Title
Three essays on dynamics, strategic default and assortative matching

Author
Rabanal, Jean Paul Paul

Publication Date
2012

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA
SANTA CRUZ

THREE ESSAYS ON DYNAMICS, STRATEGIC DEFAULT AND ASSORTATIVE MATCHING

A dissertation submitted in partial satisfaction of the
requirements for the degree of

DOCTOR OF PHILOSOPHY

in

ECONOMICS

by

Jean Paul Rabanal

June 2012

The Dissertation of Jean Paul Rabanal
is approved:

Professor Daniel Friedman, Chair

Professor Kenneth Kletzer

Professor Carl E. Walsh

Tyrus Miller
Vice Provost and Dean of Graduate Studies

Table of Contents

List of Figures

List of Tables

Abstract

Three essays on dynamics, strategic default and assortative matching

by

Jean Paul Rabanal

This dissertation consists of three chapters. Chapter 1, joint with Daniel Friedman, illustrates general techniques for assessing dynamic stability in games of incomplete information by re-analyzing two models of preference evolution. The techniques include extensions of replicator and gradient dynamics, and for both models they confirm local stability of the key static equilibria. Chapter 2 studies strategic default on home mortgages using an experimental approach. I demostrate that (i) people appear to follow the prediction of the strategic default model quite closely in the high asset volatility treatment, and that (ii) incorporating social interactions delays the strategic default beyond what is considered optimal. Finally, Chapter 3, joint with Olga A. Rabanal, shows that introducing assortative matching to one-shot Prisoners' Dilemma type games enhances cooperative behavior.

To Teffi.

Acknowledgments

My work would not be possible without the support of Dan Friedman and Olya Rabanal. I would also like to thank Professors Michael P. Dooley, Kenneth Kletzer, Phil McCalman, Carl E. Walsh, Nirvikar Singh, Jennifer Poole and Ai-ru Cheng.

Introduction

In this dissertation, I present a dynamic model used to analyze the stability of games of incomplete information as well as experimental evidence regarding two key individual decisions: (i) the decision to default on a loan and (ii) cooperation in a one-shot Prisoners' Dilemma game under pre-specified matching protocols.

Standard equilibrium concepts, such as Bayesian Nash Equilibrium and Sequential Equilibrium, offer sophisticated formulations of what one might hope to see in games of incomplete information. These equilibrium concepts, however, beg the crucial dynamical question: would human players ever actually reach such an equilibrium, or even get close? The Chapter 1 "Incomplete Information, Dynamic Stability and the Evolution of Preferences: Two Examples," joint with Daniel Friedman, addresses these questions. It does not prove new general results nor offer new models, but it does show how to extend replicator dynamics and gradient dynamics from games of complete information to games of incomplete information. We illustrate our techniques, using two models of preference formation due to Arce (2007) and Friedman and Singh (2009). For both models, we confirm local stability of the key static equilibria.

The Chapter 2 "Strategic default with neighbors: A laboratory experiment" seeks to shed more light on the factor driving the strategic default. In the United States, following the collapse of home prices in 2007-08, millions of households found themselves underwater. Assets, such as real estate, generally represent a substantial portion of household wealth. Therefore, strategic default can be especially relevant in an attempt to protect household assets. When a sufficient number of households begin to see a decline in their real wealth, the effect on the economy can become quite strong due to changing consumption and investment levels. Given the importance of real estate to household wealth and to overall economic health, it would be of interest to study the driving forces of strategic default. Greater understanding of the causes and consequences of strategic default would allow for more appropriate response and thus more efficient policy-making. In this chapter, I demonstrate that people appear to follow the prediction of the strategic default model quite closely, and that incorporating social interactions delays the strategic default beyond what is considered optimal.

