Eastern Economic Journal, **2008**, **34**, (480–503)

www.palgrave-journals.com/eej

SYMPOSIUM ARTICLE

Segregation and Strategic Neighborhood Interaction

Jason M. Barr[a] and Troy Tassier[b]
[a]Department of Economics, Rutgers University, Newark, USA.
E-mail: jmbarr@rutgers.edu
[b]Department of Economics, Fordham University, Bronx, NY 10458.
E-mail: tassier@fordham.edu

We introduce social interactions into the Schelling model of residential choice; these interactions take the form of a Prisoner's Dilemma game. We first study a Schelling model and a spatial Prisoner's Dilemma model separately to provide benchmarks for studying a combined model, with preferences over like-typed neighbors and payoffs in the spatial Prisoner's Dilemma game. We find that the presence of these additional social interactions may increase or decrease segregation compared to the standard Schelling model. If the social interactions result in cooperation then segregation is reduced, otherwise it can be increased.
Eastern Economic Journal (2008) **34,** 480–503. doi:10.1057/eej.2008.26

Keywords: Schelling tipping model; spatial Prisoner's Dilemma; cooperation; segregation

JEL: C63; C73; D62

INTRODUCTION

Racial and ethnic segregation in the United States continues to be common, despite survey results that show people to be increasingly opposed to the idea of racial segregation [Sethi and Somanathan 2004]. Schelling [1969; 1971] has provided one of the most compelling accounts of why segregation is still so widespread. His model, seen as a forerunner of the current agent-based modeling paradigm in economics, shows how even a relatively small preference for neighbors of one's own "type" can lead to neighborhood tipping and high levels of segregation.

While agents, individually, may prefer living in majority-type neighborhoods for cultural or language-based reasons, segregation is considered a bad outcome because of the external effects that it can have on the society as a whole. For instance, it is commonly argued that racial segregation, especially when mixed with income inequality, may lead to unequal education and employment opportunities, the persistence of income and wealth inequality, and poverty traps. As Cutler and Glaeser [1997] demonstrate, blacks living in urban ghettos have reduced social and economic outcomes, such as lower high school graduation rates and labor market earnings. In general, severe segregation is associated with lower social well-being.

Clearly people choose neighborhoods for many reasons beyond the racial and ethnic composition. In general, people weigh type-based preferences along with other location-based characteristics, such as the quality of the public schools, and the types of nearby stores. In addition, the utility derived from living in a specific neighborhood or community can be determined, in part, by the degree to which residents have positive interactions with their neighbors.

In this paper, we consider an extension of the Schelling model by also having agents play a repeated Prisoner's Dilemma (PD) game with their neighbors. In short, agents will determine their location choice by the outcome of both the Schelling and PD games. Our interest here is in expanding the Schelling model to include other features beyond type-based preferences that can determine the residents' quality of life. People's utility derived from their residential choice is determined, in part, by the *actions* of their neighbors. In many cases, the actions of residents and their neighbors are endogenous. One example is that of property maintenance. As Robert Frost has written, "good fences make good neighbors": if one neighbor does not contribute to the maintenance of a common fence, it will reduce the incentive of the other to maintain it as well.[1] Another example is parental involvement in the public schools. The more parents are involved, the more it will confer a positive benefit upon everyone in the school: increased student performance, a better sense of community, etc. The success of these neighborhood outcomes depends on agents' willingness to play a cooperative strategy in a neighborhood game.

Social psychologists have documented a relationship between residents' sense of community, "neighboring," and personal well-being. Sense of community — a psychological perception of how well neighbors get along — has been found to be associated with a greater sense of personal well-being [Farrell et al. 2004]. Neighboring is the exchange of goods and services among neighbors, such as the giving of information about good plumbers, the lending of power tools, or the provision of aid in an emergency. The willingness of neighbors to engage in these trades can directly influence residents, as the standard gains-from-trade models show, but can also improve the sense of community and, therefore, well-being [Farrell et al. 2004]. In addition, demographers have documented a negative relationship between people's sense of community (including the amount of neighborhood turnover) and their desire to move [Lee et al. 1994; Clark and Ledwith 2005]. Thus, one can envision an endogenous relationship between neighboring and mobility.

Recent research in economics and sociology has investigated the effect of social capital and trust on agent behavior [Glaeser et al. 2000]. At the country level, greater degrees of trust among citizens have been found to increase economic growth and to decrease corruption. As well, research findings suggest that dense social networks can sustain trust; while interactions between different racial groups are often characterized by lower degrees of trust [Glaeser et al. 2000].[2] Marschall and Stolle [2004] found, in a sample of neighborhoods in the Detroit area, that there is a strong relationship between race and feelings of trust (holding income constant). They found, for example, that "neighborhood racial heterogeneity and neighborhood sociability significantly increase blacks' propensity to trust others" [p. 146].[3] Thus, social interactions among agents of different races or ethnicities can foster trust.

It is within this context that we introduce the PD game into the Schelling model. Schelling's original model was designed to show how, even with strong preferences for integration, segregation was the only stable equilibrium. Our aim is to demonstrate that by introducing a model where cooperation (and therefore trust) among agents can develop endogenously, integrated neighborhoods can be an obtainable and stable equilibrium. Our goal then is to introduce a social dilemma game into the type-based interactions of a Schelling model in order to view how the potential maintenance of cooperation may impact the resulting levels of segregation. Of course, there are many possible social dilemmas that could be introduced. We choose to introduce a PD game due to its wide study in the social sciences and because it is a clear example of a social dilemma game.

As many models have shown (discussed below), cooperation can be a sustainable outcome in a repeated PD framework under certain conditions. We view the emergence of cooperation here as the development of neighborhood trust and, also by extension, as the gains that are available to neighbors when they engage in neighboring. For example, one can imagine doing a favor (at one's personal cost) for one's neighbor and implicitly expecting that your neighbor would do a similar favor for you (at their cost). Thus, if one were to engage in providing favors that are not reciprocated then one has sacrificed the cost of doing these favors to the benefit of their neighbor and received the traditional "sucker's payoff" in a PD game while their neighbor has received the "cheater's payoff." Or in an even simpler example, one may simply smile or wave at a neighbor as a sign of neighborly goodwill. A reciprocal greeting brings about better overall neighborhood relations but a lack of a reciprocal greeting may lead one to feel taken advantage of and potentially offer rewards of superiority to the shirker.

One could imagine modeling a game for each of the potential interactions that occurs among neighbors. But we want to keep the model as simple as possible and also use a game that is well understood. Therefore, we collapse these potential interactions into a simple PD game and allow the reader to broadly interpret the game as representing many neighborhood interactions, such as the loaning of power tools, the provision of aid, and/or just being friendly.[4]

The reason the PD can be important within the Schelling game is that cooperation can potentially offset the loss of utility that neighbors receive when they live with neighbors of a different type. Our aim is to investigate under what conditions this can hold, and to what degree we can view segregation and cooperation as substitutes. That is to say, to what extent does the emergence of trust among neighbors offset or remove negative utility from living with different-type neighbors? To simplify matters, we model an equal proportion of agent types, as well as an equal initial proportion of agents who are "cooperators" and "defectors." Certainly the interaction of agents can be more complex when one group is a minority and the other is a majority. Sociological research has found that black and whites in the US have different attitudes toward both integration and toward trust of neighbors [Marschall and Stolle 2004]. We leave this complicating variation of the model for future work.

Here, we will show that low levels of segregation can be supported in our model if high levels of cooperation can also be supported as an outcome of the PD game. On the other hand, our model may generate even higher levels of segregation than are produced in the Schelling model when all agents defecting is the outcome of the PD game. Thus, our model leads to the conclusion that increasing social interactions can be helpful in reducing segregation if the process yields cooperation. But, social interactions should be limited if the interactions lead to non-cooperative outcomes. To the best of our knowledge no other paper has explored the effect that neighborhood cooperation can have in affecting the instability of integration in the Schelling model.

The paper proceeds as follows. The next section discusses the related literature. Then the third section discusses the Schelling model with the inclusion of a utility function for agents. Next, in the fourth section we introduce the repeated PD game and the probability rules agents use in choosing whether to cooperate or not. Then, in the fifth section, we provide the model and results of the combined Schelling and PD game. Finally, the sixth section provides some concluding remarks.

RELATED LITERATURE

Schelling models

The Schelling model has become one of the central models in the development of agent-based economics [see Pancs and Vriend 2007]. It is widely cited as a pedagogical tool partly for the simplicity of the model and partly for the intrigue of the central result. Modest preferences for living near one's own type can lead to high levels of residential segregation; this result has proven to be quite robust to different versions of the model.

Most recently Pancs and Vriend [2007] subject the Schelling model to a series of computational specifications (one dimension *vs* two dimensions, line *vs* ring, checkerboard *vs* torus) and find that the central segregation result does not depend strongly on any of these specifications. Further they expand the range of preferences incorporated in the Schelling model. In the original model, the preferences of Schelling's agents were asymmetric; agents were opposed to being a member of a small minority group in their neighborhood but not to being a member of a large majority group. In short, agents did not prefer segregation but they did not oppose it either as long as they were in the majority.

A main contribution of Pancs and Vriend [2007] is in their expansion of the utility functions considered in the Schelling model. Specifically, they remove the above-described asymmetry by considering a utility function where agents prefer integration to both being in a neighborhood where they are a small minority and being in a neighborhood where they are a large majority. Agents have single peaked, tent-shaped preferences where a 50–50 integrated neighborhood is strictly preferred to all other neighborhood compositions (even one with an agent's neighborhood being composed exclusively of their own type). Somewhat amazingly, they find that the Schelling model still produces large amounts of residential segregation.

