is present. If a system with both egoist and altruist agents could be mapped to an equilibrium system, chemical potentials could be defined as $\mu_e(\rho_a, \rho_e) = u(\rho_a + \rho_e)$ and $\mu_a(\rho_a, \rho_e) = u_{\mathrm{altr}}(\rho_a + \rho_e)$. As chemical potentials derive from a free energy, their cross derivatives would be equal, $\partial \mu_e/\partial \rho_a = \partial \mu_a/\partial \rho_e$, leading to $u'(\rho) = u'_{\mathrm{altr}}(\rho)$. This equality is not satisfied as seen from Figs. 1 and 4, showing that the system reaches a non equilibrium steady state.

We are well aware that simple models do not allow to draw any rigorous conclusion about what is going on in the *real* world [4, 22, 23]. While Schelling's segregation model neatly shows that one cannot logically deduce individual racism from global segregation, it may well be that for some towns racism is one cause of segregation, for some others not; at any rate the reasons behind urban segregation are far more complex than those that any simple model can come up with. Simple models can be helpful to analyze some interesting phenomena, the origin of which may be obscured in more complicated realistic settings. Ours may help thinking about the effectiveness of coordination by an infinitesimal proportion of altruist agents, but it cannot be directly applied to real systems. Real agents do not behave like these virtual robots: they

are able to put their actions into context, to anticipate the behavior of the others and moreover, they disagree about what is the social "optimum" [4, 24].

---

* corresponding author, pablo.jensen@ens-lyon.fr
[1] C Castellano, S Fortunato, and V Loreto. Statistical physics of social dynamics. *Rev Mod Phys*, 81:591, 2009.
[2] Jean-Philippe Bouchaud. Crises and collective socio-economic phenomena: Simple models and challenges. *Journal of Statistical Physics*, 151(3):567–606, 2013.
[3] DJ Watts. *Everything is Obvious.* Crown Business, 2011.
[4] P Jensen. *Pourquoi la societe ne se laisse pas mettre en equations.* Seuil (Paris), 2018.
[5] TC Schelling. Dynamic models of segregation. *Journal of Mathematical Sociology*, 1:143–186, 1971.
[6] S Grauwin, E Bertin, R Lemoy, and P Jensen. Competition between collective and individual dynamics. *Proceedings of the National Academy of Sciences*, 106(49):20622–20626, December 2009.
[7] JP Eckmann and D Ruelle. Ergodic theory of chaos and strange attractors. *Rev Mod Phys*, 57:617–656, Jul 1985.
[8] L Hébert-Dufresne, A Allard, PA Noël, JG Young, and E Libby. Strategic tradeoffs in competitor dynamics on adaptive networks. *Scientific Reports*, 7:7576, 2017.
[9] MC Cross and PC Hohenberg. Pattern formation outside of equilibrium. *Rev Mod Phys*, 65:851–1112, Jul 1993.
[10] ID Couzin et al. Uninformed individuals promote democratic consensus in animal groups. *Science*, 334(6062):1578–1580, 2011.
[11] D Challet and YC Zhang. Emergence of cooperation and organization in an evolutionary game. *Physica A*, 246(3):407 – 418, 1997.
[12] SS Liaw and C Liu. The quasi-periodic time sequence of the population in minority game. *Physica A*, 351(2):571 – 579, 2005.
[13] SS Liaw. Maximise global gain in the minority game. *Research Letters in Physics*, 2009, 2009.
[14] M Mobilia, A Petersen, and S Redner. On the role of zealotry in the voter model. *J Stat Mech*, 2007(08):P08029, 2007.
[15] E Fehr and S Gachter. Altruistic punishment in humans. *Nature*, 415:137, 2002.
[16] R Boyd, H Gintis, S Bowles, and PJ Richerson. The evolution of altruistic punishment. *Proceedings of the National Academy of Sciences*, 100(6):3531–3535, 2003.
[17] A Kirman and M Teschl. Selfish or selfless? the role of empathy in economics. *Phil Trans Roy Soc B*, 365(1538):303–317, 2010.
[18] M Hetzer and D Sornette. An evolutionary model of cooperation, fairness and altruistic punishment in public good games. *PLOS ONE*, 8(11):1–13, 11 2013.
[19] G Szabo, A Szolnoki, and L Czako. Coexistence of fraternity and egoism for spatial social dilemmas. *Journal of Theoretical Biology*, 317:126 – 132, 2013.
[20] R Cong et al. Individual mobility promotes punishment in evolutionary public goods games. *Nat Sci Rep*, 7:14015, 2017.
[21] R Lemoy, E Bertin, and P Jensen. Socio-economic utility and chemical potential. *European Physical Letters*, 93:38002, 2012.