The Chapter 3, "Cooperative behavior in one-shot games with assortative matching" joint with Olga A. Rabanal, introduces a mechanism that enhances cooperation through group formation, or assortative matching. Consider an investment project that requires some effort on the part of the worker and likewise a wage on the part of the manager to compensate the worker's effort. A worker can choose to fully cooperate and select a high-level of effort or shirk responsibility (defect). Similarly, the manager can fully compensate the worker or choose to defect (low wage). The actions are simultaneous and the choice of cooperation or defection exemplifies a Prisoners' Dilemma game. Often, these projects can be a one-shot deal and therefore according to theory and vast experimental evidence lead to a suboptimal outcome. Preliminary results show that under assortative matching we can elicit cooperative behavior and thus achieve a pareto superior outcome as compared to the random matching group formation.

Chapter 1

Incomplete Information, Dynamic Stability and the Evolution of Preferences: Two Examples (with Daniel Friedman)

1.1 Introduction

Standard equilibrium concepts, such as Bayesian Nash Equilibrium and Sequential Equilibrium, offer sophisticated formulations of what one might hope to see in games of incomplete information. These equilibrium concepts, however, beg the crucial dynamical question: would human players ever actually reach such an equilibrium, or even get close?

For games of complete information, such questions of dynamic stability have been addressed in a principled way by evolutionary game theory. That theory shows that certain subsets of Nash equilibrium (e.g., Evolutionary Stable States, ESS) are indeed reachable by players using simple adaptation rules (e.g., replication or imitation or learning); see, for example

Weibull (1995), Friedman (1991) and Sandholm (2011) for games in normal form and Cressman (2003) for games in extensive form. For games of incomplete information, however, dynamic stability questions remain largely unresolved.

The present paper begins to address those questions. It does not prove new general results nor offer new models, but it does show how to extend replicator dynamics and gradient dynamics from games of complete information to games of incomplete information. In some cases the extensions continue to yield systems of ordinary differential equations (ODEs) but in other cases they yield partial differential equations (PDEs), and the stability properties of these systems can be investigated analytically or numerically. We illustrate, using two recent models of preference formation due to Arce (2007) and Friedman and Singh (2009).

Both papers use the indirect evolutionary approach to model preference evolution. Players of given preference types are matched pairwise to play a game with known material payoffs. Players learn (on a rapid time scale) to maximize expected utility given their type, but evolve (on a slower time scale) according to realized material payoffs. Previous investigations represented evolution in terms of a static notion such as ESS; e.g. see Güth and Yaari (1992), Ok and Vega-Redondo (2001) and Dekel, Ely and Yilankaya (2007). Here we will instead use standard dynamic specifications — replicator and gradient — to assess the stability of key equilibria of both models.

The current paper is organized as follows. Section 1.2 presents a Principal-Agent game of incomplete information due to Arce (2007). A population of Agents containing two types (one self-interested and the other autonomy-preferring) is randomly matched pairwise with a population of Principals who can not observe Agents' type. Which preference types and what sort of behavior will survive in the long run? Arce uses static concepts to analyze the dynamic stability of various equilibria, and emphasizes the result that increasing the incentive wage can destabilize an efficient equilibrium for some distributions of preference types. He also finds some cases where the standard agency equilibrium can be destabilized.

Our analysis of the model begins by writing out the expected payoff and expected utilities given all state variables including the population share of each type of Agents, which Arce (2007) takes as exogenous. Then we introduce and analyze a system of four coupled ordinary differential equations (ODEs) that characterizes the evolution of the state variables. That system uses standard replicator dynamics, with fitness given by the realized material payoffs, to model the time path of the population shares. To model adjustment of strategy mixtures, the ODE system focuses on utility and applies gradient dynamics, which in this context looks much like standard replicator dynamics. The parameters include adjustment speed, and we focus on cases where strategy mixture adjustment is faster than the evolution of types.

For some equilibria, eigenvalue techniques allow us analytically to characterize dynamic stability. However, for some key equilibria, these and other tractable analytic techniques are inconclusive. We then rely on numerical solutions of the ODE system, and explain why it is appropriate in such cases to focus on time averages. Our results complement and extend those of Arce (2007), and in particular we find that both kinds of Agents can coexist for a broad set of parameter values.

