Simulated Evolution of Selfish Herd Behavior

Timothy C. Reluga [1]

treluga@amath.washington.edu

Department of Applied Mathematics

University of Washington, Box 352420

Seattle, WA 98195-2420

Telephone: 206.543.0319

Steven Viscido

Department of Biology

University of Washington

Seattle, WA 98195-2420

February 1, 2005

[1]Author to whom correspondence should be addressed

Abstract

Single species aggregations are a commonly observed phenomenon. One potential explanation for these aggregations is provided by the selfish herd hypothesis, which states that aggregations result from individual efforts to reduce personnel predation risk at the expense of group-mates. Not all movement rules based on the selfish herd hypothesis are consistent with observed animal behavior. Previous work has shown that herd-like aggregations are not generated by movement rules limited to local interactions between nearest neighbors. Instead, rules generating realistic herds appear to require delocalized interactions. To date, it has been an open question whether or not the necessary delocalization can emerge from local interactions under natural selection. To address this question, we study an individual-based model with a single quantitative genetic trait that controls the influence of neighbors as a function of distance. The results indicate that selection increases the influence of distant neighbors relative to near neighbors and support the idea that herd behavior can evolve from localized movement rules under predation pressure.

Keywords: selfish herd, behavior, evolution, predation risk

1 Introduction

Gregarious behavior occurs in animals across a wide variety of taxa, including invertebrates (Ritz, 1994), fish (Shaw, 1970), birds (Lack, 1968), wildebeests (Gueron and Levin, 1993), lions (Bertram, 1975), and primates (Nakagawa, 1990), among others. The ubiquity of gregarious behavior across taxa implies that it has evolved independently at least several times (Williams, 1966), and therefore one might justifiably expect aggregations to confer strong advantages upon their members. Indeed, individual benefits may include improved reproductive success (Lack, 1968; Burger and Gochfeld, 1991), increased foraging success (Cody, 1971; Krebs et al., 1972), or an improved chance to survive predation (Hamilton, 1971; Vine, 1971; Watt et al., 1997). The last of these benefits, the lowering of predation risk, has received much attention over the years (Morton et al., 1994).

Social groups can decrease individual predation risk by (1) confusing the predator so that predator success declines (Hall et al., 1986; Smith and Warburton, 1992), (2) providing improved vigilance through many sets of eyes (Lima, 1995; Roberts, 1995), (3) providing a means of cooperative defense to "fight off" the predator (Wilson, 1975), (4) diluting risk so that the per-individual chance of being killed declines (Foster and Treherne, 1981), or (5) providing the individual with "cover" in the form of its group-mates (Hamilton, 1971). Which defense mechanism is employed depends on the size and mobility of the prey relative to that of the predator (Pulliam, 1973), and in some cases the mechanisms cannot operate simultaneously. For example, a very dense group might well dilute risk, while at the same time decreasing vigilance

by blocking the line of sight for all but those on the periphery (Parrish et al., 2002).

An individual's ability to derive "cover" from its group-mates was famously proposed by Hamilton (1971). According to Hamilton's (1971) "selfish herd" hypothesis, animal groups form because individuals try to interpose their group-mates between themselves and potential danger. Hamilton (1971) assumed that the predator could appear anywhere on the field occupied by the prey, and would attack the nearest target. He defined a "Domain of Danger" for each prey individual as the area closer to that individual than to any other member of the herd. Because the predator simply attacks the closest target, if it first appears inside an individual's Domain of Danger, that individual will be attacked. Smaller Domains of Danger are therefore better than larger ones, and there should be a strong advantage to moving in such a way that an individual's Domain of Danger decreases relative to those of the individual's neighbors (Hamilton, 1971; Vine, 1971; Viscido et al., 2002).

The selfish herd hypothesis provides an explanation for *why* an individual would want to be in the midst of many group-mates when threatened with danger, but it does not explain *how* an individual would go about doing so. The mechanism Hamilton (1971) originally proposed – movement toward the nearest neighbor – does not achieve the desired positioning for the individual, nor does it result in a true aggregation forming (Morton et al., 1994; Viscido et al., 2002). Thus, a key question remains unanswered: What "movement rule" must the individuals follow to obtain a densely-packed aggregation where most Domains of Danger are small? Viscido et al. (2002) proposed a rule called the "local crowded horizon" rule, where individuals would scan the horizon, mentally "tallying up" the positions of group-mates and

giving more weight to nearby individuals than to those farther away. The local crowded horizon rule achieved the desired result: a densely packed aggregation formed, where most individuals had small Domains of Danger.

However, a movement rule such as the local crowded horizon rule, where individuals average the influence of many neighbors, is rather complex. We assume that such a rule would not suddenly spring into existence, any more than a complex organ such as the vertebrate eye would. Rather, we envision a rather simplistic starting condition, where animals pay very strong attention to their nearest neighbor, and position themselves relative to only that individual. From such a simple rule, is it possible that the more complicated situation–with individuals evaluating the position of dozens of neighbors–could arise by natural selection? To date, only one study has attempted to tackle this problem. Morton et al. (1994) conducted a simulated evolutionary game between individuals moving randomly and individuals moving toward their nearest neighbor, and showed a slight benefit to those using the latter rule. The difference was judged sufficient to provide a selective advantage (Morton et al., 1994), and so we have a beginning: how nearest-neighbor movement could arise from a random movement habit is established. However, as we noted above, nearest-neighbor movement does not really produce a cohesive group, such as a fish school (Shaw, 1970) – many more neighbors must be considered before large aggregations are observed (Viscido et al., 2002; Warburton and Lazarus, 1991).

In addition to the movement rule question, a further difficulty with the selfish herd hypothesis is that, by definition, not everyone in the herd benefits from the behavior. For each individual who manages to reduce his domain of danger, there is

another individual, somewhere else in the group, whose predation risk increases as a result (Viscido, 2003). During a single predator attack, therefore, there will always be some proportion of the flock that benefits, and some proportion that is harmed, by the behavior. Since a predator only makes a single kill at a time, many individuals in a large domain of danger will survive a single attack, and the evolutionary question then becomes whether, over time, individuals benefit from the behavior more often than they are harmed. In other words, does the behavior allow one to reach a safe position more often than it leaves one at greater risk? If so then we would expect selfish herd behavior to be favored by natural selection; if not, then we would not expect it to arise. Because all previous models (Hamilton, 1971; Morton et al., 1994; Viscido et al., 2001, 2002; James et al., 2004) have only looked at changes to individual risk during a single attack, the answer to this question remains unclear. We attempt to solve this problem by tracking populations over thousands of generations, through many attacks, and following the evolution of a "gene" for selfish herd behavior.

The goal of this study is to determine whether a population composed of individuals paying attention only to one or a few neighbors when deciding where to move could, through the process of natural selection, evolve into a population that pays attention to the entire group. We assume, for the sake of simplicity, that a single quantitative locus controls how strongly an individual is influenced by its neighbors in a parthenogenic population, and that there is a small amount of genetic variance with respect to this trait in the population. We further assume that the predator can appear anywhere and, upon doing so, will kill the nearest target. Given these conditions, we use individual-based simulations to test the hypothesis that the

selective pressure provided by the predator will drive the "neighbor influence" gene such that individuals will consider increasing numbers of neighbors in their decisions.

2 Methods

We begin with a summary of our model. The model is individual-based, consisting of an asexual population and a series of iterations corresponding to predation events. At the start of each iteration, a group of individuals is randomly distributed on a surface. Each individual then moves across the surface for a fixed number of steps. At each step, individuals choose their movement direction using a weighted average of the bearings to all other group members. The weights are functions of the individual's genotype and the distances to each neighbor. After all individuals finish moving, a predator appears at a random location, and the individual nearest the predator is "eaten" (*i.e.* removed from the population). The depredated individual is replaced by a new individual with a genotype similar to that of a randomly chosen surviving individual. This completes one iteration. We monitored changes in the genotype distribution over the course of several thousand iterations. For comparison to this selection model, we construct a null model in which individuals performed uncorrelated random walks. The parameters of our models are listed in Table 1. The remainder of this section describes the implementation details of these two models.

Consider an asexual population of $N = 75$ haploid individuals, labeled $C_1 \ldots C_N$. Each individual C_i has an "influence genotype" represented by a positive real number β_i that parameterizes the individual's movement rule. Initially, all individuals have

the same genotype, $\beta_i = 7$. Each individual C_i also has a position described by a vector $\vec{x}_i \in [0, 1) \times [0, 1)$. The position vector represents the location of the individual on a 1 unit by 1 unit square which has been wrapped to form a torus. The use of a torus generalizes the approach of Hamilton (1971), who used the circular boundary of a pond in one dimension. Individuals distributed in a bounded 3-dimensional arena may collect on the boundary of the arena because it provides some refuge from predation, but the arrangement of individuals on that boundary also affects predation risk. The torus is the simplest such boundary to study. It removes edge effects associated with finite domains (Viscido et al., 2002) and avoids singularities associated with infinite domains (Morton et al., 1994) while allowing for a simple implementation (see Appendix A). Although the use of a torus introduces the possibility of biologically unrealistic paths, observed paths are similar to those expected in a planar domain.

At the start of an iteration, each individual C_i is randomly assigned an initial position $\vec{x}_i$. Positions are uniformly distributed, so that the population's initial spatial distribution is Poisson. We believe this is a reasonable representation of individual distribution before the introduction of a predatory threat (Watt et al., 1997; Viscido and Wethey, 2002).

The selection model continues as follows. Between the initial positioning and predation, each individual travels a fixed distance, D, which we will call the path length. The path is divided into $t = 20$ steps. At each step, each individual chooses a direction and moves D/t units in that direction. After t steps, every individual moves a total distance D, though not necessarily in a straight line. The number of time steps t controls the temporal resolution of the model. In practice, the number of steps must

be chosen large enough to prevent numerical instabilities and small enough for simulations to execute quickly. For efficiency, we have used a grid-free implementation of our model, which lacks collision detection on length scales of $2D/t$. On this small scale, individuals regularly "leap" over each other. These "leaps" randomize individual positions within an aggregation over time, and may weaken selective pressure on the influence genotype.

