Environment and Planning B: Planning and Design 2008, volume 35, pages 794 – 809

doi:10.1068/b33148

A comparison of regimes of policies: lessons from the two-person iterated prisoner's dilemma game

Ching-Pin Chiu
Graduate Institute of Urban Planning, National Taipei University, 67, Section 3, Min Sheng East Road, Taipei, Taiwan, Republic of China; e-mail: andrew.daniel@msa.hinet.net
Shih-Kung Lai ¶
Department of Urban Planning, National Cheng Kung University, 1 University Road, Tainan, Taiwan, Republic of China; e-mail: sklai@mail.ncku.edu.tw
Received 4 January 2007; in revised form 14 May 2007; published online 14 July 2008

Abstract. On the basis of the presumption that the effects of plans for urban development are influenced highly by the decision mechanisms under which plans function, we compare deductively four interactive strategies derived from three regimes of policies, namely, fixed, emergent, and no policies, based on the two-person iterated prisoner's dilemma game. The four strategies under consideration are tit for tat (TFT), always defect (AD), always cooperate (AC), and random actions (RA). The results show that TFT is the best strategy followed by RA, AC, and AD. The implications are that policies that take into account contingencies yield higher expected payoffs than those that do not, and that emergent policies are more effective than either fixed or no policies. The model provides an analytical approach to the issue of evaluating the potential effects of the plans.

1 Introduction

The effects of plans are a fundamental, but difficult, question in the planning field. The difficulty lies in our incomplete understanding of how plans affect actions and of how these actions in turn yield desired outcomes, mainly due to the complexity of the systems, such as cities, under consideration. With few exceptions (see, for example, Johnson, 1996; Talen, 1996), little has been written in planning literature about how plans affect urban phenomena (Hopkins, 2001, pages 48 – 53). Talen (1996) and Johnson (1996), in assessing empirically the effects of plans for city parks and the 1929 Regional Plan of New York, encounter difficulties because it is not clear whether the observed behavior is caused by these plans. Alternatively, it is possible to address the question through computer simulations and analytical deduction. Lai (1998; 2003) conducted two computer simulations to assess the effects of plans, not in a simulated urban environment, but in the context of organizational choice behavior. The results of these simulations conclude that planning brings order to the complex systems in that, as in the garbage-can model (Cohen et al, 1972), planning results in decision makers and problems tending to stick to fixed-decision situations over time.

To our knowledge, except for two articles, no analytic deduction in the literature addresses directly the question of the effects of plans. Intriligator and Sheshinski (1986) use an optimization model to compare time planning and event planning in the light of uncertainty and cost, and derive from their model five theorems on planning that specify what type of planning is best under different conditions of uncertainty and cost. Knaap et al (1998) construct a game-theoretic model to describe the interaction between a local government and a developer and derive from their model a set of behavioral propositions for an empirical test. While these two attempts enhance our understanding of the effects of planning, they also leave open the question of which plans are optimal in a more general context. We argue that the issue of planning effects

¶ Corresponding author.

can be approached through answering two related questions: (1) given the appropriate information, how can we make better plans, and (2) given the completed plans, what decision mechanisms are likely to bring about the objectives set by the plans? It is the second question that is addressed here.

The two questions imply a distinction between plans and the underlying mechanisms of how plans work. On the one hand, here plans are defined conceptually as organized sets of interdependent contingent decisions in time and space and can take different forms, either formally documented or informally conceived. On the other hand, the underlying mechanisms of how plans work are not plans themselves, but are the casual effects that plans can have on our actions and thus on the urban-development process. Put differently, plans affect directly and indirectly how decisions are made under different mechanisms that occur in the urban-development process. For example, in negotiating with a developer wanting to invest in an area of land, a local government refers to the master land-use plan to decide whether to grant development permits to the developer. The process of deciding whether to grant the permits is the mechanism through which the land-use plan functions. Thus, there can be different mechanisms of how plans work, given this strict definition of plans, and we follow an observation of Hopkins (2001) that plans can work as agendas, policies, visions, designs, and strategies. An agenda is a list of things to do. A policy is an if – then rule. A vision is an imagined future for motivating action. A design is a well-thought-through outcome to aim for. A strategy is a contingent set of related decisions, a path in a decision tree.

In this paper, we focus on the policy aspects of plans. That is, instead of providing a general approach to evaluating plans, we present an analytical model that compares the effectiveness of different decision-making mechanisms through which the objectives set externally by the completed plans are likely to be achieved.

The narrow metaphorical model we use to represent the mechanism of how plans work and the associated decision situation is the well-known two-person iterated prisoner's dilemma game (see, for example, Axelrod, 1984). This dilemma game is particularly useful in framing some planning-related situations because it addresses whether two interacting parties would cooperate in order to provide collective goods. For example, a developer when deciding whether to invest in land and a local government when deciding how to regulate land use could be formulated as the two-person iterated prisoner's dilemma game, in that, because regulated development is a collective good and both the developer and the local government have incentives not to cooperate, the best outcome, that is, regulated development, would not come about without some coercive actions. Another example is the negotiation between the land owners in a community and a local government in locating a NIMBY (not-in-my-back-yard) facility, such as a landfill site. The best outcome would be for the local government to make reasonable compensation to the landowners for locating the facility near the affected community so that the landowners concur with the local government's plan. Since the landowners and the local government maximize their self welfare and social welfare, respectively, and thus have no incentives to cooperate, the outcome of collective goods is usually difficult to achieve without some coercive actions. However, it is not clear in such situations which strategies, that is the sequence of actions taken by one party in relation to the other party's reactions, the local government should adopt in order to enhance the welfare of the existing plan.