Section 1.3 describes the basic trust game and its extension to a game of incomplete information due to Friedman and Singh (2009, henceforth FS09). They proposed a static equilibrium refinement called Evolutionary Perfect Bayesian Equilibrium (EPBE) for population games. At EPBE, each surviving type in each population has the same expected material payoff, and no potential entrant type has higher payoff.

After writing out the expected payoffs and utilities, we derive a system of six coupled differential equations that characterizes the evolution of the state variables. Five of these equations roughly parallel the ODEs for the Arce (2007) model, and the other equation uses gradient dynamics to describe evolution of the preference parameter. In line with FS09 and the previous section, we assume that individual learning enables strategy mixes to adjust more rapidly than population shares (which adjust via entry and exit, or type switching). We assume that preference parameters adjust (via genetic disposition and/or internalized norms) at an even slower rate. Our results support the implicit assumption in FS09 that EPBE is dynamically stable. More specifically, we show for reasonable parameters that the time average state converges to the relevant EPBE from initial conditions near to the equilibrium value. That is, the EPBE of the noisy trust games is locally asymptotically stable in time average.

The last section summarizes our findings and offers suggestions for applying the techniques more broadly.

1.2 The Arce (2007) Principal-Agent Model

After reviewing the model and its known equilibria, we write out the state-contingent payoffs and specify dynamics. Then we assess the dynamic stability of all equilibria.

1.2.1 Elements of the model

Each Principal (row player) in a large population is randomly matched with one Agent (column player). There are two possible types of Agents: Type 1 or self-interested (comprising a fraction $\varphi \in [0,1]$ of the population) and Type 2 or autonomy-preferring (fraction $1-\varphi$). Either type of Agent decides only to whether to work (W) or to shirk (S); the respective mixture probabilities are denoted q_j and $1-q_j$ for each type $j = 1,2$.[1] Thus Agents' type-contingent strategy space is $[0,1]\times[0,1]$.

Each Principal decides whether to monitor (M) or not (N); the mixture probabilities are p and $1-p$ respectively. Thus Principal's strategy space is $[0,1]$, and the state of the system is a vector $S = (\varphi, p, q_1, q_2) \in [0,1]^4$.

Table 1.1 shows the payoffs. The Principal receives gross payoff $v > 0$ when the Agent works, offset by the wage $w > 0$ paid to the worker and by the monitoring cost $m > 0$. Agents who work incur an effort cost $e > 0$. In addition to these material payoffs, a Type 2 Agent receives a utility increment $(+\alpha)$ when her Principal does not monitor, and a symmetric decrement $(-\alpha)$ with monitoring. The parametric restrictions $w > e, m$ and $\alpha > e, w-e$ help avoid trivialities. Arce (2007) sets wage at the value $w = \sqrt{vm}$ that maximizes the Principal's expected payoff; this implies restrictions on the gross payoff v.

[1]The *monomorphic* interpretation of a mixture probability q_j is that every type j player adopts exactly the same mixed strategy q_jW + $(1-q_j)$S. The *polymorphic* interpretation is that a fraction q_j of the type j players adopt the pure strategy W and the rest adopt the pure strategy S. The analysis below works for either interpretation, as well as for the more general interpretation that there is a distribution of pure and mixed strategies among the the type j players with overall mean q_j.

Table 1.1: **Principal (row) and Agent (column) Payoffs**. The fraction of Type 1 Agents is φ, and mixing probabilities are p for Principal and q_j for Agents of Type j.

	Type 1 (φ)		Type 2 ($1-\varphi$)	
	W (q_1)	S (1-q_1)	W (q_2)	S (1-q_2)
M (p)	$v-w-m \quad ,w-e$	$-m,0$	$v-w-m \quad ,w-e-\alpha$	$-m,0$
N (1-p)	$v-w \quad ,w-e$	$-w,w$	$v-w \quad ,w-e+\alpha$	$-w,w$

Source: Arce (2007)

Arce notes that if all Agents are known to be Type 1, then the unique Nash equilibrium is mixed: $p = p^* = e/w$; $q_1 = q_1^* = (w-m)/w$. He also notes that if all Agents are known to be Type 2, there is again a mixed NE in which $p = p^{**} = (\alpha - e)/(2\alpha - w)$, $q_2 = q_2^* = (w-m)/w$ as well as two pure NE: one at (N,W) or $p = 0$, $q_2 = 1$ and the other at (M,S) or $p = 1$, $q_2 = 0$.