Each individual independently determines a movement direction by calculating a weighted average of the directions to all the other individuals in the population. Let the unit vector $\hat{z}_{i*}$ denote the preferred movement direction of individual C_i. The new position of individual C_i currently at position $\vec{x}_i$ will be $\vec{x}_i + \hat{z}_{i*}D/t$. The calculation of the weighted average $\hat{z}_{i*}$ is dependent on the individual's influence genotype β_i, the unit direction vectors $\hat{z}_{ij}$ pointing along the shortest path from individual C_i to individual C_j, and the minimum distances r_{ij} from individual C_i to individual C_j. Specifically,

$$\hat{z}_{i*} \propto \sum_{j \neq i} \eta_{ij} \hat{z}_{ij} \ , \tag{2.1}$$

where

$$\eta_{ij} = \frac{1}{1 + (r_{ij}/\rho)^{\beta_i}} \ . \tag{2.2}$$

η_{ij} is interpreted as the "influence" of individual C_i on individual C_j (Viscido et al., 2002). The influence η_{ij} describes the importance of neighbor j in the decision of where to move: the larger η_{ij}, the greater the influence of neighbor j. Eqs. (2.1) and (2.2) are simplifications of an individual's actual decision process. A more realistic biological model should take into account an individual's body orientation, depth

perception, visual acuity, visual limitations imposed by the crowding of neighbors, and collisions with neighbors among other effects. It is also unclear whether the evolution of different behaviors can be modeled as a continuum, or should instead be described as leaps between distinct behaviors. Eqs. (2.1) and (2.2) describe one of many possible evolutionary decision models, but one that is both simple and well-suited to mathematical analysis.

The calculation of the influence η_{ij} in Eq. (2.2) uses a functional form known as the Hill equation (Weiss, 1997; Hill, 1910) in physiology. The maximum possible influence of an individual is chosen to be 1. The parameter ρ controls the half-saturation level of the Hill equation, such that $r_{ij} = \rho$ implies $\eta_{ij} = 1/2$ *i.e.* at a distance ρ, an individual has half the maximum influence. Individuals closer than ρ will have more than half the maximum, and individuals at distances greater than ρ will have less than half the maximum influence. For our simulations, the half-saturation level ρ is chosen to be small($\rho = 0.002$ usually). The genotype β_i in Eq. (2.2) is known as the Hill coefficient, and controls the shape of the function. For large β_i, a steep threshold effect occurs around the half-saturation level, after which the influence rapidly approaches zero as a function of distance. This simulates the case where individuals are "near sighted" and primarily pay attention to those neighbors nearest to themselves. For small β_i, there is no threshold near the half-saturation level and the influence approaches zero slowly. This simulates the condition where individuals are "far sighted" and pay equal attention to the entire group, with relatively little bias toward nearby neighbors. Thus, β_i effectively represents an "influence gene" that determines how far out an individual looks when

deciding where to run.

An alternative form is the piecewise-linear influence function

$$\eta_{ij} = \max\left(1 - \beta_i r_{ij} , \quad 0 \right) \tag{2.3}$$

The genotype β_i has the same qualitative interpretation in both the standard Hill influence function and the alternative piecewise-linear influence function. The piecewise-linear influence function is less accommodating to mathematical study than the Hill function, but similar simulation results were obtained with both forms (see Results).

After the movement of the individuals, a predator is introduced at a point chosen at random from a uniform spatial distribution. This is equivalent to an "ambush" attack, like a lion leaping out of tall grass (Hamilton, 1971). The individual nearest the predator is "eaten", and removed from the population. A randomly chosen surviving individual then reproduces and the offspring individual replaces the dead individual. The genotype $\beta_{\text{offspring}}$ of the offspring individual relates to the genotype β_{parent} of the parent individual by

$$\beta_{\text{offspring}} = 10^{m} \beta_{\text{parent}} , \tag{2.4}$$

where m is a random variable uniformly distributed on $[-\mu, \mu]$, and μ describes the mutation level of the influence gene. The mutation variance grows as μ increases. A multiplicative mutation process is chosen over an additive mutation process so that the coefficient of variation (the ratio of standard deviation and the mean) of $\beta_{\text{offspring}}$ remains independent of the parental genotype β_{parent}. Clearly, the heritability of β decreases as the mutation level μ increases. In the absence of selection and bounds on

the genotype, Eq. (2.4) produces a lognormal distribution of genotypes. We restrict the genotypes to the interval $[0.1, 50]$. Extreme values of the genotype were not observed to produce significantly different predation risks (see Results) but could bias calculations of mean genotype and could also increase the level of round-off error in our calculations. When $\beta_{\mathrm{offspring}}$ fell outside these bounds, the pseudo-random number generator was called repeatedly until an acceptable $\beta_{\mathrm{offspring}}$ value was generated from β_{parent}. These bounds, in combination with the multiplicative mutation process, lead, in the absence of selection, to a population mean $\beta \approx 7$. The initial condition $\beta_i = 7$ was chosen to agree with the mean of this neutral process.

The replacement event completes the iteration. The next iteration begins by randomly repositioning the individuals, and repeating the steps described above.

As a control for the selection model described above, we constructed a null model that is identical to the selection model with one exception: individuals move in uncorrelated random walks such that they are Poisson distributed over the torus. Under this modification, all individuals have the same expected predation risk, independent of genotype. This null model may be biologically appropriate when a predator selects its prey based on size or any spatially independent criteria.

Both the selection model and the null model are ergodic, discrete-time Markov processes, and their genotype distributions were observed to approach asymptotic distributions. The stochastic effects in each iteration are the initial positions of each individual, the position of the predator, and the degree of mutation in the offspring. All other dynamics are deterministic. Our simulation code is written in C++ for the GNU/Linux environment and is available on request.

3 Results

3.1 Analysis of the Influence Genotype

To understand how changes in the "influence genotype" parameter β alter the macroscopic behaviors of individuals, consider a reference individual located at the origin. We define the absolute influence η of a neighboring individual at distance ρ from our reference individual in correspondence with Eq. (2.2) as

$$\eta = \frac{1}{1 + (r/\rho)^{\beta}} , \tag{3.1}$$

where ρ is the half-saturation level, r is distance, and β is the parameter experiencing selection. Suppose we have two neighbors C_1 and C_2 with respective distances r_1 and r_2 from our reference individual. Without loss of generality, we may assume that the distance $r_1 \leq r_2$. We can calculate the absolute influences, η_1 and η_2, using Eq. (3.1). The relative influence η_e is the ratio of the influence of individual C_2 relative to the influence of individual C_1:

$$\eta_e = \frac{\eta_2}{\eta_1} = \frac{1 + (r_1/\rho)^{\beta}}{1 + (r_2/\rho)^{\beta}} . \tag{3.2}$$

Since $r_1 \leq r_2$, the relative influence $\eta_e < 1$. C_2 will have less influence on the reference individual than neighbor C_1.

The extent to which the nearer neighbors are favored controls the degree of clustering we observe in a population. Small clusters are formed by favoring only the nearest neighbors, while large clusters result from less near-neighbor bias (Viscido et al., 2002). To better understand the relationship between the influence genotype β and near-neighbor bias in the influence function, it is helpful to perform an asymptotic

analysis of the relative influence η_e. Assume the distribution of individuals over space is a Poisson distribution in n dimensions. In our model, this Poisson assumption is only valid at the beginning of each iteration. For our experiments, the initial expected nearest-neighbor distance in two dimensions is 0.06 (see Appendix B). Now, for $\beta > 0$,

$$\lim_{\rho \to 0} \eta_e = \left(\frac{r_1/\rho}{r_2/\rho} \right)^{\beta} = \left(\frac{r_1}{r_2} \right)^{\beta} . \tag{3.3}$$

When ρ is small and r_1 and r_2 are random variables, we can approximate the expected relative influence as (see Appendix C)

$$\langle \eta_e \rangle \approx \frac{1}{1 + \frac{\beta}{n}} . \tag{3.4}$$

The expected relative influence $\langle \eta_e \rangle$ of the farther of two neighbors decreases as the influence genotype β increases (Fig. 1). In higher dimensions, the decrease in relative influence is slower, suggesting that it is easier to form large aggregations in three dimensions than in two dimensions.

3.2 Effect of Influence Genotype on Domains of Danger

Before presenting the results of our simulations, it is useful to look at the selection model's qualitative dependence on the genotypes β_i in the absence of predation and mutation. The behaviors of homogeneous populations in which every individual has the same genotype $\beta_i = \beta$ are compared for different values of β. Given sufficiently long path lengths, the final positions of individuals exhibit various amounts of clustering, depending on the value of β (Fig. 2, 3). For large β, numerous small clusters of individuals are observed (Fig. 2b). For small β, a single large cluster is

observed (Fig. 2f). The distribution of individual Domain of Danger areas strongly reflects the degree of clustering (see Fig. 3). When individuals favor their nearest neighbors (large β), the Domain of Danger area distribution is similar to the null model (Fig. 3) On the other hand, when individuals pay attention to more distant neighbors (small β), the Domain of Danger distribution is bimodal: individuals on the cluster perimeter have very large Domains of Danger, while individuals in the cluster interior have very small Domains of Danger (Fig. 3). The Domains of Danger also depend on how tightly a cluster is packed (smaller interior domains result from tighter packing). For intermediate β values, we observe a spectrum of clustering behavior, from solitary individuals to large groups (Fig. 2).