In a two-person iterated prisoner's dilemma game, each player can either 'defect' or 'cooperate' at each encounter. The combination of the two players' actions yields different payoffs. There are four such combinations, namely, (C, C), (C, D), (D, C), and (D, D), where C represents 'cooperate' and D represents 'defect', and the first and second elements in the parentheses symbolize the actions taken by the first and second players, respectively, in each iteration. The resulting payoffs are arranged so that the highest payoff is given to the player who defects while the other player cooperates and then receives the lowest payoff. In addition, both players are better off when they both cooperate than when they both defect. The structure of the payoff table forces the two rational players to defect, but they would be better off if they could both cooperate. Without a commitment, however, no player would run the risk of taking the cooperate action because the other player would always tend to take the defect action in order to gain. Thus, this situation is a dilemma. A two-person prisoner's dilemma game is said to be iterative when the two players encounter each other more than once. This two-person iterated prisoner's dilemma game has generated a large amount of literature, mainly in the field of economics, concerning how cooperation would emerge, and most of this work focuses on replicating, through experimental settings and computer simulations, the conditions under which a particular interactive strategy is superior (for example, Axelrod, 1984). Little has been deduced about which strategy would yield the best outcome. Instead of delving into the theoretical implications of the two-person iterated prisoner's dilemma game, we use it here as a metaphorical analytic tool to compare three regimes of policies, namely, fixed, emergent, and no policies, in terms of the overall expected payoffs for a particular player over time, assuming that the objective of maximizing the overall expected payoffs is given externally by some plan.

In section 2 a simplified version of the two-person iterated prisoner's dilemma game is depicted, and we define the three regimes of policies in the light of this model. In section 3 we compare deductively the effects of the three regimes of policies in terms of the overall expected payoffs over time, both for limited and unlimited numbers of iterations. Drawing on real planning situations, in section 4 some of the implications and limitations of the model are discussed.

2 The simplified prisoner's dilemma game and plans

The prisoner's dilemma game has been explored thoroughly, mainly outside planning literature (for example, Jones and Zhang, 2004; Rilling et al, 2002; Seale et al, 2006; Sheldon, 1999; Stephens et al, 2002; Taiji and Ikegami, 1999; Yi et al, 2005). Most of such work focuses on comparing interactive strategies in experimental settings, real situations, and computer simulations, and explaining how cooperation emerges from similar dilemmas (Axelrod, 1984; 1997; Rockenbach and Milinski, 2006). To the best of our knowledge, no deductive proof has been provided to determine which strategy is optimal. This is because there are an infinite number of combinations of payoffs that fit the logic of the prisoner's dilemma game. Nowak and May (1993) designed a simplified version of the two-person iterated prisoner's dilemma game that serves as a basis for our deductive comparison. In their formulation, Nowak and May (1993) reduced the payoff table of the two-person iterated prisoner's dilemma game to one that contains only one parameter as follows.

In the payoff diagram below, the values represent the payoffs received by each player when player one (rows) takes a certain action, while player two (columns) takes another action.

	C	D
C	1	0
D	b	0

For example, if player one cooperates and player two also cooperates, then player one will receive a payoff of one. If player one defects while player two cooperates, then player one will receive a payoff of b. For the simplified version of the two-person iterated prisoner's dilemma game to be equivalent to the original one, b must be greater than one, so that the Nash equilibrium settles on the combination of actions where both players defect. The beauty of this simplified version of the two-person iterated prisoner's dilemma lies in its simplicity of reducing the analytic structure to a single parameter b, while retaining the generalizability of the game situation. Note that these payoffs are designed in order to manifest the logic of the prisoner's dilemma game and should not be taken literally as shown by payoffs of zero in the second column. This simplified version of the two-person iterated prisoner's dilemma game allows us to formulate a mathematical proof in a succinct and efficient way.

Given the simplified structure of the two-person iterated prisoner's dilemma game, we define a policy as a decision rule of a sequence of actions to be taken in the ensuing dynamic interactions of the two players over time. That is, a sequence of *C*s or *D*s is derived from a policy for the encounters over time, such as *CCDDCDCC* .... In addition, we assume that the objective of each player is to maximize the overall expected payoffs as given externally by an associated plan. There can be three regimes of such policies, namely, fixed, emergent, and no policies. A fixed policy is a predetermined sequence of actions to be taken over time regardless of the action taken by the other player in each iteration. An emergent policy is a contingent set of actions taken in the light of the action taken by the other player in the previous iteration. No policies implies that there is no orderly pattern or rule for how to take actions over time. Equivalently, a policy is an if–then decision rule (Hopkins, 2001). For each regime of policies, there can be an infinite number of sequences of actions that satisfy the definitions. For example, any combinatory sequence of *C*s and *D*s is a fixed plan; any contingent rule for taking actions can be an emergent policy; and a regime of no policies implies any probability distribution of taking certain actions.

In order to compare the effectiveness of the three regimes of policies, we select a representative policy for each regime, namely, the four strategies always defect (AD), always cooperate (AC), tit for tat (TFT), and random actions (RA). AD means to cooperate in the initial iteration, and once the other player defects in the current iteration, always defect in the subsequent iterations. AC means to cooperate in all encounters, regardless of what actions the other player takes. TFT means to cooperate in the initial iteration, and then respond by following the other player's action in the previous iteration. RA simply means to take any action arbitrarily in each iteration with a certain probability distribution. Intuitively, TFT and AD represent the best strategies in the emergent and fixed regimes, respectively, because TFT is found to be the best through computer simulations (Axelrod, 1984) and AD avoids the possibility of being exploited. Whether RA is the best in its regime depends on the probability distribution selected, but we assign the same probability distribution of taking actions across the policies in the different regimes for the purpose of comparison. Therefore, the four policy types are sufficiently distinctive to characterize the three regimes of policy, and comparing them deductively should yield insight into which regime of policies would be more effective. Table 1 summarizes the definitions of the four policy types.

3 The deductive comparisons

In order to compare the effects of the four policies, we formulate first, in general terms, the overall expected payoffs for player one based on the simplified game situation, and then make pairwise comparisons of these four policies in terms of the overall expected payoffs to determine the rankings of these policies. There are two situations—a limited number of iterations and an unlimited number of iterations. For each computation of the overall expected payoff for the particular player, we assume that in each iteration

Table 1. Definitions of the four policies tested.

Strategy	Definition of strategy	Regime of policy
AD (always defect)	Cooperate in the initial iteration, and once the other player defects in the current iteration, always defect in the subsequent iterations.	Fixed
AC (always cooperate)	Cooperate in all encounters, regardless of what actions the other player takes.	Fixed
TFT (tit for tat)	Cooperate in the initial iteration, and then respond by following the other players' action in the previous iteration.	Emergent
RA (random actions)	Take any action arbitrarily in each iteration with a certain probability distribution.	No policies

the probabilities that player one cooperates and defects are p and $1-p$, respectively, and that player two cooperates and defects are q and $1-q$, respectively, where $0 \leqslant p, q \leqslant 1$. Let k denote the number of iterations. Let $(r, s)_k$ stand for a combination of the actions taken by player one (action r) and player two (action s) in iteration k respectively, where $r, s \in \{C, D\}$ and C and D denote cooperate and defect, respectively. A strategy is a sequence of actions taken by a particular player over time, which is distinct from a policy of if–then decision rules for that player.