Recently the Schelling model also has been subjected to more rigorous analytical analysis from the perspective of evolutionary game theory [Zhang 2004; Dokumaci and Sandholm 2006]. This work attempts to understand the creation of segregation in the Schelling model from a formal analytical perspective.

The Prisoner's Dilemma

Studies of the maintenance of cooperation in the PD are vast. Economists are long familiar with the "folk theorem" [Fudenberg and Tirole 1991], which says, in part, that agents can maintain cooperation in an infinitely repeated PD as long as the future is not discounted too heavily.

More recently, other means of maintaining cooperation in the PD have been studied. Examples range to include reputation [Nowak and Sigmund 1998], reciprocity [Axelrod 1984], the use of tags and signals to recognize opponent types [Riolo 1997; Riolo et al. 2001; Hales 2001; Janssen 2008], and withdrawal from play [Aktipis 2004; Janssen 2008]. To keep this model as simple as possible, we do not incorporate tags or signals, but our model is amenable to these extensions, since it is reasonable to assume that agents may be likely to play different strategies based on whether they are playing with their same type or not.

Most closely related to the implementation of the PD in our paper is the work on the maintenance of cooperation in spatial models [Nowak and May 1992; 1994;

Schweitzer et al. 2002]. In these models, it is shown that repeated interaction with local neighbors may lead to the evolution of cooperation in the PD game.[5]

These spatial models are especially pertinent for our purposes because the Schelling model is clearly spatial as well. Therefore, we have chosen to follow this line of research to implement a model of the PD into the Schelling model. Specifically, we have chosen a framework close to that of Nowak and May [1994] for our model. Agents in the traditional Schelling model choose to move or stay based on agent types in a specified local neighborhood. In our model agents will also consider a second element in their utility function: the outcomes of a PD game with those agents located in the same specified neighborhood. We describe our model in more detail in the sections below.

THE SCHELLING MODEL WITH UTILITY

Our paper will consist of three models. The first will be a stand-alone version of the Schelling segregation model and the second an implementation of a neighborhood-based PD model. We present the results of these two models independently in order to have a benchmark for comparison to our third model that combines elements of both of the individual models.

In our first experiment, we implement the standard Schelling model with a utility function. We have two types of agents. Each type of agent is different in some recognizable way from the other type, and all else equal, each agent has a utility function over the proportion of agents of the same type. We assume here that all agents have the same utility function.

In our implementation we use a 12×12 grid with 14 spots (roughly 10 percent) left empty. Sixty-five of the agents are one type, and sixty-five are the other. Each agent interacts with the agents within her ‘‘Moore’’ neighborhood, that is, an agent’s neighborhood consists of the eight surrounding agents. If an agent is in a corner or on an edge, then the agent will have fewer neighbors. The lattice is not a ‘‘wrapped’’ around torus. We make this assumption because we believe this is a more accurate model of the geographic/spatial patterns observed in physical residential neighborhoods.[6]

An agent’s utility is determined based on the proportion of neighbors that are of the same type. The function is given by

$$u(s) = \left\{ \begin{array}{ll} \alpha + 2(\beta - \alpha)s & \text{if } s \in [0, 0.5) \\ \beta & \text{if } s \in [0.5, 1] \end{array} \right\}$$

where $s \in [0, 1]$ is the proportion of neighbors that are the same as the agent. This utility function is increasing from 0 to 0.5 and peaks at 0.5, where it remains flat after that. To make the game in this section comparable to the fifth section, over different runs, we change the values of α and β, but preserve the difference between them. Initially $\alpha = -1, \beta = 0$, then we raise them in fixed increments, for example, to $\alpha = -0.9$, $\beta = 0.1$, so that a portion of the utility function is above zero. Figure 1 presents a graph of the utility function for three $\alpha - \beta$ pairs.

Initially, agents are randomly distributed on the lattice. Then an agent is selected in turn, and the agent’s utility is calculated.[7] If the agent’s utility is less than zero, the agent moves to a randomly selected open location; if the utility is greater or equal to zero the agent stays. We repeat this process for each agent in the model. Each ‘‘run’’

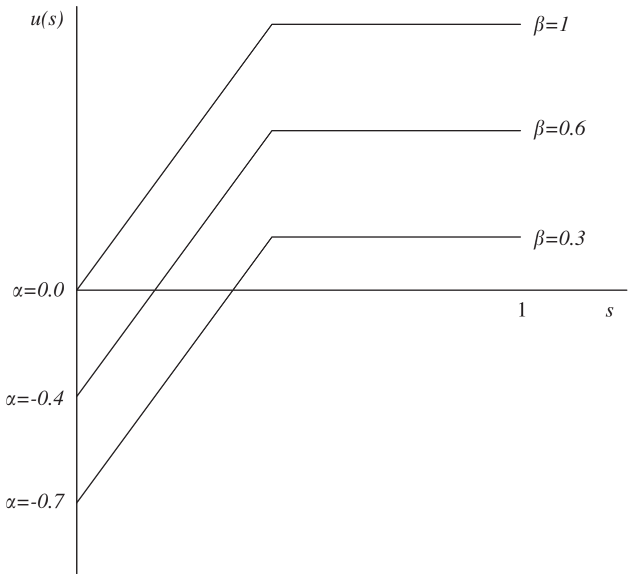

Figure 1. Utility *vs* segregation levels for different α–β combinations.

of a game goes until 100,000 iterations or when agents stop moving, whichever comes first.[8]

In Schelling's [1971] model, agents determine if they are satisfied by looking at the proportion of neighbors who are like them; if the proportion of like-typed agents is below a certain threshold (e.g., 50 percent) they move. The process then continues until no agent wants to move; in the end, the board is typically highly segregated. As is now well known, the ability of agents to leave neighborhoods when they are not satisfied with the proportion of agents like them causes a "tipping" dynamic that generates a highly segregated outcome.

Notice how the utility function that we use creates a threshold similar to the original Schelling model. Here, agents move if $u(s) < 0$, which implies an agent moves if $s < (-\alpha/2(\beta - \alpha))$; this is the intersection of the utility function and the horizontal axis. Thus, by changing α and β we create different Schelling thresholds. For instance, $\alpha = 0$, $\beta = 1$ gives a threshold of 0 and $\alpha = -0.5$, $\beta = 0.5$ gives a threshold of 1/4.

As a measure of segregation we define the *Similarity Rate* as the average proportion of same-type neighbors for each agent across all neighborhoods.[9,10] Denote the agent type as t. Let the agent of one type be $t = 0$; and an agent of the other type be $t = 1$, then we have the following definition for the Similarity Rate:

$$\textit{Similarity Rate} = \frac{1}{N}\sum_{i=1}^{N} s_i$$

$$s_i = \left[\frac{1}{n_i}\sum_{j=1}^{n_i} |1 - t_{i-} t_{ij}|\right]$$

where t_i is agent i's type and t_{ij}, $j = 1, \ldots, n_i$ are the types of agent i's neighbors; N is the total number of agents. Thus, s_i is the proportion of neighbors that are the same as agent i.

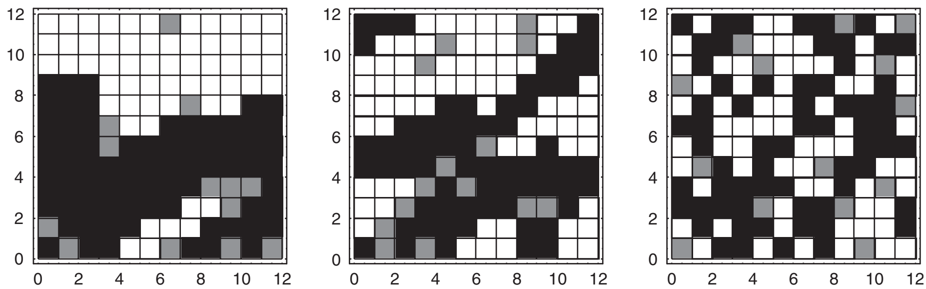

Figure 2. Three outcomes of the Schelling game for different $\alpha-\beta$ combinations. Left: $\alpha=-1$, $\beta=0$. Middle: $\alpha=-0.7$, $\beta=0.3$. Right: $\alpha=-0.2$, $\beta=0.8$

Table 1 Percent Similarity as a function of β ($\alpha=\beta-1$)

β	*Implied threshold*	*% Similarity*
0.0	0.50	87.54
0.1	0.45	87.48
0.2	0.40	82.85
0.3	0.35	77.26
0.4	0.30	76.10
0.5	0.25	60.67
0.6	0.20	56.49
0.7	0.15	55.56
0.8	0.10	52.06
0.9	0.05	52.08
1.0	0.00	49.79

Average of 200 runs.

Figure 2 gives examples of typical runs for three $\alpha-\beta$ combinations to demonstrate how the utility functions can affect segregation. We can see that $\alpha-\beta$ directly affects the equilibrium level of segregation. The white squares are type 0 agents, the black squares are type 1 agents, and the gray squares are empty.

Average levels of Similarity over 200 replications are given in Table 1. The table illustrates a few facts. Segregation is monotonically increasing in β and decreasing in the threshold, as would be expected. Notice that when $\beta=0$ and the threshold is 0.50, 87.5 percent of neighbors are of the same type on average.