The influence genotype's effect on Domain of Danger area distributions can be tested quantitatively. Pairwise Kolmogorov-Smirnov tests with 100 replicates indicate that values of $\beta < 6.4$ produce Domain of Danger distributions that are significantly different from those produced by values of $\beta \geq 6.4$ (Kolmogorov-Smirnov two-sample tests, $P < 0.001$). Furthermore, Domain of Danger distributions are different for all pairwise comparisons (Kolmogorov-Smirnov two-sample tests, $P < 0.001$) except for $\beta = 1.0$ vs. $\beta = 0.5$ (Kolmogorov-Smirnov two-sample tests, $P > 0.58$), and $\beta = 6.4$ vs. $\beta = 9.0$ (Kolmogorov-Smirnov two-sample tests, $P > 0.37$). Additionally, $\beta = 9.0$ produces distributions that are not significantly different from the null model (Kolmogorov-Smirnov two-sample tests, $P > 0.28$). Thus, there is no improvement in individual safety (with respect to the Domain of Danger) until the influence genotype is reduced to $\beta < 6.4$, but below this value (*i.e.* when the perception range is widened), the Domain of Danger is highly dependent on β. Clearly, for $\beta > 6.4$, the

movement reduces danger little better than random movement, but when β falls below 6.4, even small differences in β will cause a large change in the predation risk experienced by an individual. Then, at some lower threshold, presumably when the animal is paying nearly equal attention to the entire population ($\beta \approx 1$), further reductions in β do not affect predation risk.

3.3 Simulation Results

To address whether the assumptions of the selfish herd hypothesis are sufficient for the emergence of herding as proposed in Hamilton (1971), we can observe the evolution of the influence genotype β over time. Since, for any given iteration, every individual in a population is likely to have a unique genotype, a full description of the population's state requires an N dimensional vector. For simplicity, we summarize the population's state at each time step using the mean genotype of the population,

$$\text{mean}\,\beta = \frac{1}{N}\sum_{i=1}^{N}\beta_i\,, \tag{3.5}$$

and track this mean genotype over time (Fig. 4). The variance of mean β is inversely proportional to the population size N, so mean β is only appropriate when comparing populations of the same size. Although the stochastic nature of the model leads to large oscillations and potentially long initial transients in mean genotype, the asymptotic distribution of mean β in the selection and null models are significantly different (Two-sample t-test assuming unequal variances on log-transformed data, $df = 79$, $P < 0.001$; Fig. 5). Selection consistently decreases mean β in comparison to the null model, though the extent of this decrease is dependent upon the population

size, mutation level, and path length parameters. Thus, selection appears to favor individuals with less localized movement rules.

The effect of predation on mean β strongly depends on the level of mutation μ. For large mutation levels, the selection model and the null model produce mean β values that are similar, though still significantly different (Paired t-test assuming unequal variances for mutation level $\mu = 2$, $df = 35$, $P < 0.001$; Fig. 6). The similarity of the models in the case of large mutation levels corresponds to a loss of heritability. The mutation level is so large that offspring fitness is independent of parental fitness. As the mutation level μ decreases, the selection model's asymptotic behavior becomes distinct from that of the null model (Fig. 6). For the parameters we use, the selection and null models lead to significantly different asymptotic distributions of mean β when $\mu < 2.0$ (Paired t-tests assuming unequal variances, $df = 35$, $P < 0.001$ for all comparisons; Fig. 6). We must note, in passing, the somewhat counterintuitive phenomenon that the variance of mean β in the null model increases as the mutation level decreases. This is an artifact of using mean β to summarize the population's state. In the absence of selection, genotypes may experience large transients. For large mutation levels, these transients average out among individuals and have a stabilizing effect on the distribution of mean β. For small mutation levels, however, most individuals in a finite population will have similar genotypes, and there is little cancellation of stochastic effects in the calculation of mean β. Thus, lower variation among individuals results in a form of within-population synchrony that allows wide variation in mean β and the illusion of wider variation in β across populations.

The path length D also effects mean β. When the path length is less than the

expected nearest-neighbor distance, about 0.06 in two dimensions for $N = 75$ (see Appendix B), the selection model behaves like the null model. For path lengths greater than the maximum separation distance (0.7 on the unit torus), changes in the path length do not greatly affect the asymptotic distribution of mean β. In this later case, selection holds mean β at or below 1 most of the time, given sufficiently small mutation levels. For path lengths between the nearest-neighbor distance and the maximum separation distance, we observe strong dependence of the asymptotic mean β distribution on the specific path length D (Fig. 7). The same result is obtained when we use the alternate piecewise-linear influence function (Fig. 8). Interestingly, mean β obtained its minimum values around $D = 0.4$ rather than at $D = 0.7$ as expected.

Three other parameters in our model are the population size N, the half-saturation level ρ, and the number of steps t. Variations in population size N have visible impacts on the genotype distributions (Fig. 9). Selection causes little change in genotype distribution when the expected nearest neighbor distance is greater than the path length. For large populations, selection reduces the observed β values.

Parameters ρ and t are model-dependent and may be absent in alternative formulations. The distribution of mean β is insensitive to the half-saturation level ρ, provided it was sufficiently small (Fig. 10). The selective pressure on the influence genotype β does change when the half-saturation level is large: many neighbors will have similar influences, independent of that individual's genotype. The distribution of mean β is also insensitive to the number of steps t, provided t is sufficiently large (Fig. 11).

4 Discussion

Hamilton's selfish herd hypothesis is an *a priori* model for evolution (Maynard Smith, 1972). Like most *a priori* models, the selfish herd hypothesis was not formulated independent of observational data, can only explain a limited range of behavioral adaptation, and is only as good as its assumptions. There are many two dimensional movement rules consistent with the selfish herd hypothesis, but only some of these lead to behaviors matching field observations. An important but unaddressed question has been if natural selection will favor delocalized movement rules that facilitate herd-like behavior. In this paper, we test whether the simple assumptions of the selfish herd hypothesis are compatible with the gradual evolution of large-scale aggregations observed in nature. In our individual-based simulations, this is the case. Initially, most individuals pay almost exclusive attention to their immediate neighbors and ignore distant neighbors. However, selection favors those individuals with movement rules that pay more attention to distant neighbors, because those individuals are better at reducing their Domains of Danger, and die less frequently. In a short time, the population becomes dominated by individuals with delocalized influence functions (represented in the model by a small β value; Fig. 4), and these individuals are good at both gathering into a single tight group (Fig. 2) and reducing their Domains of Danger (Fig. 3). We conclude that herd behavior can arise through the gradual evolution of delocalized interactions under the selfish herd hypothesis.

With regard to the dynamics of aggregation, our model disagrees with one of Hamilton's 1971 original proposals. Hamilton suggested that cluster sizes would

increase incrementally over time as small clusters coalesced into larger clusters, and that members of smaller groups would see the "collective benefit" of joining with other small groups (Hamilton, 1971). In our model, cluster size is relatively stable over time, with most clusters remaining at the size of initial formation over the simulation's timescale. These dynamics closely agree with field observations of fiddle crabs (Viscido and Wethey, 2002). Contrary to Hamilton's proposal, collective decision-making was not necessary on any level for the creation of a single unified cluster from an initially scattered population.

One surprising result from our model is the lack of monotonicity in mean β under variable path lengths D. As the path length increases from $D = 0$, corresponding to the null model, the mean influence genotype's distribution decreases until it reached a minimum around $D = 0.4$. This decrease was expected, but the subsequent weak increase in mean β as the path length increases further is unexpected (Fig. 7). The increase is much larger in the piecewise-linear influence function (Fig. 8). There are three potential explanations for this increase: (1) it is a result of weaker selective pressures corresponding to less need for efficiency, (2) it is a result of a reversal in selective pressures, such that the smallest β values are no longer optimal, as appears to be the case near $D = 0.4$, or (3) it is a numerical artifact potentially associated with absence of collision detection. The last of these three appears most likely. Doubling the number of time steps from 20 to 40 weakens the increase of mean β for larger path lengths, more so for the piecewise-linear influence function than for the Hill influence function, but does not remove the increase. Distinguishing between a weakening of selection and a reversal of selection is difficult, even in simulation experiments. If this

increase does, in fact, correspond to a reversal in selection, it suggests some advantage to preemptive local aggregation at large D similar to Hamilton's original suggestion. Although the phenomenon does not affect the conclusions of the present study, it is unexpected and interesting and worth note for future work.

The geometry of the inhabited space may play an important role in the evolution of individual behavior under some circumstances. Replacing the torus with a bounded arena, for example, introduces potentially important edge effects (Viscido et al., 2002). If individuals are Poisson distributed over a square domain, an individual near an edge has a Domain of Danger half that of an interior individual, and an individual in a corner may have an even smaller proportional Domain of Danger. In essence, edges introduce refuge effects. But edge predation risks may be elevated compared to interior predation risk if predators also prefer the edges. The evolutionary tradeoff between selfish herd behavior and attraction to edges and other refuges remains a topic for future research.

We emphasize that our model is an idealized biological representation. For instance, we noted previously that the model is not accurate on length scales smaller than $2Dt$ because there is no mechanism for collision detection. Unfortunately collision detection would seriously complicate our model implementation. We do not feel this significantly affects our results, but collision detection may be useful in more applied studies. In addition, our single-locus description of an influence function in an asexual population is certainly different from the real-life decision-making processes of gregarious species. The modeling of a more realistic mechanism will require species-specific research into behavioral and physiological genetics. On the other

hand, any implementation that leads to similar macroscopic behaviors will experience similar selective pressures. An alternative model, such as a symmetric $\mathfrak{n}$-person zero-sum game for example, should yield a similar conclusion. Thus, despite the abstractions of our model, we feel that it sheds useful light on the potential for the evolution of selfish herds. Whether predation pressure is the actual cause of aggregation, however, must be assessed biologically on a case-by-case basis.

Acknowledgements

The authors would like to thank two anonymous reviewers, the participants of the Mathematical Biology seminar series for helpful discussion, and T. Daniel for suggesting Eq. (2.3). M. Miller and M. Kot provided helpful critical comments. This work was supported in part by NSF VIGRE grant DMS-9810726.

References

Abramowitz, M., Stegun, I., 1972. Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables, 10th Edition. National Bureau of Standards, Washington, D.C.

Bertram, C. R., 1975. Social factors influencing reproduction in wild lions. Journal of Zoology 177, 463–482.

Burger, J., Gochfeld, M., 1991. The Common Tern: Its Breeding Biology and Social Behavior, 1st Edition. Columbia University Press, New York.

Cody, M. L., 1971. Finch flocks in the mohave desert. Theoretical Population Biology 2, 142–158.