3.1 The case of a limited number of iterations

3.1.1 *Evaluation of TFT*

For TFT, when $k = 1$, indicating the initial iteration, player one cooperates with the probability of one, whereas player two can either cooperate or defect with probabilities of q and $1-q$, respectively. The possible combinations of actions in this iteration are $(C, C)_1$ and $(C, D)_1$. Referring to figure 1, where the paths with bold lines are permissible under the definition of TFT, the expected payoff for player one at this stage is

$$(1 \times q \times 1) + [1 \times (1-q) \times 0] = q \ . \tag{1}$$

Note that the resulting payoffs are given at the end of leaves in each iteration.

When $k = 2$, given the definition of TFT and if player two cooperates in iteration one, then player one will cooperate in iteration two with a probability of one; otherwise, player one will defect with a probability of one. Referring to figure 1, there are four possible combinations of sequential actions in iteration two, that is, $[(C, C)_1, (C, C)_2]$, $[(C, C)_1, (C, D)_2]$, $[(C, D)_1, (D, C)_2]$, and $[(C, D)_1, (D, D)_2]$, and the overall expected payoff for player one in iteration two is

$$\begin{aligned}(1 \times q \times 1 \times 1 \times q \times 1) + [1 \times q \times 1 \times (1-q) \times 0] + [1 \times (1-q) \times 1 \times q \times b] \\ + [1 \times (1-q) \times 1 \times (1-q) \times 0] = q^2 + qb - q^2 b \ .\end{aligned} \tag{2}$$

In the same way, when $k = 3$ there are eight possible combinations of sequential actions in iteration three, and the overall expected payoff for player one in this iteration is

$$\begin{aligned}&(1 \times q \times 1 \times q \times 1 \times q \times 1) + [1 \times q \times 1 \times q \times 1 \times (1-q) \times 0] \\ &+ [1 \times q \times 1 \times (1-q) \times 1 \times q \times b] + [1 \times q \times 1 \times (1-q) \times 1 \times (1-q) \times 0] \\ &+ [1 \times (1-q) \times 1 \times q \times 1 \times q \times 1] + [1 \times (1-q) \times 1 \times q \times 1 \times (1-q) \times 0] \\ &+ [1 \times (1-q) \times 1 \times (1-q) \times 1 \times q \times b] \\ &+ [1 \times (1-q) \times 1 \times (1-q) \times 1 \times (1-q) \times 0] = q^2 + qb - q^2 b \ .\end{aligned} \tag{3}$$

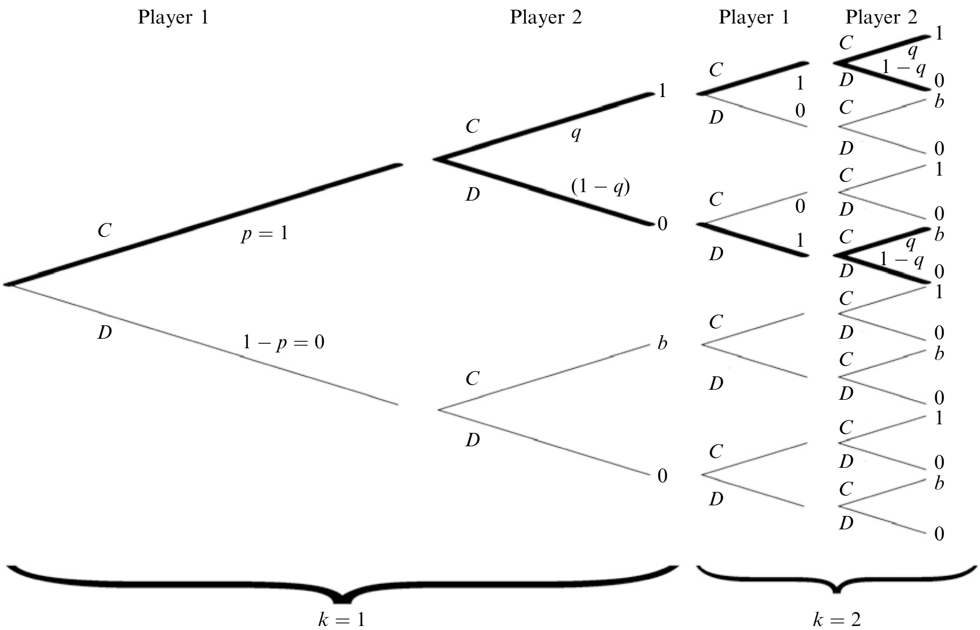

Figure 1. The game tree of tit for tat in the two-person iterated prisoner's dilemma game for iterations $k = 1$ and 2. C = cooperates; D = defects; p is the probability player one will cooperate and q the probability that player two will cooperate; b is a payoff value > 1.

It can be proved, as shown in the appendix, that when $k = n$ the expected payoff for player one in iteration n is $T_n = qT_{n-1} + (1-q)T_{n-1} = T_{n-1}$, and we have

$$T_n = \begin{cases} q\,, & n = 1 \\ q^2 + qb - q^2 b\,, & n \geqslant 2 \end{cases} .$$

Let S_t be the sum of the expected payoffs for player one across the iterations up to n, and we have

$$S_t = q + (q^2 + qb - q^2 b)(n-1) \ . \tag{4}$$

3.1.2 *Evaluation of AD*

For AD, player one cooperates initially with the probability of one, and will keep defecting once player two defects. When $k = 1$ player one cooperates initially, and there are two possible combinations of actions $(C, C)_1$ and $(C, D)_1$. The overall expected payoff for player one in this iteration is

$$(1 \times q \times 1) + [1 \times (1-q) \times 0] = q \ . \tag{5}$$

If player two defects in iteration one, then player one will defect in return in iteration two with a probability of one. Starting from iteration two, player two, noticing that player one defected in iteration one and to avoid being exploited, will also defect in the following iterations, and the process enters into a punishment phase where both players defect (Dixit and Skeath, 2002). According to this scenario, when $k = 2$ the possible combinations of sequential actions are $[(C, C)_1, (C, C)_2]$, $[(C, C)_1, (C, D)_2$, and $[(C, D)_1, (D, D)_2]$, and the overall expected payoff for player one in iteration two is

$$(1 \times q \times 1 \times q \times 1) + [1 \times q \times 1 \times (1-q) \times 0] + [1 \times (1-q) \times 1 \times 1 \times 0] = q^2 . \tag{6}$$