REPEATED PRISONER'S DILEMMA

Now consider a spatial repeated Prisoner's Dilemma (RPD) game. As is standard for the PD game, agents choose to play one of two actions: Cooperate or Defect. The general payoffs of the game are given in Table 2.

As is well known, a one-shot PD has only one Nash equilibrium: {Defect, Defect}. However, as is also well known, in a RPD game when the game will not end with certainty, the folk theorem says that any outcome in a RPD is a sustainable equilibrium given appropriate discount values. As mentioned in the Introduction, there is now a large literature beyond the folk theorem on the maintenance of cooperation in the PD. Here we use a framework similar to that of Nowak and May

Table 2 PD payoffs for an agent, where $C > A > D > B$, and $(C+B) < 2A$

	Rival cooperates	*Rival defects*
Agent cooperates	A	B
Agent defects	C	D

[1994] to combine an RPD game into the Schelling model. We note that this is just one possible choice of many for incorporating a PD game that has the possibility of resulting in cooperative outcomes. Although we will report results on the outcome of the RPD game our primary focus is on the segregation outcomes when this game is combined with the standard Schelling model. Thus, we restrict attention in this paper to our chosen implementation of the RPD and save additional comparisons to other implementations for future work.

We begin with a stand-alone version of the RPD in order to have a benchmark for later comparison. In this paper, we explore a game where in each round each agent plays a one-shot PD game with each of their neighbors. To simplify matters, we will set $C = A + \varepsilon$ and $D = B + \mu$, where μ, $\varepsilon > 0$, and we fix A, B, and μ (μ is fixed at a small value so as to preserve the PD payoff structure), and we change ε over different experiments. As ε increases, agents have a greater incentive to defect or "cheat" against a cooperating opponent; thus sustaining cooperation becomes less likely.

THE RPD ON THE LATTICE: THE MODEL

In this paper, agents play with their neighbors who, as discussed above, are within each agent's "Moore" neighborhood. Each agent has a probability, $p_i \in [0, 1]$, of playing cooperate (and $1 - p_i$ is the agent's probability of playing defect). In this section, agents do not move.[11] When an agent is selected, she chooses an action $\{C, D\}$ according to her probability distribution function. Once an action is selected, she plays that same action with all her neighbors for that round. She plays the PD with each neighbor, who selects an action based on their own probability distribution function.

For any two players, define the payoff to agent i when playing a neighbor agent j as

$$\pi_i(x_i; x_{ij}) = Ax_i x_{ij} + Bx_i(1 - x_{ij}) + C(1 - x_i)x_{ij} + D(1 - x_i)(1 - x_{ij})$$

where x_i, $x_{ij} = 1$ if a player cooperates; 0 if the player defects. x_{ij} represents the action of agent i's neighbor, agent $j \in \{1, \ldots, 8\}$. If we set $C = A + \varepsilon$ and $D = B + \mu$, and rearrange terms, equation (1) can be written as

$$\pi_i(x_i; x_{ij}) = Ax_{ij} + B(1 - x_{ij}) + [\varepsilon x_{ij} + \mu(1 - x_{ij})](1 - x_i)$$

If agent i plays against n_i neighbors, then we have the average payoff as:

$$\bar{\pi}_i(x_i; x_{ij}) = \frac{1}{n_i}\sum_{j=1}^{n_i} \pi_i(x_i; x_{ij}) = A\rho_{ij} + B(1 - \rho_{ij}) + [\varepsilon\rho_{ij} + \mu(1 - \rho_{ij})](1 - x_i) \quad (2)$$

where $\rho_{ij} = 1/n_i \sum_{j=1}^{n_i} x_{ij}$. Thus $\bar{\pi}_i(x_i; x_j) = \pi_i(x_i; \rho_{ij})$. That is, agent *i's* average payoff is determined by the proportion of neighbors playing cooperate. Notice that the agent will receive a weighted sum of $A\rho_{ij} + B(1-\rho_{ij})$ regardless of her action. However, if the agent chooses to defect, she also will receive a "defectors bonus" given by $\varepsilon\rho_{ij} + \mu(1-\rho_{ij})$, a weighted sum of the additional payoffs from defection.

Each time an agent finishes playing with her neighbors she updates her p_i. To update p_i we use a rule similar to that of Nowak and May [1994], which is given by the following formula (we do not include time subscripts to simplify notation):

$$p_i = \frac{\sum_{j=1}^{n_i} \pi_{ij}(x_{ij}; x_i) x_{ij}}{\sum_{j=1}^{n_i} \pi_{ij}(x_{ij}; x_i)}$$

where x_i, $x_{ij} = 1$, if an agent cooperates, 0 otherwise.[12] The rule is an imitation rule where agents change their probability based on the strategies of their neighbors. Note how this rule allows for the development of social capital in that it implicitly allows for reciprocity: if a high proportion of neighbors cooperate, then the agent will be more likely to cooperate the next time the agent and her neighbors interact. Reciprocity has been found to be a common behavior in economic activities [Fehr and Gächter 2000].

Thus, as "cooperate" ("defect") becomes a more lucrative action for the neighbors of agent *i*, p_i increases (decreases). Also note that if all of an agent's neighbors play cooperate (defect) then the agent will play cooperate (defect) in the next round with probability 1. The system reaches an absorbing state whenever all agents play the same action in a given round.

This formula can be rewritten according to equation (2) as

$$p_i = \frac{[Ax_i + B(1 - x_i)]\rho_{ij}}{[Ax_i + B(1 - x_i)] + [\varepsilon x_i + \mu(1 - x_i)](1 - \rho_{ij})} \tag{3}$$

Notice that, *ceteris paribus*, $\partial p_i/\partial \rho_{ij} > 0$, which means that an increasing level of cooperation in the neighborhood of agent *i* increases the probability that agent *i* will cooperate in the next round. This can generate a type of "cooperative tipping." As demonstrated below, if ε is not too large, all agents having $p_i = 1$ can become an equilibrium (absorbing state).

Equilibria

For the probability updating rule given by equation (3), there are two pure strategy equilibria (everyone plays cooperate or everyone plays defect) and no mixed strategy equilibrium.

Claim 1 There are two pure strategy absorbing states of the model, $p_i = 1$ for all i and $p_i = 0$ for all i, and no mixed strategy absorbing states.

Proof To prove this, we begin with identifying the value of ρ_{ij} where player *i* will be indifferent between playing $x_i = 1$ or $x_i = 0$; we set $(p_i|x_i = 1; \rho_{ij}) = (p_i|x_i = 0; \rho_{ij})$. This gives the condition

$$\frac{B\rho_{ij}}{B + (1 - \rho_{ij})\mu} = \frac{A\rho_{ij}}{A + (1 - \rho_{ij})\varepsilon}$$

Rearranging terms gives

$$\rho_{ij}(1-\rho_{ij})\mu A - \rho_{ij}(1-\rho_{ij})\varepsilon B = 0 \quad (4)$$

Assuming $A/B \neq \varepsilon/\mu$, then the only solutions to equation (4) are $\rho_{ij}=1$ and $\rho_{ij}=0$, $\forall_i$. ■

Next we turn to the selection of these equilibria over repeated plays.

THE RPD ON THE LATTICE: EXPERIMENTS

Here, we fix the payoffs (given in Table 3) and vary $\varepsilon \in \{0, 0.01, 0.02, \ldots, 0.10\}$. Given the probability updating rule, there is a range of ε and μ where everyone cooperating is an attainable absorbing state (where "attainable" here means that it occurs with non-zero probability over repeated runs of the RPD game).

To perform the experiment, we assign half the agents an initial probability of 1 and the other half a probability of 0.[13] As discussed above, each agent is chosen in turn, and he randomly picks an action according to his probability function. He then plays all his neighbors, who also randomly pick an action. Note the agent selects an action once and plays the same action for that round. After he plays the game with all his neighbors, he observes his rivals' payoffs and updates his probability, p_i. We run the system for 100,000 iterations or until the system reaches an absorbing state, whichever comes first.[14] We then take averages of 200 runs to smooth out fluctuations. Two examples are given below in Figure 3, where $C=5.05$. The figures present, for each iteration, the average of all agents' probability of cooperation.

Table 4 presents the proportion of runs (out of 200) that end in the game hitting an absorbing state where all agents playing defect with probability 1. As the results show, the smaller the value of ε (the smaller the value of the "cheat" payoff, C), the smaller the proportion of runs that end in all agents defecting; the remaining runs

Table 3 Payoffs for RPD experiment

	Rival cooperates	*Rival defects*
Agent cooperates	5	3
Agent defects	$5+\varepsilon$	3.01

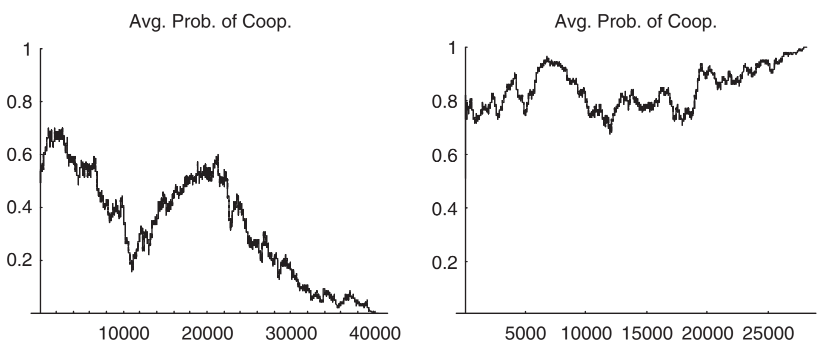

Figure 3. Two outcomes for the RPD game, when $C=5.05$. Average probabilities of all agents.