Fortune, S., 1997. Voronoi diagrams and Delaunay triangulations. In: Goodman, J. E., O'Rourke, J. (Eds.), Handbook of Discrete and Computational Geometry, 1st Edition. CRC Press, Boca Raton, FL.

Foster, W. A., Treherne, J. E., 1981. Evidence for the dilution effect in the selfish herd from fish predation on a marine insect. Nature 293, 466–467.

Gueron, S., Levin, S. A., 1993. Self-organization of front patterns in large wildebeest herds. Journal of Theoretical Biology 165, 541–552.

Hall, S. J., Wardle, C. S., MacLennan, D. N., 1986. Predator evasion in a fish school: test of a model for the fountain effect. Marine Biology 91, 143–148.

Hamilton, W. D., 1971. Geometry for the selfish herd. Journal of Theoretical Biology 31, 295–311.

Hill, A. V., 1910. The combinations of haemoglobin with oxygen and with carbon monoxide i. Journal of Physiology 40, iv–vii.

James, R., Bennett, P. G., Krause, J., 2004. Geometry for mutualistic and selfish herds: the limited domain of danger. Journal of Theoretical Biology 228, 107–113.

Krebs, J. R., MacRoberts, M. H., Cullen, J. M., 1972. Flocking and feeding in the great tit *Parus major* – and experimental study. Ibis 114, 507–534.

Lack, D., 1968. Ecological Adaptations for Breeding in Birds, 1st Edition. Chapman and Hall, London.

Lima, S. L., 1995. Collective detection of predatory attack by social foragers: fraught with ambiguity? Animal Behaviour 50, 1097–1108.

Maynard Smith, J., 1972. The Theory of Evolution. Penguin, Baltimore, MD.

Morton, T. L., Haefner, J. W., Nugala, V., Decino, R. D., Mendes, L., 1994. The selfish herd revisited: do simple movement rules reduce relative predation risk? Journal of Theoretical Biology 167, 73–79.

Nakagawa, N., 1990. Decisions on time allocation to different food patches by japanese monkeys (*Macaca fuscata*). Primates 31, 459–468.

Parrish, J. K., Viscido, S. V., Grunbaum, D., 2002. Self-organized fish schools: an examination of emergent properties. Biological Bulletin 202 (3), 296–305.

Pulliam, R. H., 1973. On the advantages of flocking. Journal of Theoretical Biology 38, 419–422.

Ritz, D. A., 1994. Social aggregation in pelagic invertebrates. Advances in Marine Biology 30, 156–216.

Roberts, G., 1995. A real-time response of vigilance behaviour to changes in group size. Animal Behaviour 50, 1371–1374.

Shaw, E., 1970. Schooling in fishes: critique and review. In: Aronson, L. R., Tobach,

E., Lehrman, D. S., Rosenblatt, J. S. (Eds.), Development and evolution of behavior, 1st Edition. Vol. 1. W. H. Freeman, pp. 452–480.

Smith, M. F. L., Warburton, K., 1992. Predator shoaling moderates the confusion effect in blue-green chromis, *Chromis viridis*. Behavioral Ecology and Sociobiology 30, 103–107.

Vine, I., 1971. Risk of visual detection and pursuit by a predator and the selective advantage of flocking behavior. Journal of Theoretical Biology 30, 405–422.

Viscido, S. V., 2003. The case for the selfish herd hypothesis. Comments on Theoretical Biology 8, 665–684.

Viscido, S. V., Miller, M., Wethey, D. S., 2001. The response of a selfish herd to an attack from outside the group perimeter. Journal of Theoretical Biology 208, 315–328.

Viscido, S. V., Miller, M., Wethey, D. S., July 2002. The dilemma of the selfish herd: the search for a realistic movement rule. Journal of Theoretical Biology 217 (2), 183–194.

Viscido, S. V., Wethey, D. S., 2002. Quantitative analysis of fiddler crab flock movement: evidence for "selfish herd" behaivour. Animal Behaviour 63, 735–741.

Warburton, K., Lazarus, J., 1991. Tendency-distance models of social cohesion in animal groups. Journal of Theoretical Biology 150, 473–488.

Watt, P. J., Nottingham, S. F., Young, S., 1997. Toad tadpole aggregation behaviour: evidence for a predator avoidance function. Animal Behaviour 54, 865–872.

Weiss, J. N., 1997. The Hill equation revisited: uses and misuses. The FASEB Journal 11 (11), 835–841.

Williams, G. C., 1966. Adaptation and Natural Selection. Princeton University Press, Princeton, NJ.

Wilson, E. O., 1975. Sociobiology: The New Synthesis, 1st Edition. Harvard University Press, Cambridge, Mass.

A Technical issues associated with a Toroidal domain

The use of a torus as the spatial domain removes boundary effects associated with finite domains and avoids some singularities associated with infinite domains, but creates some complications in the calculations of distances and directions. These complications can not be accommodated in other geometries such as spheres without expensive computation. The complications parallel the issues that arise in trigonometry, where the angular measures -2π, 0, and 2π represent the same angle. For instance, the position of an individual C_i can be equally well described as being $(0.5, 0.4)$, $(1.5, 1.4)$, $(-3.5, 4.4)$, or, in general, $(0.5, 0.4) + \mathbb{Z} \times \mathbb{Z}$ (where $\mathbb{Z}$ represents the set of integers, and $\mathbb{Z} \times \mathbb{Z}$ represents the two dimensional integer lattice). The ambiguities are resolved by the convention that measurements are always made based on the nearest image of a neighboring individual. Thus, the distance r_{ij} between individual C_i and individual C_j is defined as

$$r_{ij} = \min \|\vec{x}_i - \vec{x}_j + \mathbb{Z} \times \mathbb{Z}\| \,, \tag{A.1}$$

where $\| \bullet \|$ represents the standard Euclidian magnitude of a vector. Similarly, the direction vector $\hat{z}_{ij}$ from individual C_i to individual C_j is

$$\hat{z}_{ij} = \frac{\vec{\delta}}{r_{ij}} \text{ where } \vec{\delta} \in \vec{x}_i - \vec{x}_j + \mathbb{Z} \times \mathbb{Z} \text{ and } \|\vec{\delta}\| = r_{ij} \,. \tag{A.2}$$

While r_{ij} is always unique, $\hat{z}_{ij}$ is not necessarily unique. However, double precision floating point arithmetic makes cases of non-uniqueness so rare (less than once in

every 10^{12} calculations) that no attempt was made to define $\hat{z}_{ij}$ in the ambiguous cases.

B Expected Nearest Neighbor Distance

In this appendix, we derive the initial expected nearest neighbor distance $\langle r_* \rangle$ when individuals are Poisson distributed in space. Consider two points randomly placed on a regular n-dimensional torus by a generalized Poisson process. The probability that the distance r_* between these two points is less than a given distance r is

$$P(r_* < r) = 1 - A(r) \ , \tag{B.1}$$

where $A(r)$ is the fraction of the area of the torus enclosed in a ball of radius r. More generally, let r_* be the distance to the nearest neighbor in a population of N points. Then

$$P(r_* < r) = 1 - (1 - A(r))^{N-1} = F(r, N) \ , \tag{B.2}$$

where $A(r)$ is as defined above. For given N, $F(r, N)$ constitutes a cumulative probability distribution for r_*.

The density function $\partial F/\partial r$ is usually used in calculation of expected values, but because F is not differentiable in general, it is useful to use integration by parts to obtain the following alternative formula for the expectation of r_*:

$$\langle r_* \rangle = r_f - \int_{r_i}^{r_f} F(r) dr \ . \tag{B.3}$$

$$= r_f - \int_0^{r_f} 1 - [1 - A(r)]^{N-1} \, dr \tag{B.4}$$

$$= \int_0^{r_f} [1 - A(r)]^{N-1} \, dr \ . \tag{B.5}$$

Since $A(r)$ is a monotonically increasing function of r, it is reasonable to asymptotically expand this integral around the left endpoint for large N, in which case,

$$\langle r_* \rangle = \int_0^{r_f} e^{(N-1)\ln[1-A(r)]} dr \sim \int_0^{\infty} e^{-(N-1)A(r)} dr \; . \tag{B.6}$$

In n dimensions, on a locally Euclidian manifold,

$$A(r) \approx \frac{2}{n} \frac{\Gamma^n (1/2)}{\Gamma (n/2)} r^n \; , \tag{B.7}$$

for small r, where $\Gamma()$ is the gamma function (Abramowitz and Stegun, 1972). After integration of Eq. (B.6), we find

$$\langle r_* \rangle \sim \frac{\Gamma \left(1 + \frac{1}{n}\right)}{\sqrt{\pi}} \sqrt[n]{\frac{\Gamma \left(\frac{n+2}{2}\right)}{N-1}} \; . \tag{B.8}$$

In $n = 2$ and $n = 3$ dimensions, the expected nearest neighbor distances are respectively

$$\frac{1}{2\sqrt{N-1}} \quad \text{and} \quad \frac{1}{1.80\sqrt[3]{N-1}} \; . \tag{B.9}$$

C Derivation of the Expected Relative Influence

Here, we calculate the expected value of the relative influence of the second nearest neighbor, as shown in Eq. (3.4). Mathematically, the relative influence

$$\eta_e = \left(\frac{\min\{r_1, r_2\}}{\max\{r_1, r_2\}} \right)^{\beta} , \tag{C.1}$$

where r_1 and r_2 are independent, identically distributed random variables describing the distances to two neighbors that are Poisson distributed on a sphere of radius R. Similar results can be obtained on a torus for sufficiently large populations. To

determine $\langle \eta_e \rangle$, we first observe that the volume of an n-dimensional sphere of radius r is

$$V_n(r) = \frac{2\pi^{n/2}}{n\Gamma(n/2)} r^n \,, \tag{C.2}$$

where $\Gamma()$ is the gamma function (Abramowitz and Stegun, 1972). The joint probability density for finding two neighbors at distances r_1 and r_2 respectively is then

$$P(r_1, r_2) = \frac{1}{V_n(R)^2} \frac{dV_n(r = r_1)}{dr} \frac{dV_n(r = r_2)}{dr} = \frac{n^2}{R^2} \left(\frac{r_1}{R}\right)^{n-1} \left(\frac{r_2}{R}\right)^{n-1} . \tag{C.3}$$