[22] T Venturini, P Jensen, and B Latour. Fill in the gap: A new alliance for social and natural sciences. *JASSS*, 18(2):11, 2015.

[23] E Ostrom. Beyond markets and states: Polycentric governance of complex economic systems. *American Economic Review*, 100(3):641–72, June 2010.

[24] B Latour. *Politics of explanation, in Knowledge and Reflexivity, ed. by Steve Woolgar*. SAGE, 1988.

(a) t = 0

(b) t = 2

(c) t = 5

(d) t = 10

(e) t = 12
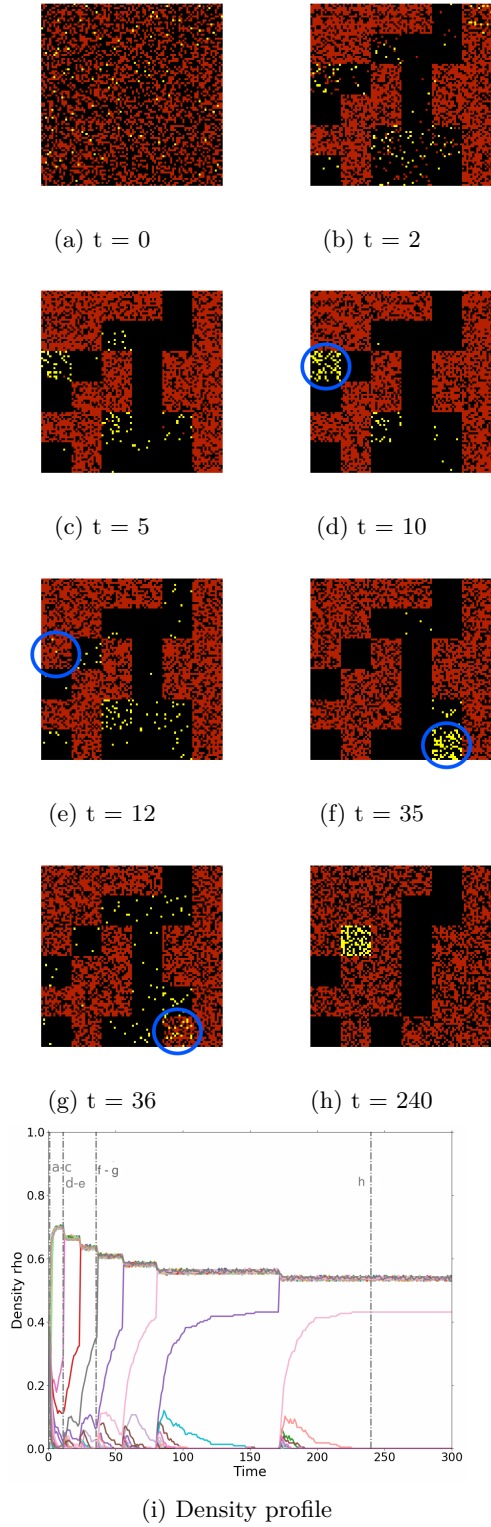
(f) t = 35

(g) t = 36

(h) t = 240

(i) Density profile

Figure 3: Evolution of the city for $p = 0.03, Q = 36$ and $H = 225$. Panels (a-h) show the occupation of the different neighborhoods at different times. Egoists are represented in red, altruists in yellow, empty sites in black. (a) initial; (b) first steps; (c) usual segregation; (d-e): first invasion and altruist escape from the block surrounded in blue; (f-g): final invasion of the block surrounded in blue; (h): stationary state. In panel (i), each continuous line represents the evolution of the density of a single neighborhood. Vertical dashed lines show the times corresponding to panels (a-h).
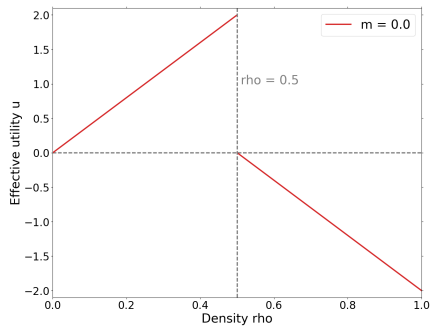
Figure 4: *Effective* utility function of altruistic agents.