When $k=3$ there are four possible combinations of sequential actions, and the expected payoff for player one in iteration three is

$$(1\times q\times 1\times q\times 1\times q\times 1)+[1\times q\times 1\times q\times 1\times(1-q)\times 0] + [1\times q\times 1\times(1-q)\times 1\times 1\times 0]+[1\times(1-q)\times 1\times 1\times 1\times 1\times 0] = q^3 . \quad (7)$$

Let P_n denote the expected payoff in iteration n. It can be proved, as shown in the appendix, that $P_n = qP_{n-1} + [(1-q)\times 0] = q^n$. Let S_p denote the sum of the expected payoffs across iterations up to n, and we have

$$S_p = \frac{q(1-q^n)}{1-q} . \quad (8)$$

3.1.3 *Evaluation of AC*

Under the AC policy, player one always cooperates with a probability of one, regardless of the action taken by player two in the previous iteration. When $k=1$ there are two possible combinations of actions, $(C, C)_1$ and $(C, D)_1$, and the expected payoff for player one in this iteration is

$$(1\times q\times 1)+[1\times(1-q)\times 0] = q . \quad (9)$$

When $k=2$ there are four possible combinations of sequential actions, namely, $[(C, C)_1, (C, C)_2]$, $[(C, C)_1, (C, D)_2]$, $[(C, D)_1, (C, C)_2]$, and $[(C, D)_1, (C, D)_2]$. The expected payoff for player one in iteration two is

$$(1\times q\times 1\times q\times 1)+[1\times q\times 1\times(1-q)\times 0]+[1\times(1-q)\times 1\times q\times 1] + [1\times(1-q)\times 1\times(1-q)\times 0] = q . \quad (10)$$

The expected payoff for player one in iteration three is

$$(1\times q\times 1\times q\times 1\times q\times 1)+[1\times q\times 1\times q\times 1\times(1-q)\times 0] + [1\times q\times 1\times(1-q)\times 1\times q\times 1]+[1\times q\times 1\times(1-q)\times 1\times(1-q)\times 0] + [1\times(1-q)\times 1\times q\times 1\times q\times 1]+[1\times(1-q)\times 1\times q\times 1\times(1-q)\times 0] + [1\times(1-q)\times 1\times(1-q)\times 1\times q\times 1] + [1\times(1-q)\times 1\times(1-q)\times 1\times(1-q)\times 0] = q . \quad (11)$$

Let F_n denote the expected payoff for player one in iteration n, and it can be proved, as shown in the appendix, that $F_n = qF_{n-1} + (1-q)F_{n-1} = F_{n-1} = q$. Let S_f denote the sum of the expected payoffs for player one across iterations up to n, and we have

$$S_f = nq . \quad (12)$$

3.1.4 *Evaluation of RA*

Under the RA policy, both players either defect or cooperate with probabilities p or $1-p$ and q or $1-q$, respectively. When $k=1$ there are four possible combinations of actions, namely, $(C, C)_1$, $(C, D)_1$, $(D, C)_1$, and $(D, D)_1$. The expected payoff for player one in this iteration is

$$(p\times q\times 1)+[p\times(1-q)\times 0]+[(1-p)\times q\times b]+[(1-p)\times(1-q)\times 0] = pq+qb-pqb . \quad (13)$$

When $k=2$ there are 16 possible combinations of sequential actions. These are $[(C, C)_1, (C, C)_2]$, $[(C, C)_1, (C, D)_2]$, $[(C, C)_1, (D, C)_2]$, $[(C, C)_1, (D, D)_2]$, $[(C, D)_1, (C, C)_2]$, $[(C, D)_1, (C, D)_2]$, $[(C, D)_1, (D, C)_2]$, $[(C, D)_1, (D, D)_2]$, $[(D, C)_1, (C, C)_2]$,

$[(D, C)_1, (C, D)_2]$, $[(D, C)_1, (D, C)_2]$, $[(D, C)_1, (D, D)_2]$, $[(D, D)_1, (C, C)_2]$, $[(D, D)_1, (C, D)_2]$, $[(D, D)_1, (D, C)_2]$, and $[(D, D)_1, (D, D)_2]$. The expected payoff for player one in this iteration is

$$\begin{aligned}(p \times q \times p \times q \times 1) &+ [p \times q \times p \times (1-q) \times 0] + [p \times q \times (1-p) \times q \times b] \\ &+ [p \times q \times (1-p) \times (1-q) \times 0] + [p \times (1-q) \times p \times q \times 1] \\ &+ [p \times (1-q) \times p \times (1-q) \times 0] + [p \times (1-q) \times (1-p) \times q \times b] \\ &+ [p \times (1-q) \times (1-p) \times (1-q) \times 0] + [(1-p) \times q \times p \times q \times 1] \\ &+ [(1-p) \times q \times p \times (1-q)] \times 0] + [(1-p) \times q \times (1-p) \times q \times b] \\ &+ [(1-p) \times q \times (1-p) \times (1-q) \times 0] + [(1-p) \times (1-q) \times p \times q \times 1] \\ &+ [(1-p) \times (1-q) \times p \times (1-q) \times 0] + [(1-p) \times (1-q) \times (1-p) \times q \times b] \\ &+ [(1-p) \times (1-q) \times (1-p) \times (1-q) \times 0] \;=\; pq + qb - pqb \ . \end{aligned} \tag{14}$$

When $k = 3$ the expected payoff for player one in this iteration is also

$$(pq + qb - pqb) \ . \tag{15}$$

Let M_n denote the expected payoff for player one in iteration n, and it can be proved, as shown in the appendix, that $M_n = M_{n-1} = M_{n-2} = \ldots = M_2 = M_1 = pq + qb - pqb$. Let S_m denote the sum of the expected payoffs across iterations up to n, and we have

$$S_m = n(pq + qb - pqb) \ . \tag{16}$$

Table 2 summarizes the evaluation results of the four strategies.