Table 4 Percentage of RPD games that end in all agents defecting with probability 1

C Payoff	*% Defect*
5.00	30.5
5.01	46.6
5.02	55.4
5.03	58.3
5.04	72.1
5.05	81.2
5.06	82.6
5.07	89.8
5.08	91.5
5.09	93.5
5.10	97.9

end with all agents cooperating. As shown in the table, cooperation can be attained in a large percentage of runs as long as the incentives to cheat are sufficiently small.

COMBINED SCHELLING AND PD GAMES

After having demonstrated the results in the stand-alone Schelling model and our implementation of the RPD we now move to analyzing the combined game. Recall that the purpose of this section is to analyze the effect of non-type based social interactions in a Schelling model. Here we model the additional social interactions as a RPD game with neighbors as defined in the Schelling model.

In the combined game, we have total per-period utility given by:

$$U(s,x) = \phi u(s) + \pi(x;\rho) \quad (5)$$

where $x=1$ if an agent cooperates, and $x=0$ if an agent defects; ρ is the proportion of neighbors who cooperate. $\phi>0$ scales the relative importance of the Schelling portion of utility to the RPD portion of utility. As ϕ increases the importance of the Schelling game increases. Notice a few simplifying assumptions that we make. First, the utility function is simply the sum of two game payoffs; there is no interaction effect between the two parts. Certainly we could imagine a situation where the payoff to the PD would be determined, in part, by the number of like-type agents; and perhaps cooperation could be more likely if an agent is playing with mostly her own type as in the signalling or tag-based examples of cooperation in the RPD mentioned in the Introduction. In addition there is no asymmetry in utility across types. We leave these complicating variations for future work.

As in the original Schelling model, we will define a threshold utility above which agents are satisfied and remain at their current location; agents whose utility is below the threshold are not satisfied and subsequently move. In this implementation we set the entire Schelling portion of utility to be negative and the RPD portion of utility to be positive and set the threshold at 0. Thus, agents need to have sufficiently positive RPD utility in order to be satisfied. Further, we can use ϕ to scale the importance of the two components of the utility function. One also can think of changes to ϕ as an indirect adjustment of the movement threshold. As we increase ϕ we increase the weight on the negative portion of utility and thereby increase the movement threshold indirectly. As ϕ increases, either agents need to achieve larger

RPD payoffs or live in a neighborhood with more like-typed agents in order to be satisfied.

This general framework means that agents gain utility by moving to neighborhoods with higher concentrations of their type, and/or by learning to cooperate or defect depending on the actions of their neighbors. We show here that tradeoffs exist between the two portions of utility: a high cooperating neighborhood can offset living in a neighborhood with few of an agent's same type.

Recall that for the Schelling game, the utility function is rising with the fraction of the same type until 1/2, where it is flat thereafter. Here, to make the utility from the two games directly comparable, we set α and β equal to minus the RPD payoffs, $-A$ and $-B$, respectively, and scale the Schelling utility according to the parameter ϕ:

$$u(s) = \left\{ \begin{array}{ll} -\phi[A + 2(B - A)s] & \text{if } s \in [0, 0.5) \\ -\phi B & \text{if } s \in [0.5, 1] \end{array} \right\} \tag{6}$$

where $\phi > 0$ is the relative weight given to $u(s)$ in the combined utility function.

COMBINED UTILITY

If we substitute the Schelling utility and the RPD payoff function into equation (5), and rearrange terms, we can write the combined utility function as (with subscripts dropped)

$$U(s, x; \phi) = \left\{ \begin{array}{ll} (A - B)\rho + A(1 - \phi) + (A - B)(\phi 2s - 1) + [\varepsilon\rho + \mu(1 - \rho)](1 - x) & \text{if } s \in [0, 0.5) \\ (A - B)\rho + B(1 - \phi) + [\varepsilon\rho + \mu(1 - \rho)](1 - x) & \text{if } s \in [0.5, 1] \end{array} \right\}$$

Here combined (total) utility may be positive or negative depending on the values of ρ, s, and ϕ. First, consider the case where $s \in [0, 0.5)$. The first and last terms are clearly positive, while the middle two terms can be positive or negative depending on the size of ϕ. Below we trace through the various elements of the utility function more carefully in order to illuminate the computational results that follow.

The case of $\phi = 1$

For a moment, suppose that each of the two components of utility have equal weight, this is the case when $\phi = 1$. Therefore, we have

$$U(s, x; \phi) = \left\{ \begin{array}{ll} (A - B)\rho + (A - B)(2s - 1) + [\varepsilon\rho + \mu(1 - \rho)](1 - x) & \text{if } s \in [0, 0.5) \\ (A - B)\rho + [\varepsilon\rho + \mu(1 - \rho)](1 - x) & \text{if } s \in [0.5, 1] \end{array} \right\}$$

Recall that if an agent cooperates then $x = 1$ and if an agent defects then $x = 0$. Consider the case when $s \geqslant 0.5$. In this case the Schelling utility is a constant and only the level of ρ and the agent's action (cooperate or defect) determine utility. In the case where $\phi = 1$ and $s \geqslant 0.5$ the agent always has positive utility and will not move regardless of the outcome of the PD game with her neighbors.

Now, consider the case where an agent cooperates ($x = 1$) and we have $s < 0.5$. In this case, the utility function is given by $(A-B)\rho + (A-B)(2s-1)$. Therefore, utility will be positive and an agent will not move if $(A-B)(\rho + 2s-1) > 0$. This implies that the agent needs $\rho + 2s > 1$ in order to be satisfied. Note that there is a substitution effect between the amount of cooperation in the neighborhood and the fraction of like-typed neighbors. High levels of cooperation can induce an agent to remain in a neighborhood in which she is a minority member. Further, if $\rho = 1$ then any level of s will yield utility greater than or equal to 0. In this case the agent will not move. In sum, when $\phi = 1$, the all-cooperate absorbing state makes an agent satisfied even if he is a part of a very small minority in his neighborhood.

Now, if the agent defects ($x = 0$) and $s < 0.5$, then the agent receives an added utility bonus of $\varepsilon\rho + \mu(1-\rho)$. Thus, a defecting agent is more likely to have positive utility and stay than a cooperating agent, but the tradeoff between cooperation and like-typed agents still exists. Agents will stay in a non-cooperating neighborhood if they have sufficient like-typed neighbors and agents will stay in a neighborhood that has few like-typed neighbors if there is sufficient cooperation.

Notice however that the dynamics of this process are complex. Even though there exists a tradeoff between cooperation and having sufficient like-typed neighbors, a neighborhood with high levels of cooperation is attractive to defecting agents. A defecting agent in a highly cooperative neighborhood will receive large payoffs and thus lead to other agents becoming more likely to defect themselves. This would lead to agents wanting to move to neighborhoods with more like-typed neighbors. In the remainder of this paper we will investigate the equilibrium conditions and the dynamics of this process.

Equilibria when $\phi = 1$

We proceed to discuss an equilibrium of the combined game when $\phi = 1$. An equilibrium requires that (1) the RPD component of the game has reached an absorbing state, and (2) all agents in the population have utility greater than or equal to 0 so that no agent changes location. Above, we sketched the requirements for agents not to move. We can now state these requirements a bit more formally as functions of s and ρ for a cooperating and a defecting agent in a combined game equilibrium.

In an equilibrium where all agents cooperate, $x = \rho = 1$ for each agent. Recall that when $x = 1$, for no movement to occur we need $\rho + 2s > 1$ when $\phi = 1$. Since $\rho = 1$ in equilibrium, this inequality always holds in an all-cooperate equilibrium. This means that an agent will accept any level of integration or segregation at this equilibrium. But the level of segregation that emerges will be a result of the complex, stochastic dynamics that unfold.

On the other hand, if we have the all-defect equilibrium then $x = \rho = 0$. In this equilibrium, an agent will not move if $s \geqslant 1/2(A-D/A-B)$, since this is the value of s for which utility is greater than or equal to zero. Since $D > B$, this threshold will be achieved when s is close to, but less than, $s = 1/2$. That is to say, even in the case where everyone defects, less segregation than in the "pure" Schelling case is possible. Notice that if D were to be increased, it would allow an even lower level of segregation to exist.

If we compare this to the equilibrium with all agents cooperating we see that an equilibrium where all agents defect requires a sufficient level of like-typed neighbors. The all-cooperate equilibrium can be sustained for any level of s. If we think about these equilibrium outcomes in the spirit of Schelling's original model we see that an equilibrium with all-defect being the RPD outcome implies that there will be a non-zero threshold for like-typed neighbors for each agent. Thus, we are in a situation very similar to the original Schelling model and should expect high levels of segregation as in the original model. On the other hand, if we reach an all-cooperate absorbing state of the RPD then this threshold for like-typed neighbors is removed thereby creating the possibility of integration in the Schelling model.

This leads to the following questions: (1) Under what conditions can the all-cooperate absorbing state be reached in the combined game? (2) If the all-cooperate absorbing state is reached, what level of segregation will be attained? Recall that the all-cooperate absorbing state can co-exist with any level of segregation when $\phi = 1$. Integration is possible but segregation is possible as well. We next describe the equilibrium conditions for the more general model (we allow for $\phi \neq 1$) and then proceed to explore the parameter space of our model using simulations.