Now,

$$\langle \eta_e \rangle = \int_0^R \int_0^R \left(\frac{\min\{r_1, r_2\}}{\max\{r_1, r_2\}} \right)^{\beta} P(r_1, r_2)\, dr_2\, dr_1 \,. \tag{C.4}$$

Noting that the integral is invariant under the permutation of r_1 and r_2,

$$\langle \eta_e \rangle = 2\frac{n^2}{R^{2n}} \int_0^R \int_{r_1}^R \left(\frac{r_1}{r_2}\right)^{\beta} r_1^{n-1} r_2^{n-1}\, dr_2\, dr_1 \tag{C.5}$$

$$= 2\frac{n^2}{R^{2n}} \int_0^R r_1^{n-1+\beta} \int_{r_1}^R r_2^{n-1-\beta}\, dr_2\, dr_1 \tag{C.6}$$

$$= \frac{n}{\beta + n} \,. \tag{C.7}$$

Thus, the expected relative influence of the second nearest neighbor is

$$\langle \eta_e \rangle = \frac{1}{1 + \frac{\beta}{n}} \,. \tag{C.8}$$

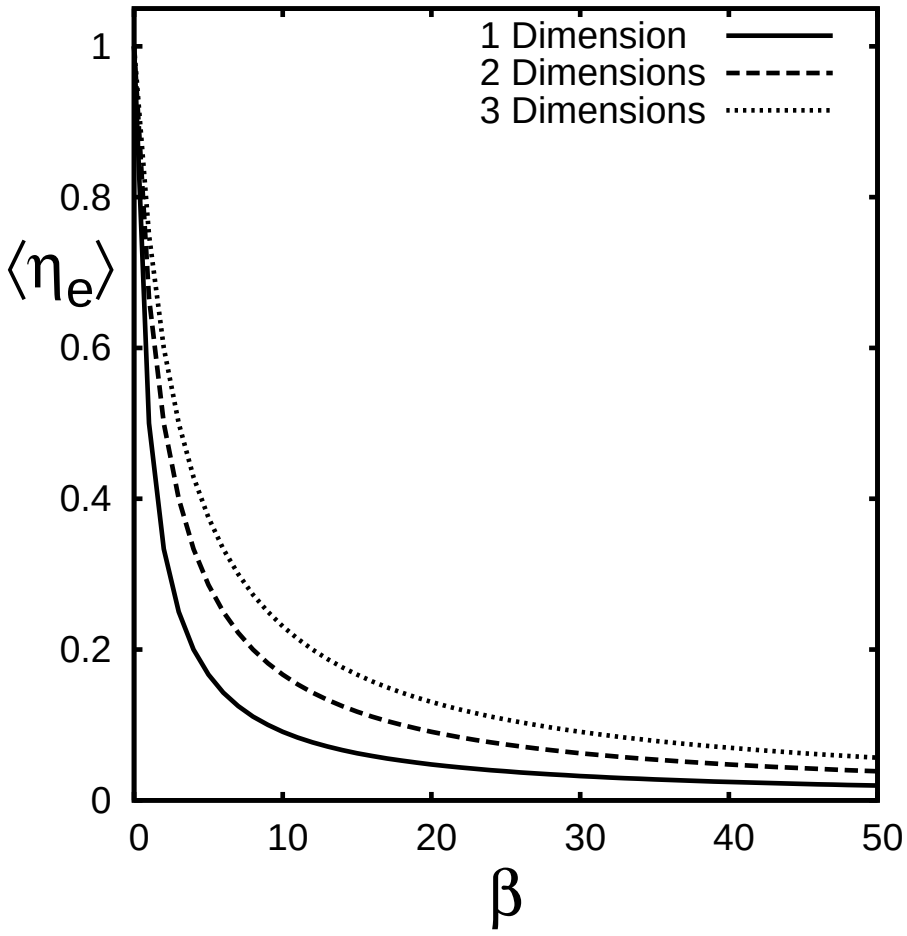

Figure 1: The expected relative influence $\langle \eta_e \rangle$ of the next-nearest neighbor as a function of β when individuals are Poisson distributed. $\langle \eta_e \rangle$ decays algebraically in β, with a rate inversely related to the number of dimensions n (see Eq. (3.4)).

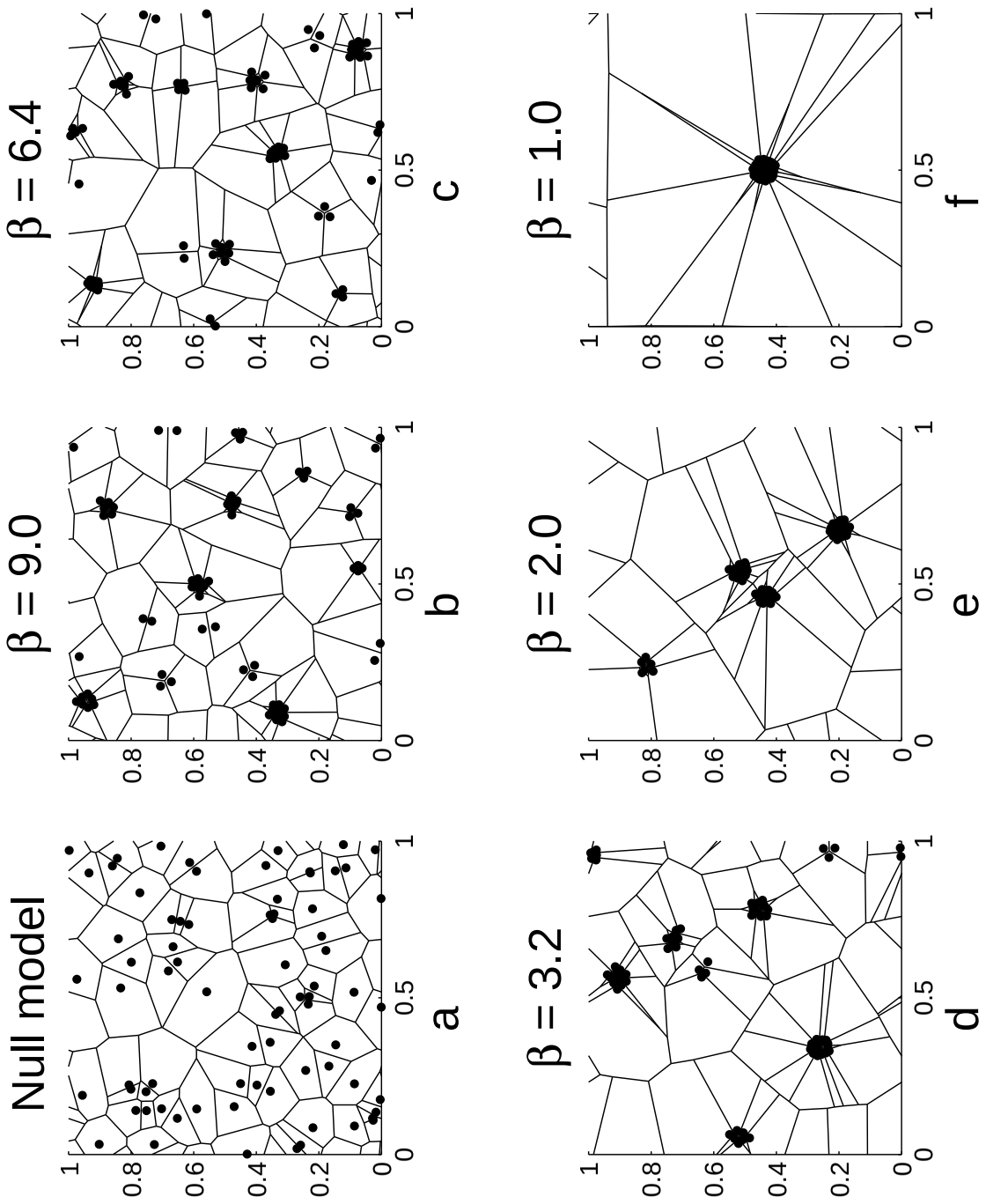

Figure 2: A Voronoi diagram from the null model (Poisson distribution) and five Voronoi diagrams (Fortune, 1997) for uniform populations of 75 individuals with $D = 0.8$ and $\rho = 0.08$. For each plot, all 75 individuals have the same genotype β given at the top of the plot, and were initially Poisson distributed. Cluster sizes increase as β decreases.