Table 2. Evaluation results of the four strategies.

Strategy	Sum of the expected payoffs across iterations up to n	Equation
TFT (tit for tat)	$S_t = q + (q^2 + qb - q^2 b)(n-1)$	(4)
AD (always defect)	$S_p = \dfrac{q(1-q^n)}{1-q}$	(8)
AC (always cooperate)	$S_f = nq$	(12)
RA (random actions)	$S_m = n(pq + qb - pqb)$	(16)

Note. p and q are the probabilities that player one and player two, respectively, cooperate, and b is a payoff value >1.

3.1.5 *Pairwise comparison between TFT and AD*

Given these generalized formulae for the overall expected payoffs for player one for each of the four policies, we can compare these payoffs in pairs to determine the rankings of the four policies in terms of the overall expected payoffs. The pairwise comparisons are conducted through the logic of induction. Consider first the TFT and AD strategies. When $k = 2$ the difference between the overall expected payoffs for player one is given by equations (2) – (6), that is,

$$(q^2 + qb - q^2 b) - q^2 = qb(1-q) \ , \tag{17}$$

which is positive because $1 - q > 0$, and it can be concluded that TFT is better than AD in terms of the overall expected payoff for iteration two.

Assume that the payoff difference between TFT and AD derived from equations (4) and (8) is positive when $k = n$, that is,

$$S_t - S_p \;=\; q + (q^2 + qb - q^2 b)(n-1) - \frac{q(1-q^n)}{1-q} \;>\; 0 \;. \tag{18}$$

When $k = n+1$ this difference between the overall expected payoffs for player one derived from TAT and AD is equal to

$$q + (q^2 + qb - q^2 b)(n-1) + (q^2 + qb - q^2 b) - \frac{q(1-q^n)}{1-q} - q^{n+1}$$
$$= \left[q + (q^2 + qb - q^2 b)(n-1) - \frac{q(1-q^n)}{1-q}\right] + (q^2 + qb - q^2 b) - q^{n+1} \;.$$

This can be rewritten as

$$\left[q + (q^2 + qb - q^2 b)(n-1) - \frac{q(1-q^n)}{1-q}\right] + q\left[q(1-q^{n-1}) + qb(1-q)\right] \;,$$

and because

$$\left[1 + (q^2 + qb - q^2 b)(n-1) - \frac{q(1-q^n)}{1-q}\right] > 0,\; (1-q^{n-1}) > 0, \;\text{ and }\; (1-q) > 0 \;,$$

we can see that it is positive. Thus we have

$$\left[q + (q^2 + qb - q^2 b)(n-1) - \frac{q(1-q^n)}{1-q}\right] + q\left[q(1-q^{n-1}) + qb(1-q)\right] \;>\; 0 \;. \tag{19}$$

Based on the logic of induction, we have proved that TFT is better than AD in terms of the overall expected payoff for player one.

3.1.6 *Pairwise comparison between TFT and AC*

For the comparison between TFT and AC, when $k = 2$ the difference between the overall expected payoffs derived from the two policies, given equations (4) – (12) for $n = 2$, is

$$q + (q^2 + qb - q^2 b) - q - q \;=\; q \times (q-1)(1-b) \;. \tag{20}$$

By assumption $(1-b) < 0$, and $(q-1) < 0$, therefore the payoff difference expressed in equation (20) is positive, meaning that TFT is better than AC in terms of the overall expected payoff for player one when $k = 2$.

Assume that the payoff difference between TFT and AC derived from equations (4) and (12) is positive when $k = n$, that is,

$$q + (q^2 + qb - q^2 b)(n-1) - nq \;>\; 0 \;. \tag{21}$$

When $k = n+1$ the difference in payoffs for player one is

$$q + (q^2 + qb - q^2 b)(n) - (n+1)q$$
$$= q + (q^2 + qb - q^2 b)(n-1) + (q^2 + qb - q^2 b) - nq - q \;.$$

Rewriting this, and because we know from equation (21) that

$$[q + (q^2 + qb - q^2 b)(n-1) - nq] > 0$$

and from equation (20) $[(q^2 + qb - q^2 b) - q] = (b-1)(q - q^2) > 0$, we have

$$[q + (q^2 + qb - q^2 b)(n-1) - nq] + [(q^2 + qb - q^2 b) - q] \;>\; 0 \;. \tag{22}$$

Based on the logic of induction we can conclude that TFT is better than AC in terms of the overall expected payoff for player one.

3.1.7 *Pairwise comparison between TFT and RA*
For the comparison between TFT and RA, when $k = 2$ the payoff difference between equations (4) and (16) for $n = 2$ is

$$q + (q^2 + qb - q^2b) - 2(pq + qb - pqb) \;=\; q \times (1 - b) \times (1 + q - 2p) \; . \tag{23}$$

Assume that $(1 + q) < 2p$ and because $(1 - b) < 0$, we have $q \times (1-b) \times (1 + q - 2p) > 0$, meaning that TFT is better than RA in terms of the overall expected payoff for player one when $k = 2$.

Let the payoff difference be positive when $k = n$, that is,

$$S_t - S_m \;=\; q + (q^2 + qb - q^2b)(n - 1) - n(pq + qb - pqb) \;>\; 0 \; . \tag{24}$$

When $k = n + 1$ we have

$$\begin{aligned} S_t - S_m \;&=\; q + (q^2 + qb - q^2b)(n) - (n + 1)(pq + qb - pqb) \\ &=\; [q + (q^2 + qb - q^2b)(n - 1) - n(pq + qb - pqb)] \\ &\quad + (q^2 + qb - q^2b) - (pq + qb - pqb) \; . \end{aligned}$$

This payoff difference can be rewritten as follows and, because equation (24) is positive, $(1 + q) < 2p$ and thus $p > (1 + q)/2 > q$ so that $(q - p) < 0$ and by definition $(1 - b) < 0$, we have

$$[q + (q^2 + qb - q^2b)(n - 1) - n(pq + qb - pqb)] + q(1 - b)(q - p) \;>\; 0 \; . \tag{25}$$

This leads to the conclusion, through the logic of induction, that TFT is better than RA in terms of the overall expected payoff for player one.