1.2.2 Expected payoffs and utilities

Which equilibria, if any, are dynamically stable? Before introducing evolutionary dynamics to answer that question, we set the stage by writing out expected payoffs and utilities.

The Principal's expected payoff ω^P in equation (1.1) below arises from receiving v when the Agent works (probability $\varphi q_1 + (1-\varphi)q_2$), minus the monitoring costs incurred (with probably p) and the wages paid to the Agent. Recall from Table 1.1 that the Principal pays w unless he monitors and the Agent shirks, an event of probability $p(\varphi(1-q_1)+(1-\varphi)(1-q_2))$.

Thus

$$\omega^P = (\varphi q_1 + (1-\varphi)q_2) \cdot v - p \cdot m - [1 - p(\varphi(1-q_1) + (1-\varphi)(1-q_2))] \cdot w \tag{1.1}$$

Both types of Agent receive material payoff $w-e$ if they work (probability q_i), or w in the event that they do not work $(1-q_i)$ and the Principal does not monitor $(1-p)$. For the self-interested Agent (type 1), expected utility coincides with expected material payoff ω^{A1}, which is therefore

$$\omega^{A1} = q_1 \cdot (w-e) + (1-q_1)(1-p) \cdot w. \tag{1.2}$$

Similarly, material payoff for Type 2 Agent is

$$\omega^{A2} = q_2 \cdot (w-e) + (1-q_2)(1-p) \cdot w, \tag{1.3}$$

while her expected utility includes the preference parameter α and is

$$\omega_\alpha^{A2} = q_2 \cdot [p \cdot (w-e-\alpha) + (1-p) \cdot (w-e+\alpha)] + (1-q_2)(1-p) \cdot w. \tag{1.4}$$

1.2.3 Dynamic adjustment equations

Recall that the state space is four dimensional, and specifies the fraction of type 1 Agents (φ), Principal's mixing probability (p) and Agents' mixing probabilities (q_j). We therefore specify dynamics as a system of four coupled ordinary differential equations (ODEs), derived from expected payoffs using standard evolutionary principles.

Arce (2007, p. 718) comments, "This then begs the question, what determines the initial distribution of agent types (ρ)?" and cites several exogenous factors. For our purposes it is better to complete the model by endogenizing ρ, or φ in our notation. We invoke the basic principle of evolution that the type with higher material payoff (= fitness) will increase its share of the population. More specifically, we impose standard continuous time replicator dynamics (Taylor and Jonker, 1978; Hofbauer and Sigmund, 1988), which postulate that the growth rate $\dot{\varphi}/\varphi$ of the share of self-interested Agents is proportional (with rate constant β_1) to its payoff ω^{A1} relative to the population average ($\bar{\omega}$). Equation (1.5) and other equations below use the fact that relative payoff can be written $\omega^{A1} - \bar{\omega} = \omega^{A1} - \varphi\omega^{A1} - (1-\varphi)\omega^{A2} = (1-\varphi)(\omega^{A1} - \omega^{A2})$. Thus $\dot{\varphi}$ is equal to $\varphi(1-\varphi)(\omega^{A1} - \omega^{A2})$ times a positive adjustment speed parameter β_1.

The remaining equations use hybrid gradient-replicator dynamics for the mixture probabilities p, q_1 and q_2. Gradient dynamics are standard for evolution of continuous biological traits (e.g., Wright (1949), Lande (1976) and Kauffman (1993)), and for continuous strategy sets seem more common in economics (e.g., Sonnenschein (1982), Friedman and Ostrov (2010)) than alternative specifications. In these dynamics, the adjustment rate is proportional to its payoff gradient $\frac{\partial \omega}{\partial p}$. To shrink to the range $[0,1]$ of valid mixtures, we include a factor of the form $(1-p)p$, analogous to the binomial variance $(1-\varphi)\varphi$ that keeps φ in the interval $[0,1]$. Of course, the mixing probability q_2 responds to expected utility ω_{α}^{A2}, rather than to the material payoff that guides the evolution of preference types.