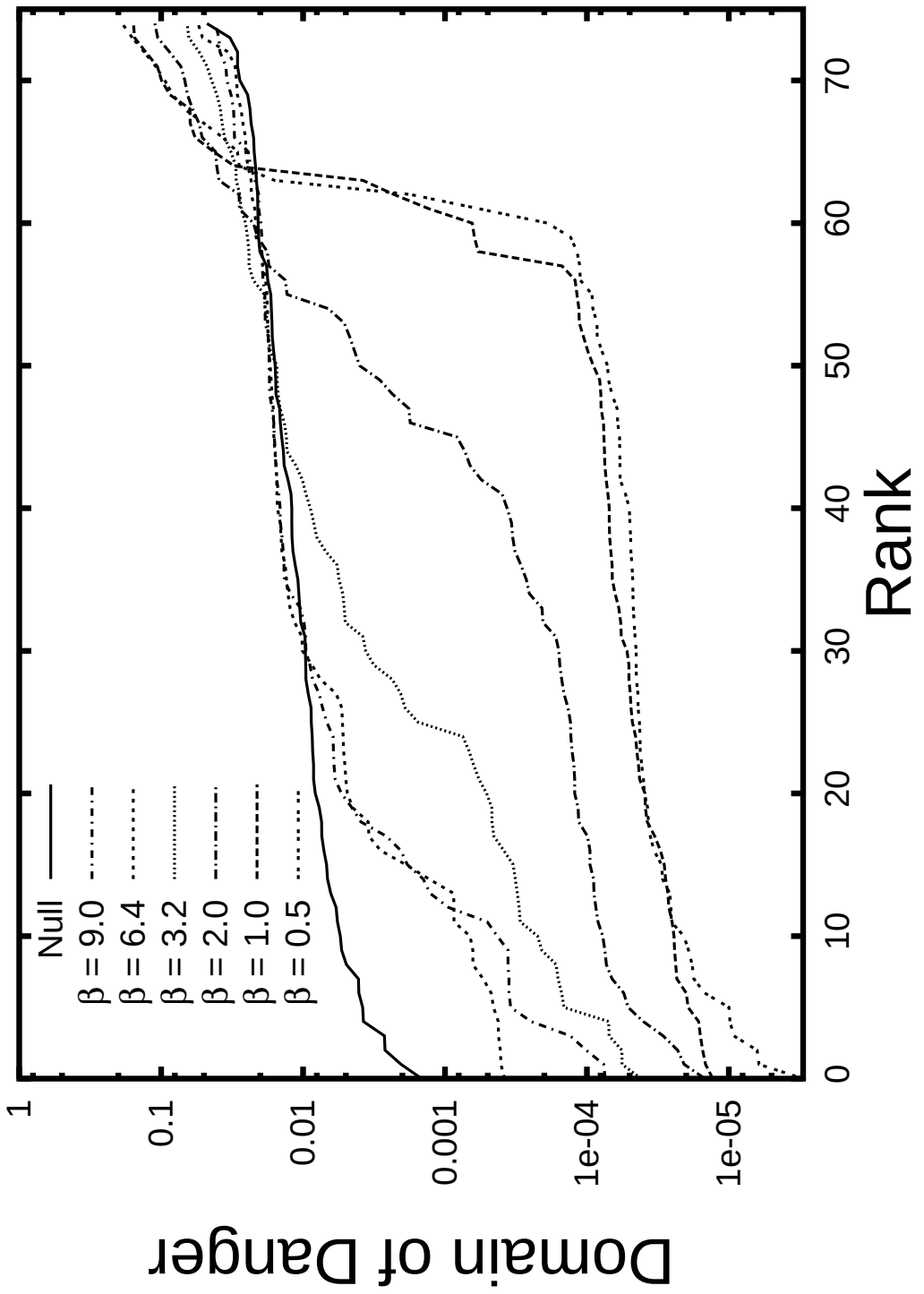

Figure 3: Ranked Domain of Dangers for the Voronoi plots in Fig. 2 and a random example of $\beta = 0.5$. As β decreases, the distribution varies from a relatively flat distribution to a hollowed out distribution where most individuals have small domains of danger, but a few individuals have large domains of danger. Note that $\beta = 1$ is very similar to $\beta = 0.5$.

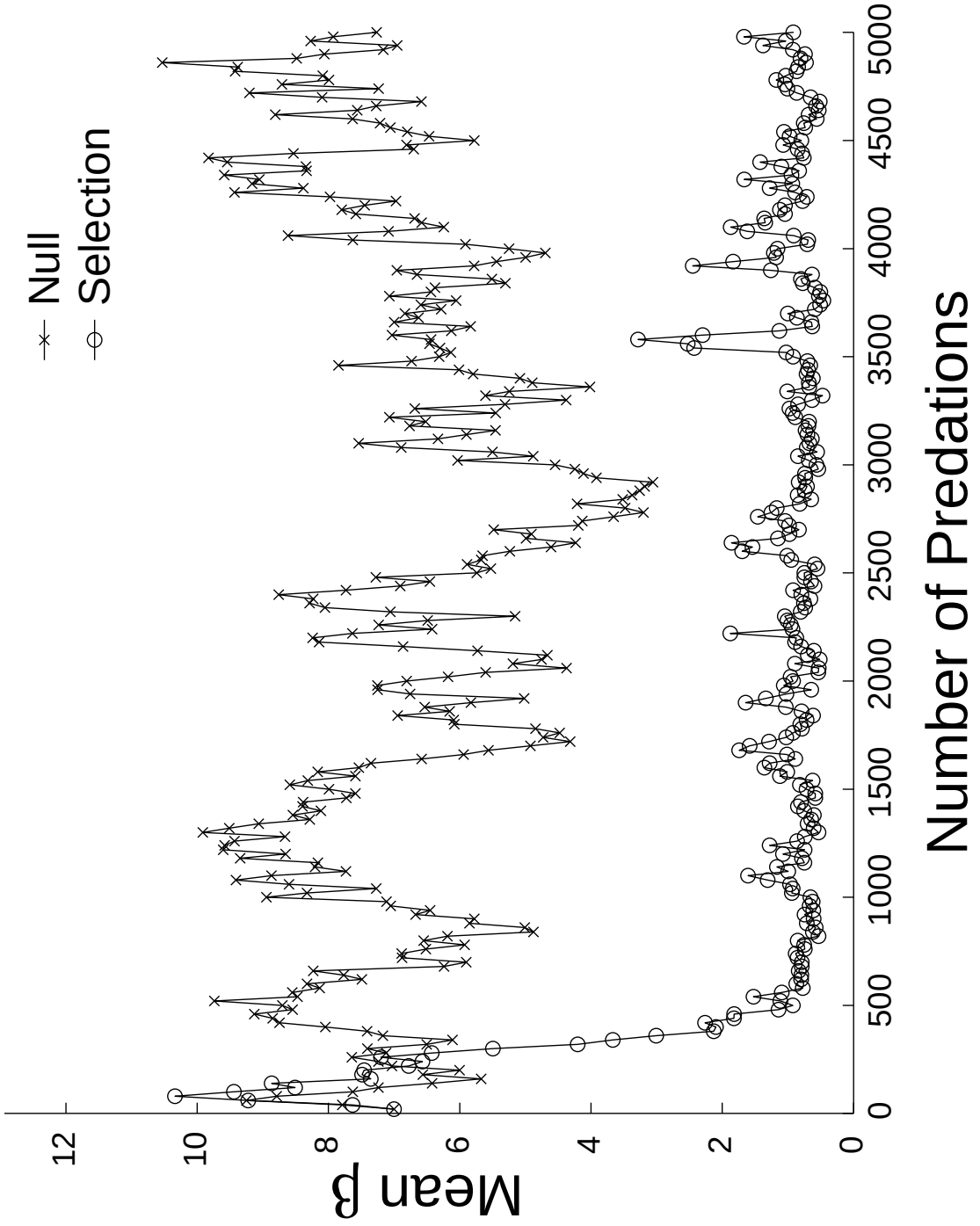

Figure 4: Two example time series for a population's mean β under the null model and the selection model ($N = 75$, $\mu = 1$, $D = 0.4$). The mean was observed every 20 predations for the duration. The initial transient of the selection model, in which the mean increases before decreasing to asymptotic levels, is a common feature attributable to the population's uniform initial conditions.

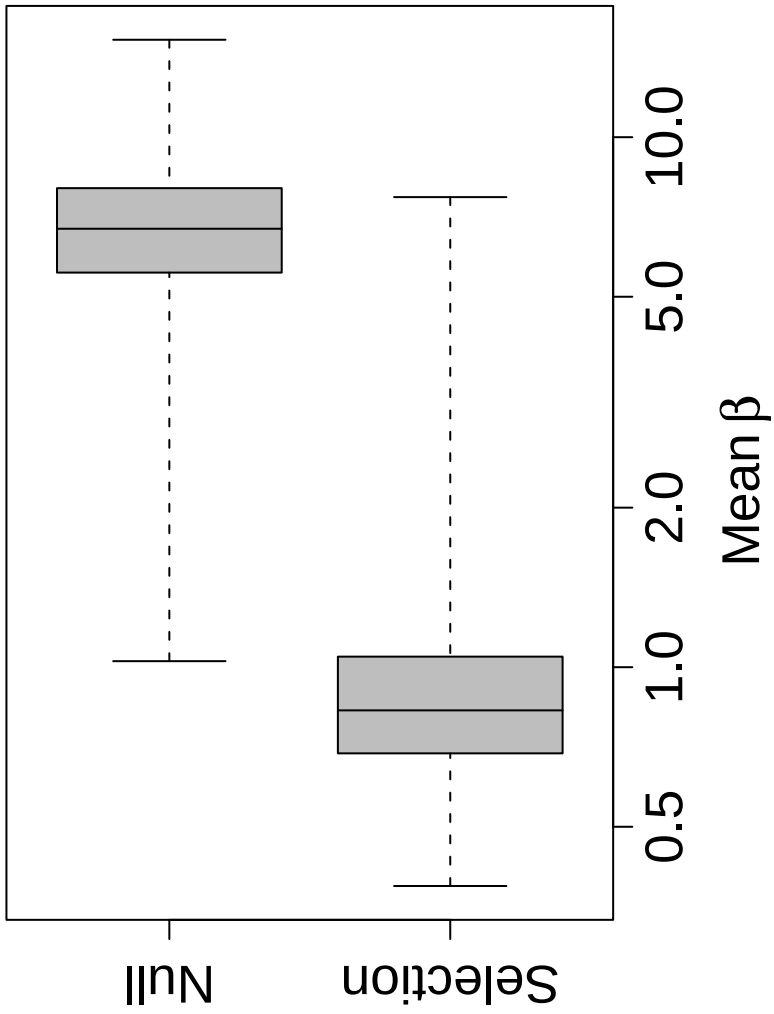

Figure 5: Box plots of mean β distribution under the null and selection models ($\mu = 1$, $D = 0.4$). In the null model, the expected predation risk is independent of genotype. Distributions are based on 63 independent runs observed every 20 predations for a duration of 10^4 predations. Initial transients were discarded. In all box plots, bars extend to the minimum and maximum observed values.