3.1.8 *Pairwise comparisons between RA and AC*
For the comparison between RA and AC, when $k = 2$ the payoff difference under the two policies can be expressed as $S_m - S_f$ for $n = 2$, that is,

$$2(pq + qb - pqb) - 2q \;=\; 2q(p - 1)(1 - b) \; , \tag{26}$$

which is positive because, by definition, $(1 - b) < 0$ and $(p - 1) < 0$. This means that RA is better than AC in terms of the overall expected payoff for player one when $k = 2$.

Let the payoff difference be positive when $k = n$, that is,

$$S_m - S_f \;=\; n(pq + qb - pqb) - nq \;>\; 0 \; . \tag{27}$$

When $k = n + 1$ the payoff difference becomes

$$(n + 1)(pq + qb - pqb) - (n + 1)q \;=\; (n + 1)q(p - 1)(1 - b) \;>\; 0 \; . \tag{28}$$

We can conclude, based on the logic of induction, that RA is better than AC in terms of the overall expected payoff for player one.

3.1.9 *Pairwise comparison between AC and AD*
For the comparison between AC and AD, when $k = 2$ the payoff difference under the two policies is

$$(q + q) - (q + q^2) \;=\; q - q^2 \;>\; 0 \; . \tag{29}$$

Therefore, AC is better than AD in terms of the overall expected payoff for player one when $k = 2$. Now, let the inequality obtain when $k = n$ that is

$$S_f - S_p = nq - \frac{q(1-q^n)}{1-q} > 0 \; . \tag{30}$$

When $k = n+1$ the payoff difference between AC and AD changes to

$$(n+1)q - \frac{q(1-q^n)}{1-q} - q^{n+1} \; .$$

As before, we rewrite this and because $[nq - q(1-q^n)/(1-q)] > 0$ and $(q - q^{n+1}) > 0$, we have

$$\left(nq - \frac{q(1-q^n)}{1-q}\right) + (q - q^{n+1}) > 0 \; . \tag{31}$$

Based on the logic of induction, we can conclude that AC is better than AD in terms of the overall expected payoff for player one.

We have proved through the pairwise comparisons that the ranking of the four policies in terms of the overall expected payoff for player one in descending order is TFT, RA, AC, and AD. Table 3 summarizes the evaluation results of the five pairwise comparisons in the case of a limited number of iterations.

Table 3. Evaluation results of the pairwise comparisons of payoffs for player one with a limited number of iterations.

Strategies compared	Payoff difference for player one in iteration $n+1$	Inequality
TFT versus AD	$\left[q + (q^2 + qb - q^2b)(n-1) - \frac{q(1-q^n)}{1-q}\right] + q[q(1-q^{n-1}) + qb(1-q)] > 0$	(19)
TFT versus AC	$[q + (q^2 + qb - q^2b)(n-1) - nq] + [(q^2 + qb - q^2b) - q] > 0$	(22)
TFT versus RA	$[q + (q^2 + qb - q^2b)(n-1) - n(pq + qb - pqb) + q(1-b)(q-p) > 0$	(25)
RA versus AC	$(n+1)(pq + qb - pqb) - (n+1)q > 0$	(28)
AC versus AD	$\left[(n+1)q - \frac{q(1-q^n)}{1-q} - q^{n+1}\right] = \left[nq - \frac{q(1-q^n)}{1-q}\right] + (q - q^{n+1}) > 0$	(31)

Note. TFT = tit for tat; AD = always defect; AC = always cooperate; RA = random actions. p and q are the probabilities that player one and player two, respectively, cooperate, and b is a payoff value >1.

3.2 The case of an unlimited number of iterations

In this section, the case of an unlimited number of iterations is considered for pairwise comparisons of the four policies.

3.2.1 *Pairwise comparison between TFT and AD*

For the comparison between TFT and AD, when $k = n$ we know that the payoff difference under the two policies is given in equation (18). When $n \to \infty$, implying an unlimited number of iterations, the payoff difference between TFT and AD is expressed as

$$\lim_{n \to \infty} \left[q + (q^2 + qb - q^2b)(n-1) - \frac{q(1-q^n)}{1-q}\right] ,$$

where

$$\lim_{n\to\infty}\left[q+(q^2+qb-q^2b)(n-1)\right] = \lim_{n\to\infty}\left\{q+\left[q^2+b(q-q^2)\right](n-1)\right\} . \tag{32}$$

Because $0 < q < 1$, we have $0 < q^2 < 1$ and $(q - q^2) > 0$. Thus

$$\lim_{n\to\infty}\left\{q+\left[q^2+b(q-q^2)\right](n-1)\right\} \rightarrow \infty .$$

In addition, because

$$\lim_{n\to\infty}\frac{q(1-q^n)}{1-q} = \frac{q}{1-q}$$

is a constant, we have

$$\lim_{n\to\infty}\left[q+(q^2+qb-q^2b)(n-1)-\frac{q(1-q^n)}{1-q}\right] = \infty-\frac{q}{1-q} > 0 , \tag{33}$$

which implies that TFT is better than AD in terms of the overall expected payoff for player one when the number of iterations is unlimited.

3.2.2 *Pairwise comparison between TFT and AC*

For the comparison between TFT and AC, we know that the payoff difference under the two policies is given in equation (21). When $n \rightarrow \infty$, the limit of this payoff difference is expressed as

$$\lim_{n\to\infty}\left[q+(q^2+qb-q^2b)(n-1)-nq\right] .$$

Because $n \rightarrow \infty$, this can be rewritten as

$$\lim_{n\to\infty}\left[(q^2+qb-q^2b)n-nq\right] = \lim_{n\to\infty}\left[n(q^2+qb-q^2b-q)\right] = \lim_{n\to\infty} n(b-1)(q-q^2) ,$$

By definition, $(b - 1) > 0$ and $(q - q^2) > 0$ and as a result we have

$$\lim_{n\to\infty}\left[q+(q^2+qb-q^2b)(n-1)-nq\right] > 0 . \tag{34}$$

We conclude that TFT is better than AC in terms of the overall expected payoff for player one when the number of iterations is unlimited.