THE GENERAL CASE

We now move to the general case with $\phi > 0$. Recall that an agent's utility function is:

$$U(s,x;\phi) = \left\{ \begin{array}{ll} (A-B)\rho + (A-B)(\phi 2s - 1) + A(1-\phi) + [\varepsilon\rho + \mu(1-\rho)](1-x) & \text{if } s \in [0, 0.5) \\ (A-B)\rho + B(1-\phi) + [\varepsilon\rho + \mu(1-\rho)](1-x) & \text{if } s \in [0.5, 1] \end{array} \right\}$$

Given whether the game is in an all-cooperate or all-defect state, we can solve for the levels of s as a function of ϕ and the PD payoffs that will create positive utility and thus give agents the desire to stay. We call this level of s the *Implied Schelling Threshold* of the combined game:

> **Definition** The Implied Schelling Threshold, s^*, is the minimum level of like-typed neighbors that can support a given equilibrium of the RPD game (all-cooperate or all-defect) in the combined game.

One should think of s^* as creating the same effect as the threshold in the original Schelling model. A larger threshold should produce more segregation than a smaller threshold. As we will see below s^* will be very important in understanding the outcomes of the model.

For an all-cooperate equilibrium s^* is given by:

$$s^* = \left\{ \begin{array}{ll} \frac{A(\phi-1)}{2\phi(A-B)} & \text{if } \phi \in \left(1, \frac{A}{B}\right] \\ 0 & \text{if } \phi \in (0, 1] \end{array} \right\} \quad (7)$$

In short, for $\phi \in (0, 1]$, any level of liked-typed neighbors can be supported in the all-cooperate equilibrium. For $\phi \in (1, A/B]$, the ability to sustain integration will depend on both the value of A/B and how close ϕ is to it. Thus, there is a non-zero Implied Schelling threshold level of like-typed neighbors that must occur to support the all-cooperate equilibrium for $\phi \in (1, A/B]$. But for $\phi < 1$, $s^* = 0$, that is, there is no threshold. Note that this meets with intuition: as ϕ increases the Schelling game becomes more important and we should expect a "Schelling-style" threshold to emerge as ϕ increases. Also note that for $\phi > A/B$, no all-cooperate equilibrium exists since utility is always less than 0 for all levels of s.

For an all-defect equilibrium, s^* is given by:

$$s^* = \left\{ \begin{array}{ll} \frac{A\phi - D}{2\phi(A-B)} & \text{if } \phi \in \left(\frac{D}{A}, \frac{D}{B}\right] \\ 0 & \text{if } \phi \in \left(0, \frac{D}{A}\right] \end{array} \right\} \quad (8)$$

Note that $s^* = 0$ at $\phi = D/A$. Thus, at this level of ϕ and below, any level of like-typed neighbors can be supported in an all-defect equilibrium. The threshold increases as ϕ increases up to $\phi = D/B$ where $s^* = 1/2$. Also note that the range of $\phi \in (0, D/B]$ is just the range for ϕ which preserves the structure of the Schelling and PD games. No all-defect equilibrium exists for $\phi > D/B$ since utility is always less than 0 for all levels of s.

If we compare s^* in the all-cooperate equilibrium to s^* in the all-defect equilibrium, we see that a lower ϕ is required in the all-defect equilibrium to have $s^* = 0$. This implies that there must be less weight on the Schelling portion of utility in an all-defect equilibrium in order to remove the threshold. This leads one to believe that in order to achieve integration there has to be a smaller weighting on the Schelling portion of utility in an all-defect equilibrium than in an all-cooperate equilibrium. We can also directly compare s^* for the entire range of feasible levels of ϕ.

Claim 2 (1) The Implied Schelling Threshold, s^*, is equal to 0 for $\phi \leqslant D/A$ in both the all-cooperate equilibrium and the all-defect equilibrium.
(2) The Implied Schelling Threshold, s^*, is lower in an all-cooperate equilibrium than in an all-defect equilibrium for $\phi \in (D/A, A/B)$.

Proof (1) is shown directly in equations (7) and (8) above. For (2) we need to consider two ranges of ϕ. For $\phi \in (D/A, 1]$, $s^* = 0$ in an all-cooperate equilibrium and $s^* = (A\phi - D/2\phi(A-B)) > 0$ for an all-defect equilibrium. For $\phi \in (1, A/B)$, we have $s^* = A\phi - A/2\phi(A-B)$ for the all-cooperate equilibrium and $s^* = A\phi - D/2\phi(A-B)$ for the all-defect equilibrium. Now, $A\phi - A < A\phi - D$ since $A > D$. Thus, the Schelling threshold is strictly lower in the all-cooperate equilibrium over this range of ϕ. ■

This claim states that s^* in an all-cooperate equilibrium is less than or equal to s^* in an all-defect equilibrium for all levels of ϕ. Thus, we should expect lower levels of segregation when an all-cooperate equilibrium is reached compared to when an all-defect equilibrium is reached.

COMBINED GAME RESULTS

We now present simulation results of the combined game. All results are averages over 1,000 runs of the simulations for each value of C. To review, agents are randomly distributed over the lattice; one-half of the agents are type 0, and the other one-half are type 1. Initially, one-half of the agents play cooperate with probability zero; the other one-half play cooperate with probability 1. The PD game payoffs are fixed according to Table 3 (note ε is a parameter that we change over different simulations). Equation (6) is the part of utility from the Schelling game.

Each agent is selected in turn, and the agent plays the PD with her neighbors according to her probability distribution function. The neighbors also choose an action according to their probability distribution functions. After the agent plays with her neighbors, her average payoff is determined. Next, the agent determines her "Schelling utility," given by equation (6). Then total utility is determined. If total utility is less than zero, the agent moves to a randomly chosen open spot. The game continues until an absorbing state is reached in the RPD *and* all agents have utility greater than or equal to 0 (thus no agent wishes to move.)

Cooperation and Similarity *vs* the cheating payoff

Figure 4 demonstrates the relationship between cooperation and segregation as we increase C, the payoff to "cheating," when $\phi = 1$. Here we see that as we increase C, there is a corresponding increase in the level of defection as occurred in the stand-alone RPD game. In addition we see that segregation also increases, on average, as C increases. Thus, we see a first indication of a direct relationship between segregation and defection.

The reason that Similarity and defection move in the same direction is because as we increase C there are relatively fewer games that end in everyone cooperating. When the game does hit an all-cooperation absorbing state, then, in fact, we see decreases in the Similarity Rate, for the reasons outlined above.

Figure 5 shows that games that end in all-cooperation have consistently lower Similarity values than when the game ends in all players defecting. Notice, however, that as we increase C, the all-cooperate case becomes increasingly rare, and therefore there is more variation in the Similarity measure. In sum, as we increase the temptation to cheat, if the players do in fact learn to cooperate, then they are also willing to live with relatively more agents of the other type.

To fully understand this result we need to return to the stand-alone Schelling game and the Schelling Threshold. Note that our utility function states that agents will stay if $-\phi[A + 2(B-A)s] + \pi(x;\rho) \geqslant 0$. If we insert $\phi = 1$ and the payoffs used for A and B we get that agents stay if $s \geqslant 5-\pi(x;\rho)/4$. Now consider the case where we are at an all-defect equilibrium. In this case the RPD payoff is $D = 3.01$ so the threshold would be $s \geqslant 1.99/4 \approx 1/2$. Recall from the stand-alone Schelling model that this implies a level of Similarity of 87.5 percent. The combined game yields a Similarity rate that is roughly the same for the most comparable stand-alone Schelling model. In other words, if we started the combined model with an all-defect equilibrium and an integrated population, the model would converge to roughly the same level of segregation as we get when starting with integrated neighborhoods.

Now, in the case of the all-cooperate equilibrium, the RPD portion of utility yields a payoff of five and agents will stay if $s \geqslant (5-5)/4 = 0$. In other words, the Schelling threshold is 0 and agents will be willing to live in a neighborhood with any

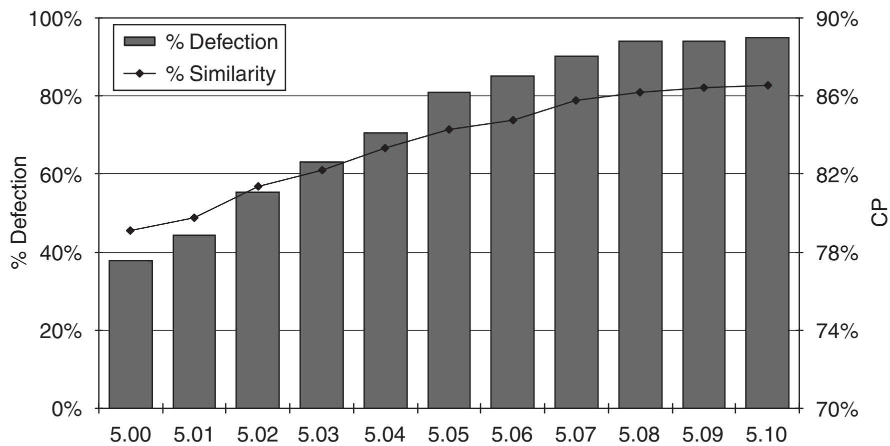

Figure 4. Percent Defection and percent Similarity *vs* C Payoff in Combined Game, $\phi = 1$.