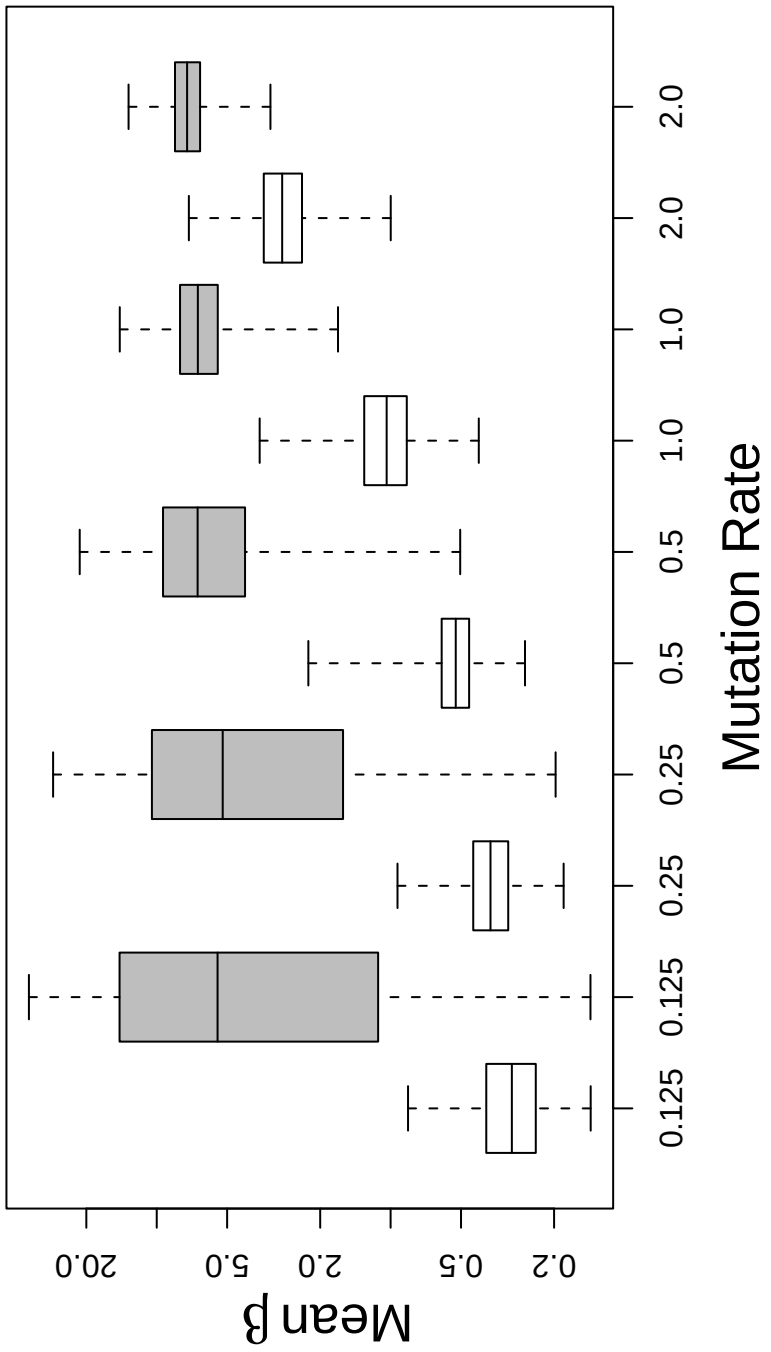

Figure 6: Box plots showing the dependence of mean β distribution on mutation level μ in both the null model (gray) and the selection model (white) ($N = 75$, $D = 0.4$). The variance of the null model increases as the mutation level decreases because the smaller mutation levels correspond to longer autocorrelation times in the observed time-series. For each box, 63 simulations were observed every 50 predations for a duration of $21,000$ predations after initial transients.

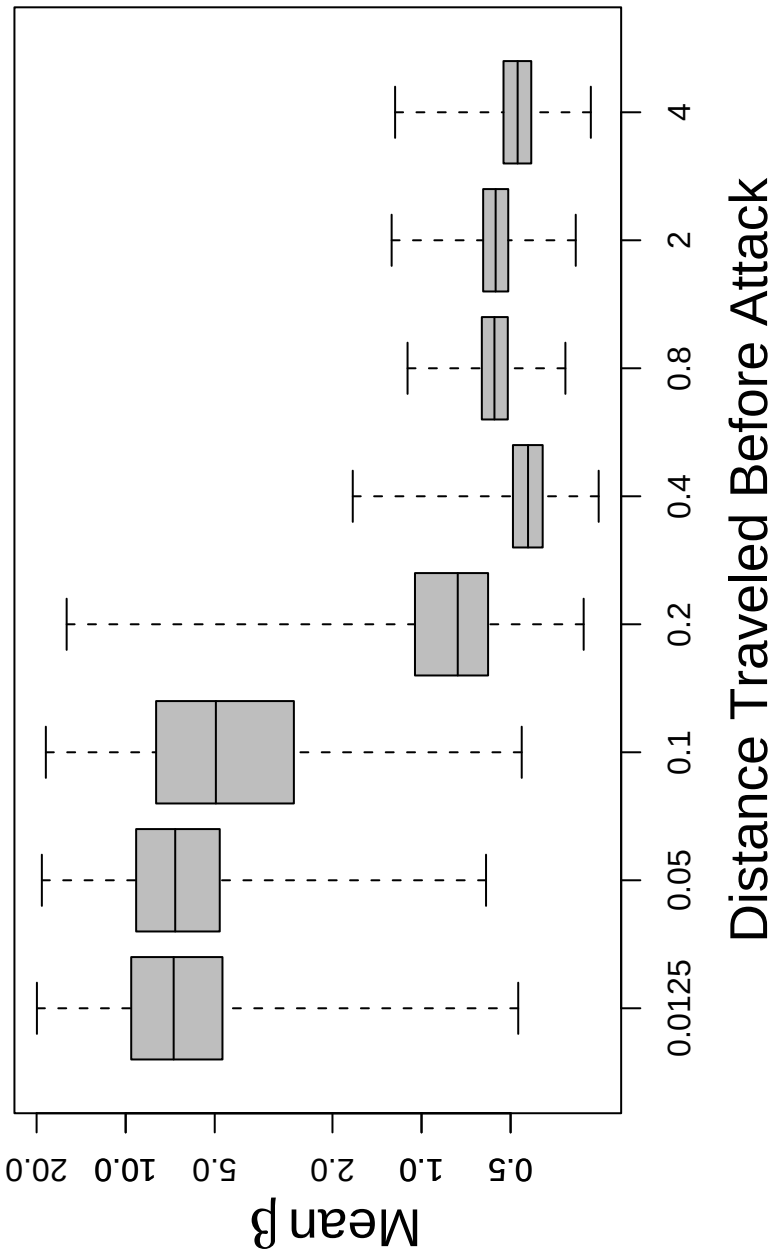

Figure 7: Box plots showing the dependence of mean β distribution on the distance D traveled between warning and attack ($N = 75$, $\mu = 0.5$, $\rho = 0.002$). For small distances, mean β distribution is similar to that of the null model. Selection had the strongest effects when $D \approx 0.4$. The expected nearest neighbor distance is about 0.06 (see Appendix B). The maximum separation distance between any two individuals is 0.7. For each box, 63 simulations were observed every 20 predations for a duration of $4,000$ predations. Initial transients were discarded.

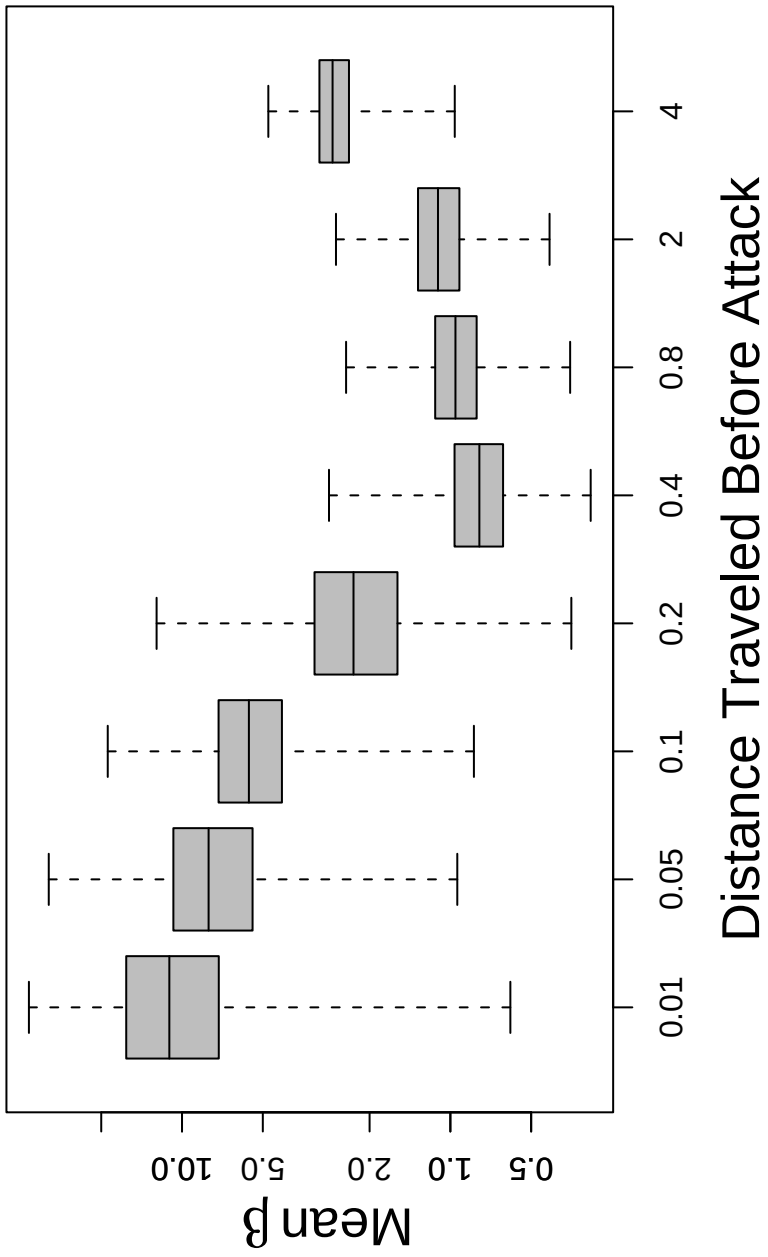

Figure 8: Box plots showing the dependence of mean β distribution on the distance D traveled between warning and attack for the alternate piecewise-linear influence function, Eq. (2.3) ($N = 75, \mu = 0.5$). Genotypes β were restricted to the interval $[0.2, 200]$. For each box, 63 simulations were observed every 20 predations for a duration of $4,000$ predations. Initial transients were discarded. The pattern is similar to that seen in Fig. 7.