3.2.3 *Pairwise comparison between TFT and RA*

For the comparison between TFT and RA, the payoff difference for $k = n$ is given in equation (24). When $n \rightarrow \infty$, this payoff difference becomes

$$\begin{aligned}
&\lim_{n\to\infty}\left[q+(q^2+qb-q^2b)(n-1)-n(pq+qb-pqb)\right] \\
&\quad = \lim_{n\to\infty}\left[q-(q^2+qb-q^2b)+n(q^2+qb+q^2b-pq-qb+pqb)\right] \\
&\quad = \lim_{n\to\infty}\left[q-(q^2+qb-q^2b)+n(b-1)(pq-q^2)\right] .
\end{aligned}$$

$[q - (q^2 + qb - q^2b)]$ is a constant, and $[n(b - 1)(pq - q^2)] > 0$ because $b > 1$ and $1 + q < 2p$, so that $p > (1 + q)/2$ and thus $pq > (q + q^2)/2 > q^2$.

$$\lim_{n\to\infty}\left[q+(q^2+qb-q^2b)(n-1)-n(pq+qb-pqb)\right. > 0 . \tag{35}$$

Hence, from above we can conclude that TFT is better than RA in terms of the overall expected payoff for player one when the number of iterations is unlimited.

3.2.4 *Pairwise comparison between RA and AC*

For the comparison between RA and AC, the payoff difference for $k = n$ is given in equation (27). When $n \rightarrow \infty$, the payoff difference can be expressed as

$$\lim_{n\to\infty}\left[n(pq+qb-pqb)-(nq)\right] = \lim_{n\to\infty}\left[nq(p-1)(1-b)\right] .$$

This tends to ∞ because $0 < p < 1$, $b > 1$, and both are constants. Thus

$$\lim_{n\to\infty} [n(pq + qb - pqb) - (nq)] > 0 \qquad (36)$$

and we can conclude that RA is better than AC in terms of the overall expected payoff for player one when the number of iterations is unlimited.

3.2.5 *Pairwise comparison between AC and AD*

For the comparison between AC and AD, the payoff difference when $k = n$ is given in equation (30). When $n \to \infty$, that payoff difference can be expressed as

$$\lim_{n\to\infty} \left[nq - \frac{q(1-q^n)}{1-q} \right] \to \infty \ ,$$

because as $n \to \infty$ $nq \to \infty$, whereas $q(1-q^n)/(1-q)$ is a constant. Thus

$$\lim_{n\to\infty} \left[nq - \frac{q(1-q^n)}{1-q} \right] > 0 \ , \qquad (37)$$

and we can conclude that AC is better than AD in terms of the overall expected payoff for player one when the number of iterations is unlimited.

In short, in the case of an unlimited number of iterations, the ranking of the four policies in descending order in terms of the overall expected payoff is TFT, RA, AC, and AD, which is consistent with the ranking in the case of a limited number of iterations. Table 4 summarizes the evaluation results of the five pairwise comparisons in the case of an unlimited number of iterations.

Table 4. Evaluation results of the pairwise comparisons of payoffs for player one with an unlimited number of iterations.

Strategies compared	Payoff difference for player one in iteration $n \to \infty$	Inequality
TFT versus AD	$\lim_{n\to\infty} \left[q + (q^2 + qb - q^2 b)(n-1) - \frac{q(1-q^n)}{1-q} \right] > 0$	(33)
TFT versus AC	$\lim_{n\to\infty} [q + (q^2 + qb - q^2 b)(n-1) - nq] > 0$	(34)
TFT versus RA	$\lim_{n\to\infty} [q + (q^2 + qb - q^2 b)(n-1) - n(pq + qb - pqb)] > 0$	(35)
RA versus AC	$\lim_{n\to\infty} [n(pq + qb - pqb) - (nq)] > 0$	(36)
AC versus AD	$\lim_{n\to\infty} \left[nq - \frac{q(1-q^n)}{1-q} \right] > 0$	(37)

Note. TFT = tit for tat; AD = always defect; AC = always cooperate; RA = random actions. p and q are the probabilities that player one and player two, respectively, cooperate, and b is a payoff value >1.

4 Discussion

We have shown that TFT is better than RA only under the condition that $(1+q) < 2p$ or $p > (1+q)/2 > q$. This condition sets a behavioral constraint on p and because $0 \leqslant q \leqslant 1$, $p > (1+q)/2 \geqslant 0.5$, implying that player one is more inclined to cooperate than to defect and that player one's inclination to cooperate is greater than player two's. Cooperation seems to be a useful action, when taken appropriately, even when faced with an opponent who is less cooperative.

Intuitively, AD seems better than AC because AD avoids being exploited by the opponent and AC runs the risk of being exploited. The comparisons show, however, that AC is better than AD. This conclusion is derived, on the one hand, from the behavioral assumption that when player one adopts the AD policy, player two will defend himself or herself by taking the defect action, reducing player one's overall expected payoff. On the other hand, if player one adopts the AC policy, player two is free to cooperate or defect with probabilities of q and $1 - q$, respectively, increasing player one's overall expected payoff compared with the AD policy. If we assume that player two will always defect when player one adopts the AC policy, then AD might be better than AC under this behavioral assumption. In addition, the overall expected payoff is computed from the point of view of a particular player. It is easy to compare these policies in terms of the sum of the overall expected payoffs of both players. This way, it would be positive to assess which policies are more effective in enhancing social welfare. Nevertheless, the comparative approach depicted in this paper provides a useful way to analyze deductively the effectiveness of different policies in the two-person iterated prisoner's dilemma game.

On the basis of the distinction between the definition of plans and the underlying mechanisms of how plans work, this paper does not address directly the issue of measuring the effectiveness of plans and how we should make better plans. Instead, it shows that given an external planning objective of maximizing the overall expected payoff of a player, when faced with the situation reminiscent of the two-person iterated prisoner's dilemma game, the best strategy for the player to adopt is TFT. The player may well be a local government in negotiation with land owners on the location of a NIMBY facility or with a developer on the granting of development permits, and the payoff for the local government may be either social welfare or votes, depending on how the planning situation is perceived by the local government. The implications of the mathematical analysis are, however, that measuring completely the effectiveness of plans is an almost inextricable issue both analytically and empirically, and that in order to do so, we need to take into account the evaluation of both making plans and how such plans work under different mechanisms. The model and the associated findings provide a starting point to address this issue analytically.