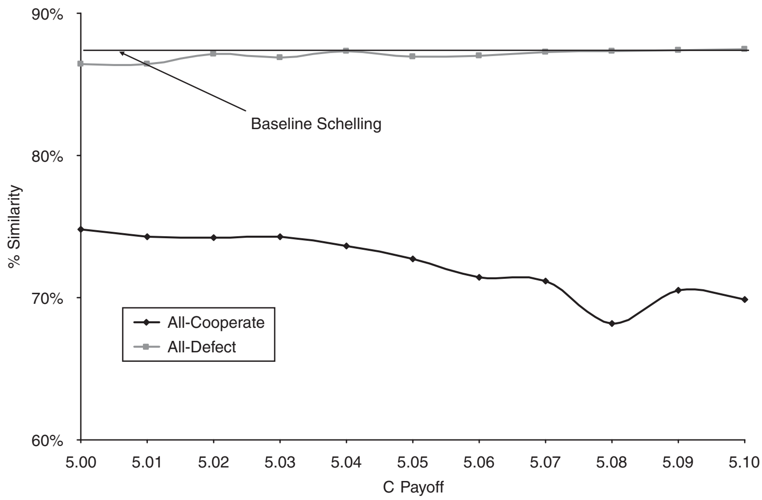

Figure 5. Percent Similarity *vs* C payoff, for all-coopering and all-defecting outcomes, $\phi = 1$.

composition of neighbors, if they all get the cooperation payoff. Thus, if we started the combined model with an all-cooperate equilibrium and an integrated population, the model would remain at complete cooperation and perfect integration. We observe in the simulations, however, that we get levels of segregation higher than perfect integration but also below that which is achieved in the all-defect equilibrium.

The analysis of the previous paragraphs tells us two things. First, the level of segregation achieved in the all-cooperate equilibrium is not simply the sum of the two processes. There is an interaction between the two games in the attainment of an equilibrium. Second, we can understand the level of segregation achieved in the combined game all-cooperate equilibrium by considering the dynamics of the process toward equilibrium. If we start with a mix of cooperators and defectors and an integrated population, there will be at least some agents who will be unsatisfied and move. As agents move, the population evolves toward an equilibrium in the RPD game, and individual agents have a propensity to increase their level of segregation as in the stand-alone Schelling game.

Consider the cases that converge to the all-defect equilibrium. These cases need to also converge to the level of segregation supported by the Schelling threshold, which is what we observe in the combined game. Now consider the case of convergence to the all-cooperate equilibrium. Here, agents could stop with perfect integration if they could immediately move to an all-cooperate equilibrium in the RPD game. But it takes time to reach this equilibrium and agents are creating greater levels of segregation as they move to the all-cooperate equilibrium. When they finally reach the all-cooperate equilibrium the process of creating more segregation immediately stops. Thus, the segregation that is reached is a direct function of the amount of time taken to reach the equilibrium. The faster the convergence to full cooperation the lower the level of segregation that is attained. This suggests that efforts to reach cooperation quickly may lead to lower levels of segregation.

Figure 6 illustrates the relationship between the percent Similarity reached in equilibrium and the number of periods needed to achieve this equilibrium for the all-cooperate outcome.[15] As an illustration of this, we present the case where $C = 5.05$,

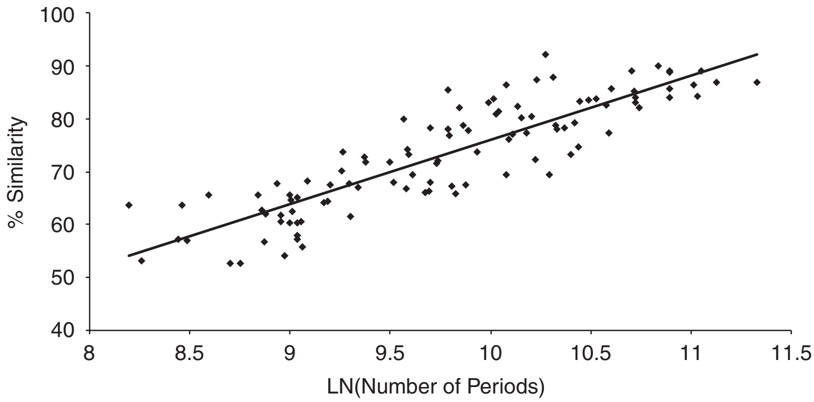

Figure 6. Percent Similarity *vs* number of periods till reaching an equilibrium, for the all-cooperate outcome. $C=5.05$, $\phi=1$. The regression line is percent Similarity $=-45.4+12.1Ln(\#\ periods)$. $\#obs.=104$, $R^2=0.76$. Both coefficients are stat. sig. at greater than the 99 percent level.

$\phi=1$ for 104 (out of 500) runs that ended with all cooperation. As we can see there is a positive and statistically significant relationship between segregation and time to convergence. In fact, the regression line shows that, on average, a doubling of the time to reach equilibrium is associated with a 12.1 percent increase in Similarity. Also note that in some extreme cases, when convergence is relatively rapid, the game ends in a situation that is quite close to perfect integration, which is about 50 percent Similarity.

Further, as C increases it becomes increasingly rare for the agents to coordinate on the all-cooperate outcome since with large cheat payoffs there are large incentives to move toward the all-defect absorbing state. Thus, in order to move to the all-cooperate absorbing state the system must get there relatively quickly because of the large incentives to move toward defection. This is what we observe in our simulations. As C increases the runs that converge to all-cooperate get there more quickly, on average, than when C is small.[16]

Cooperation and Similarity vs ϕ

Next we compare changes in segregation and defection rates as we increase ϕ. As ϕ increases more weight is placed on the Schelling portion of utility. Thus, one should expect more segregation as compared to the original Schelling model due to the increase in weight placed on the type-based portion of utility. As shown in Figure 7 this is exactly what we find: an increase in ϕ leads to more segregation. The figure presents the results for three different values of $\phi \in \{0.9, 1.0, 1.003\}$.[17]

Further, as we saw in the Equilibria section, the implied Schelling threshold of the combined game is increasing in ϕ. And as shown earlier in the paper, there is a direct relationship between segregation and defection. Again we see results that match expectations.[18]

Figure 8 shows the relationship between Similarity and the C payoff for the three different values of ϕ when the game ends in the all-cooperate equilibrium. In general, we see that a larger C value is associated with a lower Similarity value, and that increasing ϕ increases the Similarity rate.

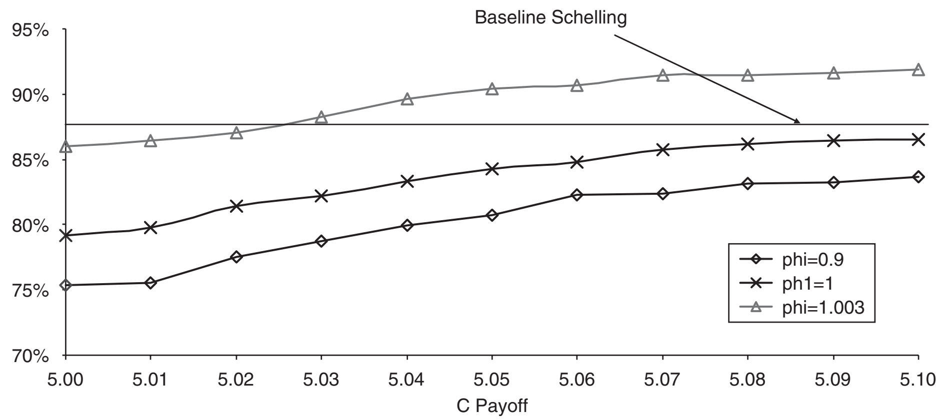

Figure 7. Percent Similarity *vs* C payoff for three different values of ϕ.

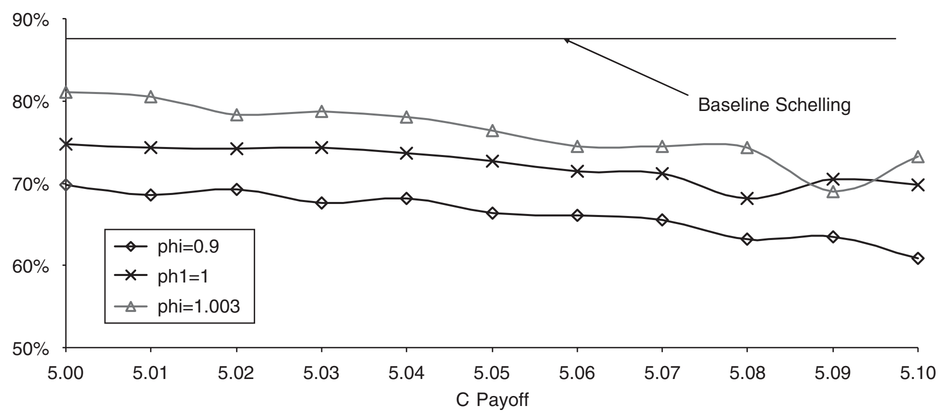

Figure 8. Percent Similarity *vs* C payoff for three different values of ϕ, when the game ends in all-cooperate.