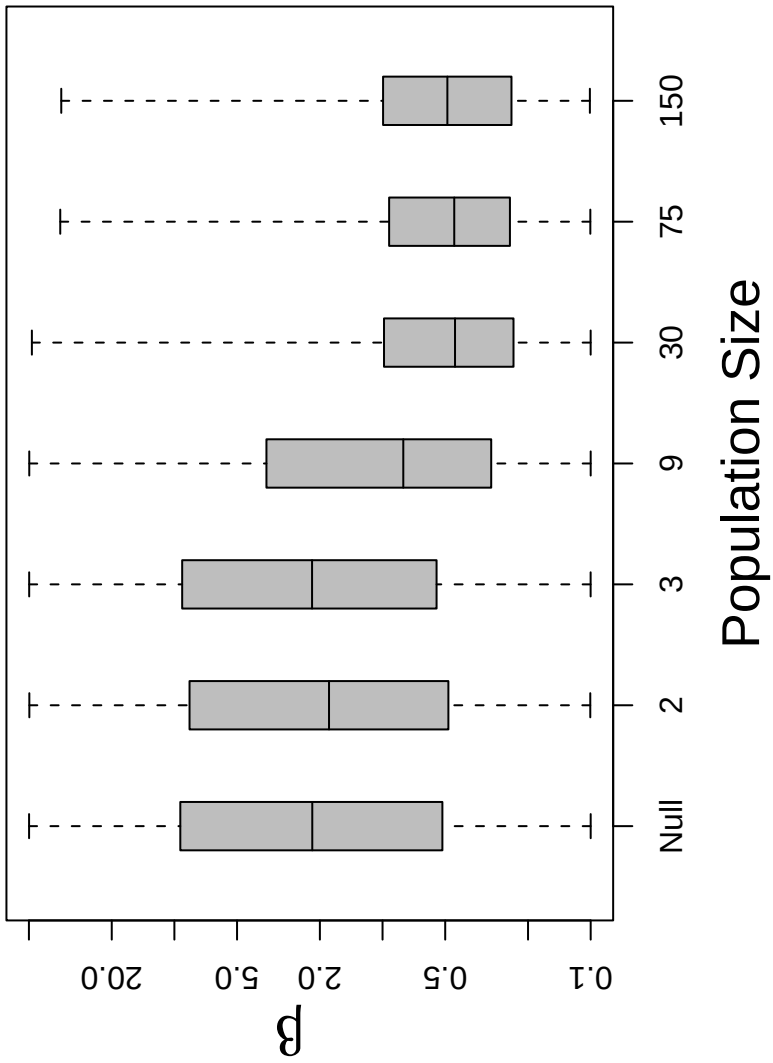

Figure 9: Box plots showing the dependence of β distribution on the population size N ($D = 0.16$, $\mu = 0.5$). When there are few individuals, such that the expected nearest neighbor distances are greater than the distance traveled before attack ($N < 8$), selection has little or no effect on β distribution. When the population is sufficiently large, the effect of selection is pronounced. The null model (far left box, $N = 75$) is included for comparison. Boxes represent 1000 samples.

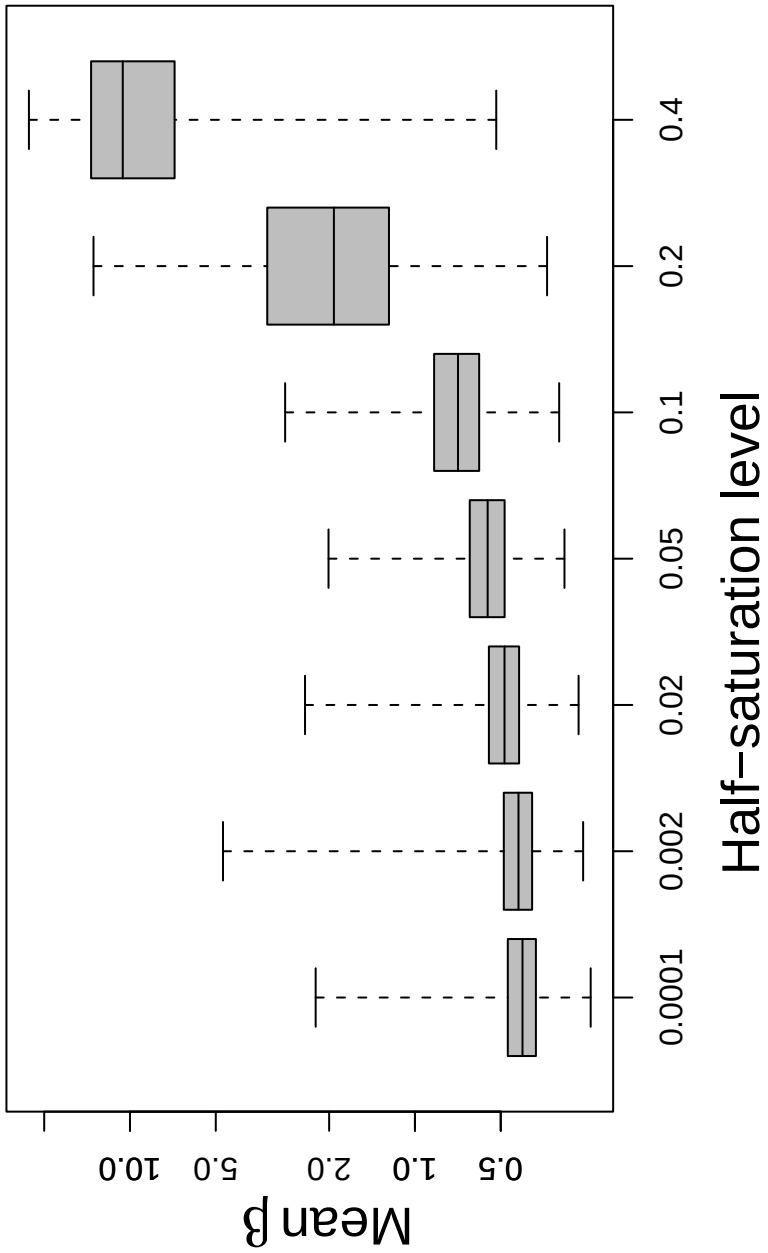

Figure 10: Box plots of mean β distribution depending on the half-saturation level ρ ($N = 75$, $\mu = 0.5$, $D = 0.4$). Initial transients were discarded. The distributions of mean β are similar for $\rho \leq 0.1$ but mean β consistently increases as ρ increases, indicating weaker of selection. For each box, 64 simulations were observed every 20 predations for a duration of $5,000$ total predations.

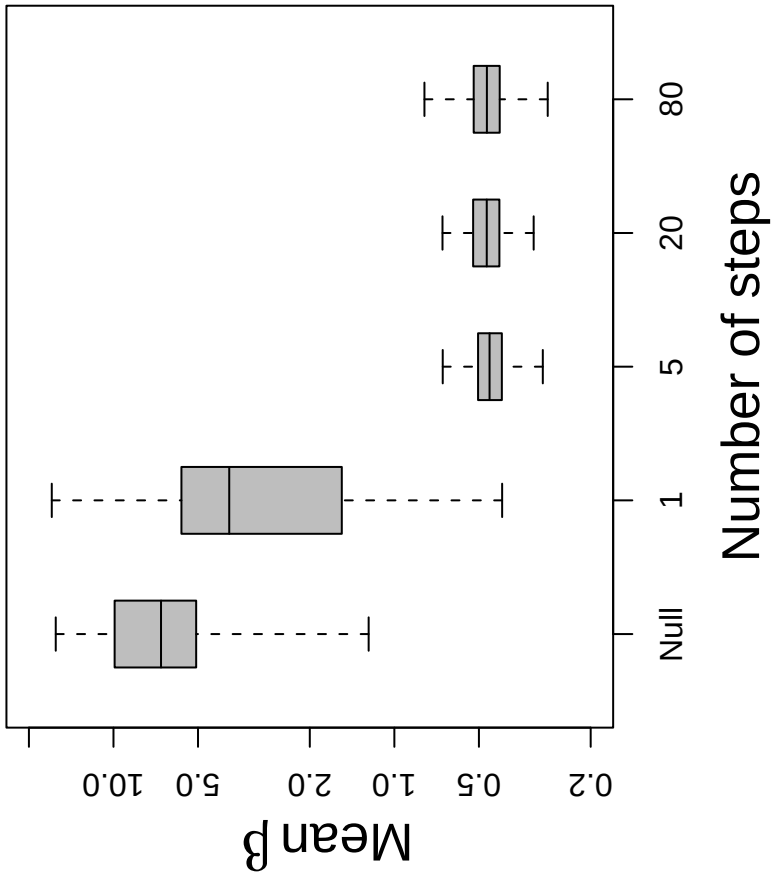

Figure 11: Box plots showing the dependence of mean β distributions on the number of steps t ($N = 75$, $D = 0.7$, $\mu = 0.5$). For $t = 1$, numerical instabilities limit selection on the genotype, but the effects of selection are insensitive to t for sufficiently large values. The distributions of mean β in the null model (far left box) was independent of t and is included for comparison. Boxes represent 240 samples.

Variable	Meaning	Units
N	Number of individuals in population	individuals
D	Distance traveled by each individual before an attack.	meters
t	Number of time steps before an attack	dimensionless
μ	Mutation level of the influence gene	per generation
β_i	Influence genotype of individual i	dimensionless
mean β	Average influence genotype of a population	dimensionless
$\vec{x}_i$	Position of individual i	meters
$\hat{z}_{i*}$	Heading vector of individual i	dimensionless
$\hat{z}_{ij}$	Direction vector from individual i to individual j	dimensionless
η_{ij}	Influence of individual j on individual i	dimensionless
r_{ij}	Distance between individuals i and j	meters
ρ	Half saturation level of the influence function	meters

Table 1: : Symbols used in this paper.