Thus the system of four coupled ODEs is

$$\dot{\varphi} = \beta_1 \varphi(1-\varphi)[\omega^{A1} - \omega^{A2}] \tag{1.5}$$

$$= \beta_1 \varphi(1-\varphi)[(pw-e)(q_1-q_2)]$$

$$\dot{p} = \beta_2 p(1-p)\frac{\partial \omega^P}{\partial p} \tag{1.6}$$

$$= \beta_2 p(1-p)[-m+(\varphi(1-q_1)+(1-\varphi)(1-q_2))\cdot w]$$

$$\dot{q}_1 = \beta_3 q_1(1-q_1)\frac{\partial \omega^{A1}}{\partial q_1} \tag{1.7}$$

$$= \beta_3 q_1(1-q_1)(pw-e)$$

$$\dot{q}_2 = \beta_4 q_2(1-q_2)\frac{\partial \omega^{A2}_{\alpha}}{\partial q_2} \tag{1.8}$$

$$= \beta_4 q_2(1-q_2)[(w-2\alpha)p+\alpha-e]$$

where the parameters w, e, m, α, and β_i are exogenous.

We assume that individual learning enables the mixing probabilities to adjust more rapidly than the population fraction φ, which adjusts via entry and exit, or type switching. Thus $0 < \beta_1 < \beta_2 = \beta_3 = \beta_4$. The restrictions noted earlier apply to parameters w, e, m, α (or to v, e, m, α if w is chosen by the formula mentioned earlier). To complete the dynamic specification, take the initial state as given and impose the boundary conditions $0 \leq p \leq 1$, $0 \leq q_j \leq 1$ and $0 \leq \varphi \leq 1$.

It might seem at first that we are dealing with a system of partial differential equations, but the gradients on the right hand side can be expressed in terms of the state variables only, using equations (1.1-1.3). For example, using equation (1.2) in equation (1.7) we have $\frac{\partial \omega^{A1}}{\partial q_1} = \omega^{A1}|_{[q_1=1]} - \omega^{A1}|_{[q_1=0]} = pw - e$. Indeed, since the functions are linear in the relevant probability, these gradients all can be replaced by fitness difference terms, and the second lines of equations (1.6 - 1.8) therefore resemble standard replicator equations.

1.2.4 Dynamic behavior

Describing the dynamic behavior of a system of 4 ordinary differential equations depending on 8 exogenous parameters sounds like a complicated task. However, for our purposes it suffices to identify the dynamically stable equilibrium (DSE) points — the subset of rest points or steady states that are Lyapunov stable. That is, we seek steady states (states for which the right hand side of the ODE system is zero) such that a solution of the ODE system with initial condition sufficiently close to the steady state will remain close to the steady state forever. The idea is that only neighborhoods of DSE are likely to be empirically relevant; elsewhere behavior is transient and will be hard to identify in field data.

Two technical remarks are in order before proceeding. First, Lyapunov stability does not guarantee asymptotic stability, i.e., does not guarantee that the solution above actually converges to the DSE as $t \to \infty$. For the systems we are studying, asymptotic stability is too much to ask for, because in two or more dimensions Liouville's theorem precludes direct convergence of replicator dynamics to any interior equilibrium from an open neighborhood of initial conditions; see e.g., Fudenberg and Levine, 1998, p.95.

Second, it is well known that Nash equilibria (NE) are a subset of steady states (or dynamic equilibria, DE); see for example Weibull, 1995, Proposition 3.4. It is also well known that DSE are a subset of NE and that, for smooth systems of ODEs like (1.5 - 1.8), a necessary condition for DSE is that the Jacobian matrix evaluated at the NE has no eigenvalues with positive real part and a sufficient condition is that all eigenvales have negative real parts; see for example Hirsch and Smale, 1974, Chapter 9. Eigenvalues with zero real part suggest (but do not guarantee) Lyapunov stability, and suggest (again with no guarantee) failure of asymptotic stability.