5 Conclusions

We have proved mathematically that TFT is the best of the four commonly used strategies, the others being AC, AD, and RA, in the two-person iterated prisoner's dilemma game. Viewing these strategies defined in game theory as representative policies in the three regimes of fixed, emergent, and no policies, the implication is that emergent policies that take into account contingencies are the most effective in terms of the overall expected payoffs. The model shows a starting approach to the issue of evaluating the potential effects of plans.

Acknowledgements. The authors thank Professor Lewis D Hopkins from the Department of Urban and Regional Planning at the University of Illinois, Urbana-Champaign for his helpful comments on clarifying the ideas of plans and policies.

References

Axelrod R, 1984 *The Evolution of Cooperation* (Basic Books, New York)
Axelrod R, 1997 *The Complexity of Cooperation: Agent-based Models of Competition and Collaboration* (Princeton University Press, Princeton, NJ)
Cohen M D, March J G, Olsen J P, 1972, "A garbage can model of organizational choice" *Administrative Science Quarterly* **17** 1 – 25
Dixit A K, Skeath S, 2002 *Game of Strategy* (W W Norton, New York)
Hopkins L D, 2001 *Urban Development: The Logic of Making Plans* (Island Press, Washington, DC)
Intriligator M D, Sheshinski E, 1986, "Toward a theory of planning", in *Social Choice and Public Decision Making* Eds W Heller, R Starr, D Starrett (Cambridge University Press, Cambridge) pp 119 – 170

Johnson D A, 1996 *Planning the Great Metropolis: The 1929 Regional Plan of New York and its Environs* (Routledge, London)
Jones M, Zhang J, 2004, "Rationality and bounded information in repeated games, with application to the iterated prisoner's dilemma" *Journal of Mathematical Psychology* **48** 334–354
Knaap G J, Hopkins L D, Donaghy K P, 1998, "Do plans matter? A framework for examining the logic and effects of land use planning" *Journal of Planning Education and Research* **18** 25–34
Lai S-K, 1998, "From organized anarchy to controlled structure: effects of planning on the garbage-can decision processes" *Environment and Planning B: Planning and Design* **25** 85–102
Lai S-K, 2003, "Effects of planning on the garbage-can decision processes: a reformulation and extension" *Environment and Planning B: Planning and Design* **30** 379–389
Nowak M A, May R M, 1993, "The spatial dilemmas of evolution" *Journal of Bifurcation and Chaos* **3** 35–78
Rilling J K, Gutman D A, Zeh T R, Pagnoni G, Bern G S, Kilts C D, 2002, "A neural basis for social cooperation" *Neuron* **35** 395–405
Rockenbach B, Milinski M, 2006, "The efficient interaction of indirect reciprocity and costly punishment" *Nature* **444** 718–723
Seale D A, Arend R J, Phelan S, 2006, "Modeling alliance activity: opportunity cost effects and manipulations in an iterated prisoner's dilemma with exit options" *Organizational Behavior and Human Decision Processes* **100** 60–75
Sheldon K M, 1999, "Learning the lessons of tit-for-tat: even competitor can get the message' *Journal of Personality and Social Psychology* **77** 1245–1253
Stephens D W, McLinn C M, Stevens J R, 2002, "Discounting and reciprocity in an iterated prisoner's dilemma" *Science* **298** 2216–2218
Taiji M, Ikegami T, 1999, "Dynamics of internal models in game players" *Physica D* **134** 253–266
Talen E, 1996, "Do plans get implemented? A review of evaluation in planning" *Journal of Planning Literature* **10** 248–259
Yi R, Johnson M W, Bickel W K, 2005, "Relationship between cooperation in an iterated prisoner's dilemma game and the discounting of hypothetical outcomes" *Learning and Behavior* **33** 324–336

Appendix

Let T_k denote a subtree in the game tree TFT for iteration k. Referring to figure 1, it can be observed easily from the game tree that the following recursive expected payoff functions obtain:

$$\begin{aligned} T_3 &= qT_2 + (1-q)T_2 = T_2 , \\ T_4 &= qT_3 + (1-q)T_3 = T_3 , \\ T_n &= qT_{n-1} + (1-q)T_{n-1} = T_{n-1} . \end{aligned}$$

Therefore, for TFT, $T_n = T_{n-1} = T_{n-2} = \ldots = T_4 = T_3 = q^2 + qb - q^2 b$ for $n \geqslant 2$.

Let P_k denote a subtree in the game tree of AD for iteration k. The recursive expected payoff functions are:

$$\begin{aligned} P_2 &= qP_1 + (1-q)0 = qP_1, \qquad \text{where } P_1 = q , \\ P_3 &= qP_2 + (1-q)0 = qP_2 , \\ P_n &= qP_{n-1} + (1+q)0 = qP_{n-1} = q^n (\forall n \in N) . \end{aligned}$$

Let F_k denote a subtree in the game tree of AC for iteration k. The recursive expected payoff functions are:

$$\begin{aligned} F_2 &= qF_1 + (1-q)F_1 = F_1 , \\ F_3 &= qF_2 + (1-q)F_2 = F_2 , \\ F_n &= qF_{n-1} + (1-q)F_{n-1} = F_{n-1} . \end{aligned}$$

Therefore, $F_n = F_{n-1} = F_{n-2} = \ldots = F_3 = F_2 = F_1 = q \times 1 + (1-q) \times 0 = q$.

Let M_k denote a subtree in the game tree of RA for iteration k. The recursive expected payoff functions are:

$$\begin{aligned} M_2 &= pqM_1 + p(1-q)M_1 + (1-p)qM_1 + (1-p)(1-q)M_1 \\ &= pM_1[q + (1-q)] + (1-p)M_1[q + (1-q)] \\ &= pM_1 + (1-p)M_1 = M_1 \ , \end{aligned}$$

$$\begin{aligned} M_n &= M_{n-1} = M_{n-2} = \ldots = M_3 = M_2 = M_1 \\ &= pq \times 1 + (1-p) \times q \times b = pq + qb - pqb \ . \end{aligned}$$