UTILITY AND WELFARE

Given our definition of utility we can investigate which outcomes provide the greatest total utility or welfare for all the agents in the population, where welfare is given by

$$W = \sum_{i=1}^{N} U_i(s, x) = \sum_{i=1}^{N} [\pi_i(x; p) - \phi u_i(s)]$$

In equilibrium, the largest payoff an agent can receive will come if she is in the majority type in the neighborhood (she will get $-\phi B$) and if everyone is cooperating (she gets A). Thus, to make comparisons across games we normalize welfare by calculating

$$W' = \sum_{i=1}^{N} \frac{U_i(s, x)}{A - \phi B} \tag{9}$$

Note that this function only measures the welfare to agents in the game, but does not include any external costs or benefits that may accrue to society.

On average, as C increases, welfare will decline, since a larger fraction of games end in all-defect. Because of the plateau in the Schelling utility function, agents will be indifferent between any neighborhoods where the fraction of like-type agents is greater than or equal to 0.5. In all cases of the combined game discussed here, all equilibria that emerge have agents, on average, living in neighborhoods that have more than half the agents like themselves, thus the increasing proportion of games that end in all-defect is driving the drop in utility as C increases; this result can be seen in Figure 9.

Furthermore, the figure shows that as we decrease ϕ, we increase the total welfare available to agents, because when the game ends in the all-defect outcome, a smaller value of ϕ increases welfare. Reducing ϕ yields both a cost and benefit to total welfare. First, it increases the denominator and thus reduces welfare. But reducing ϕ also lowers the Schelling part of total utility. However, because segregation is higher (and virtually all agents will be on the plateau part of the Schelling utility function) in the all-defect case, lowering ϕ increases the numerator more than the denominator, and thus welfare rises.

However a different story occurs in the all-cooperate outcome. When we reduce ϕ, welfare decreases. However, reducing ϕ also lowers segregation, which reduces total welfare. In the end, a lower ϕ causes a net reduction in total welfare. The relationship between ϕ and welfare is illustrated in Figure 10 where $C = 5.01$.

However, when the game ends in all-cooperate, welfare is much higher than when the game ends in all-defect. Figure 11 demonstrates how welfare evolves *vs* the C payoff, for the case when $\phi = 1$. We see that when the games end in all cooperation, total welfare is roughly 95 percent of the maximum obtainable welfare. The reason is because all the agents are cooperating and, on average, a vast majority of the agents are in neighborhoods where they are in the majority (as Figure 5 demonstrates).

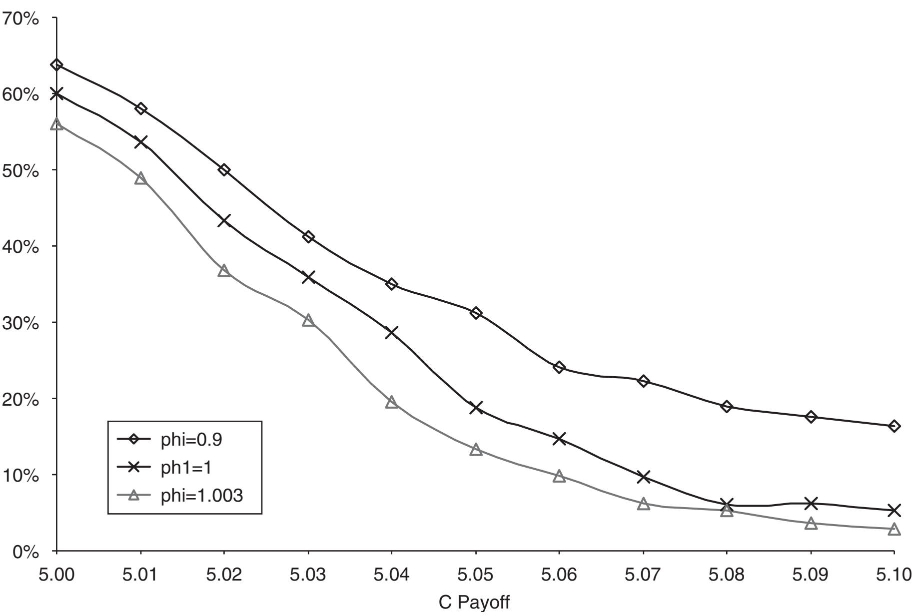

Figure 9. Welfare *vs* C payoff for different values of ϕ.

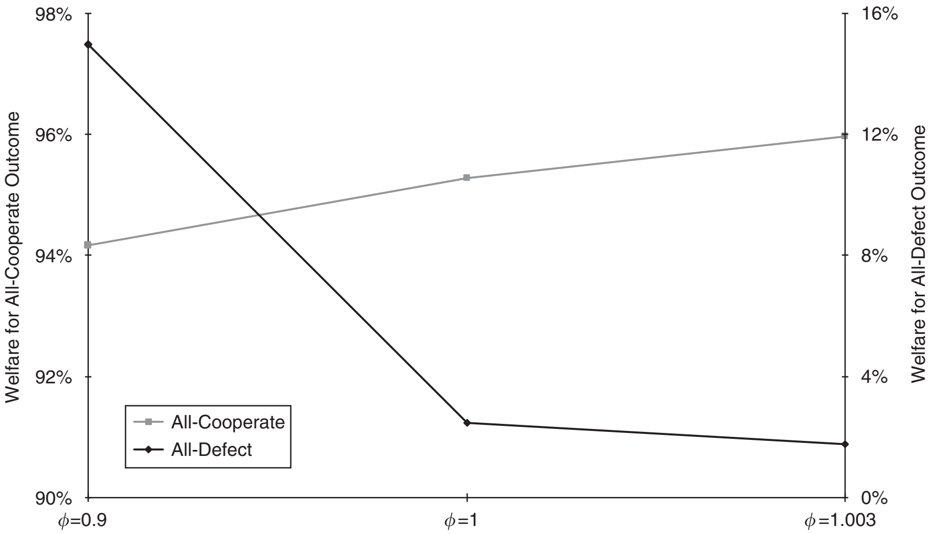

Figure 10. Welfare *vs* ϕ, $C = 5.01$.

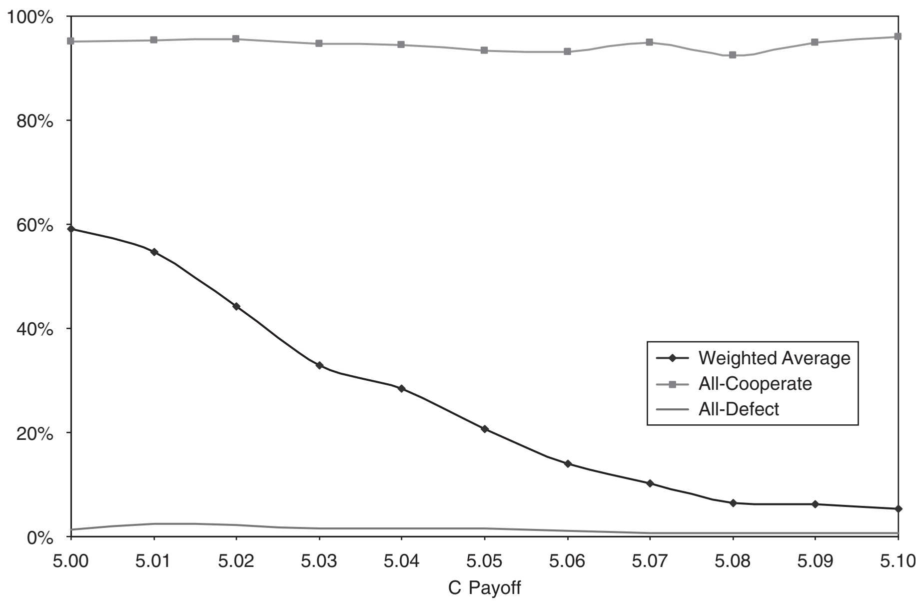

Figure 11. Welfare *vs* C payoff, for all-cooperating and all-defecting outcomes, $\phi = 1$.

To give a numerical illustration, take the case where $\phi = 1$, this gives a denominator for equation (9) of $5-3 = 2$. Furthermore, let's assume that in equilibrium, 90 percent of the agents are in neighborhoods with at least 50 percent like-type agents (which gives payoff of $-B = -3$) and 10 percent of the agents are in less-than-majority neighborhoods, and are getting $-B = -4$, on average. Thus in the all-cooperating outcome, the numerator of equation (9) is $5-0.9(3)-0.1(4) = 2.9$. This gives total welfare of $2.9/3 = 0.95$. Now let's say, the game ends in the all-defect outcome this gives a total welfare of $(3-2.9)/2 = 0.05$. In short, the outcome of the PD game is driving the total welfare.

CONCLUSION

This paper has introduced non-typed based social interactions into the Schelling model of residential segregation. In general, we find that social interactions can help to dampen the forces within the Schelling model that lead to segregation if the interactions lead to a cooperative outcome. The ability to produce cooperative social outcomes can help mitigate individual preferences for like-typed neighborhoods. On the other hand, if social interactions lead to non-cooperative outcomes, then levels of segregation can be increased beyond those found in the traditional Schelling model.

We have also demonstrated that integration and cooperation are obtainable and stable equilibria by showing how the PD game can reduce the "Schelling Threshold," that is, the proportion of like-type neighbors an agent needs to have in order to not move, and thus increase integration. We have also shown that as we increase the temptation for players to play defect, when the game ends in an all-cooperate outcome, we have increased levels of integration.

We also show that total agent welfare is, in essence, affected by the outcome of the RPD game, with the all-cooperate outcome generating roughly 95 percent of total possible welfare, while the all-defect outcome generates roughly 5 percent of total possible welfare. If agents place less relative importance on the type-based preferences of their utility segregation will be decreased, but the effect on total welfare depends on whether the game ends in all-cooperate or all-defect.

These findings suggest that if there was a "social planner" she should attempt to influence the game in the following manner. First, by attempting to reduce the importance to agents of living with the same type, the planner would decrease segregation. This would be a "good" outcome from a social point of view since, as described in the Introduction, it can increase the external benefits to society from less segregation and it can have a positive or only slightly negative effect on agent welfare. Second, the planner should encourage cooperation among agents, at least in the early stages of the game to generate an all-cooperate equilibrium, which reduces segregation as well as improves agent welfare.

There are many possible extensions to this work. For example, agents may have heterogeneous utility functions both between types and more generally within types. Agents may have heterogeneous numbers of neighbors or social interactions, or the structure of social interactions may be more complex than that studied here. Social interactions also may include other games or processes than the RPD. The most fruitful extension may be to have agents in the model recognize the type of their neighbors and react to that type in their social interactions. For instance, we could allow an agent's strategy to be a function of her opponent's type or other tags, as in the signaling RPD framework. We expect the dynamics of a model of this type to be complex but of great interest.

Acknowledgements

We thank Myong-Hun Chang for his helpful and insightful comments on an earlier draft. This paper was presented at one of the sessions sponsored by the NYC Computational Economics and Complexity Workshop at the 2007 Eastern Economic Association Meetings. We thank the session participants for their comments. Finally we thank three anonymous referees for their comments, which have helped to improve the paper. Any errors belong to the authors.

Notes

1. Another interpretation of Frost's line is that residents prefer privacy or clear boundary demarcations over social interaction with their neighbors. However, the issue of taste for privacy is not included in the model.
2. Clearly education levels of both people and the neighborhoods they live in can help determine their trust levels, their attitudes toward other races, and their willingness to engage in neighboring. Here we do not explore the role of income or other class-related variables.
3. However, interestingly, this finding did not hold for whites in the sample; heterogeneity appeared to increase feelings of distrust for whites. These findings are clearly complicated by the fact that interactions occur within the context of minority–majority relationships.
4. Frank [1999] also points out that social interaction can lead to consumption "arms races" among neighbors, which can negatively impact feelings of well-being.
5. The introduction to Schweitzer et al. [2002] provides a nice overview of the literature on cooperation in the PD.
6. In addition, wrapping or not wrapping edges has been shown not to be crucial to the emergence of segregation in the Schelling model [Pancs and Vriend 2007].
7. Since agents are randomly placed on the board, when we say that agents are selected "in turn," we mean that we begin with "agent 1" who could be located anywhere, then we go to "agent 2," who could be anywhere, then on to "agent 3," etc. Then we repeat this process starting with "agent 1."
8. A run going for 100,000 iterations means that agents have up to 790 opportunities to move. A run going for this many observations is quite rare. Furthermore, if no agents move in a round that does not mean they would not move in future rounds, but we end the game at that point because it is sufficiently close to being an equilibrium.
9. Schelling [1971] refers to this measure as the "Share."
10. Our measure of segregation produces qualitatively similar results as other measures commonly used in empirical studies such as the Dissimilarity Index [Duncan and Duncan 1955]. Also note that most of these commonly used measures are not amenable to our population. For instance, the Dissimilarity index is sensitive to neighborhood size. If one is working with a city population in the thousands or hundred-thousands with hundreds or thousands of neighborhoods (as is common in many large cities) the sensitivity is reduced. However, the Dissimilarity index would be sensitive to the choice of neighborhood size for the population size used here.
11. Note that agent movement can create different outcomes if agents can compare PD payoffs across locations. We do not address this type of movement here since it is not directly relevant for the comparison with the combined game section below. Furthermore, we also find that if agents move to new locations at random, there is no qualitative difference with the results presented here.
12. Our rule differs slightly from that of Nowak et al. [1994]. Their rule has an agent's probability coming from the sum of the payoffs of agent *i*'s neighbors over all of the games played by the neighbors. In contrast, we only use the neighbor's payoffs resulting from the game played with agent *i*. The primary reason is that we assume that agent *i* can directly view his neighbor's payoff from playing the PD game with agent *i* but cannot view the payoff that a given neighbor receives from playing with his additional neighbors. Thus, we are restricting agent *i*'s probability updating rule to include only items that *i* can directly observe from his own experience. As we demonstrate, our version of the rule only has two equilibria: all-cooperate or all-defect. Their rule has an additional mixed equilibria where defectors and cooperators co-exist.
13. We have checked other initial conditions such as $p_i = 1/2$ for all *i*, and other similar initial distributions, and have found the results to be qualitatively similar.
14. Note that the system running for 100,000 iterations is statistically rare, occurring roughly 5 percent or less across C payoffs.
15. There is no relationship between percent Similarity and number of periods till equilibrium in the all-defect case. Results available upon request.
16. For brevity, we do not present the results of convergence rates *vs* C; they are available upon request.
17. Note that we limit ϕ to a value less than but close to $D/B = 1.00333$; this is the largest value for ϕ for which both the all-cooperate and all-defect equilibria exist.
18. Interestingly, the relationship between ϕ and defection appears to be relatively weak. For $\phi > 1$, defection rates are slightly above those for $\phi \leqslant 1$, for all values of C. We do not present these results here.

References

Aktipis, C. Athena. 2004. Know When to Walk Away: Contingent Movement and the Evolution of Cooperation. *Journal of Theoretical Biology*, 231: 249–260.

Axelrod, Robert. 1984. *The Evolution of Cooperation*. New York: Basic Books.

Clark, William.A.V., and Valerie Ledwith. 2005. Mobility, Housing Stress and Neighborhood Contexts: Evidence from Los Angeles, California Center for Population Research Working Paper, CCPR–005-005.

Cutler, David M., and Edward L. Glaeser. 1997. Are Ghettos Good or Bad? *Quarterly Journal of Economics*, 112(3): 827–872.

Dokumaci, Emin, and William H. Sandholm. 2006. Schelling Redux: An Evolutionary Dynamic Model of Segregation, University of Wisconsin Working Paper.

Duncan, Otis, and Beverly Duncan. 1955. A Methodological Analysis of Segregation Indexes. *American Sociological Review*, 20: 210–217.

Farrell, Susan J., Tim Aubry, and Daniel Coulombe. 2004. Neighborhoods and Neighbors: Do They Contribute to Personal Well-being? *Journal of Community Psychology*, 32(1): 9–25.

Fehr, Ernst, and Simon Gächter. 2000. Fairness and Retaliation: The Economics of Reciprocity. *Journal of Economics Perspectives*, 14(3): 159–181.

Frank, Robert H. 1999. *Luxury Fever: Why Money Fails to Satisfy in an Era of Excess*. New York: Free Press.

Fudenberg, Drew, and Jean Tirole. 1991. *Game Theory*. Cambridge: The MIT Press.

Glaeser, Edward L., David I. Laibson, Jose A Scheinkman, and Christine L. Soutter. 2000. Measuring Trust. *Quarterly Journal of Economics*, 115(3): 811–846.

Hales, David. 2001. Tag-based Cooperation in Artificial Societies, Ph.D. Thesis, University of Essex.

Janssen, Marco A. 2008. Evolution of Cooperation in a One-shot Prisoner's Dilemma Based on Recognition of Trustworthy and Untrustworthy Agents. *Journal of Economic Behavior and Organization*, 65(3–4): 458–471.

Lee, Barret. A., R.S. Oropesa, and James W. Kanan. 1994. Neighborhood Context and Residential Mobility. *Demography*, 31(2): 249–270.

Marschall, Melissa J., and Dietlind Stolle. 2004. Race and the City: Neighborhood Context and the Development of Generalized Trust. *Political Behavior*, 26(2): 125–153.

Nowak, Martin A., and Karl Sigmund. 1998. Evolution of Indirect Reciprocity by Image Scoring. *Nature*, 393: 573–577.

Nowak, Martin A., and Robert M. May. 1992. Evolutionary Games and Spatial Chaos. *Nature*, 359: 826–829.

———. 1994. Spatial Games and the Maintenance of Cooperation. *Proceedings of the National Academy of Science*, 91: 4877–4881.

Pancs, Romans, and Nicolaas J. Vriend. 2007. Schelling's Spatial Proximity Model of Segregation Revisited. *Journal of Public Economics*, 91: 1–24.

Riolo, Rick L. 1997. The Effects of Tag Mediated Selection of Partners in Evolving Populations Playing the Iterated Prisoner's Dilemma, Santa Fe Institute Working Paper # 97-02-016.

Riolo, Rick L., Michael Cohen, and Robert Axelrod. 2001. Evolution of Cooperation without Reciprocity. *Nature*, 414: 441–443.

Schelling, Thomas C. 1969. Models of Segregation. *American Economic Review, Papers and Proceedings*, 59: 488–493.

———. 1971. Dynamic Models of Segregation. *Journal of Mathematical Sociology*, 1(2): 143–186.

Schweitzer, Frank, Laxmidhar Behera, and Heinz Mühlenbein. 2002. Evolution of Cooperation in a Spatial Prisoners Dilemma. *Advances in Complex Systems*, 5: 269–299.

Sethi, Rajiv, and Rohini Somanathan. 2004. Inequality and Segregation. *Journal of Political Economy*, 112(6): 1296–1321.

Zhang, Junfu. 2004. A Dynamic Model of Residential Segregation. *Journal of Mathematical Sociology*, 28(3): 147–170.