We will therefore use the following algorithm to identify DSE:

- find all DE, separately checking corners, edges, faces and interior of the state space;
- identify the subset of DE that are NE, and eliminate the others;
- find the eigenvalues of the Jacobian matrix of (1.5 - 1.8) evaluated at each NE, and eliminate any NE which yields an eigenvalue with positive real part; and
- identify as locally stable (and therefore empirically relevant) any NE whose eigenvalues all have negative real parts, and use numerical methods to assess the dynamic stability of any remaining NE that yields an eigenvalue with zero real part.

To begin, recall that by definition a DE for the present model is a solution to

$$\dot{\varphi} = \dot{p} = \dot{q}_1 = \dot{q}_2 = 0. \tag{1.9}$$

To sort out the many solutions, recall that our state vector $S = (\varphi, p, q_1, q_2) \in [0,1]^4$ is a four dimensional hypercube. Each of the $2^4 = 16$ corners represents a pure strategy profile, and (by virtue of the binomial factors) is a solution to (1.9). One strategy is mixed along each of the $16 \cdot 4/2 = 32$ edges, two are mixed in each 2-d face ($4 \cdot (4 \cdot 3)/2 = 24$ of them), and three are mixed in each 3-d face ($4 \cdot 2 = 8$ of them), while interior points represent strictly mixed strategy profiles.

The first step in the algorithm, then, gives us 16 corner DE. Checking all edges and faces sounds tedious, but the special structure of the model allows shortcuts. When $\varphi = 0$ (or 1), the value of q_1 (or q_2) is irrelevant, so 8 of the edges and all 16 corners are subsumed in the DE subset $\{(1,0,0,\cdot),(1,0,1,\cdot),(1,1,0,\cdot),(1,1,1,\cdot);(0,0,\cdot,0),(0,1,\cdot,1),(0,1,\cdot,0),(0,0,\cdot,1)\}$. Recall from Section 1.2.1 that only the last two cases are pure NE in the restricted ($\varphi = 0,1$ face) games, so we can eliminate the other six cases as dynamically unstable. Indeed, the same argument allows us to eliminate all edge DE in either of these faces. So the only remaining edges are of the form (i) $\varphi \in (0,1)$ and (ii) $p, q_1, q_2 \in \{0,1\}$. From equations (1.9) and (1.5) we see that (i) entails either $p = w/e$ (which is inconsistent with (ii)) or $q_1 = q_2$, which is "pure pooling" by (ii). But $p = 1$ and $q_1 = 0$ are not mutual best responses, nor are $p = 0$ and $q_1 = 1$. Hence the only remaining edge DE are of the form $(\varphi,1,1,1)$.

Appendix A collects arguments of a similar character that show the full set of NE is

$$\{(0,1,\cdot,0),(0,0,\cdot,1),(x,1,1,1),(1,p^*,\frac{w-m}{w},\cdot),(0,p^{**},\cdot,\frac{w-m}{w}),(\frac{w-m}{w},p^*,1,0)_{[\mathrm{lw}]},$$
$$(\frac{m}{w},p^*,0,1)_{[\mathrm{hw}]},(x,p^*,\frac{-m+wx}{wx},1),(x,p^*,\frac{-m+w}{wx},0)\}. \quad (1.10)$$

Here x is in some parameter-dependent subset of [0,1] specified as needed below, while $p^* = e/w$, $p^{**} = \frac{\alpha-e}{2\alpha-w}$, and [hw] (or [lw]) in subscripts indicates that the state is a NE in the high wage region of parameter space $w-2e > 0$ (or in the low wage region $w-2e < 0$).

The next step in the algorithm is to write out the Jacobian matrix $((\frac{\partial \text{RHS eq. } i}{\partial \text{state var. } j}))$, evaluate at each NE, and compute the eigenvalues. The Jacobian of ODE system (1.5 - 1.8) is

$$J = \begin{pmatrix} \beta_1(1-2\varphi)(pw-e)(q_1-q_2) & \beta_1\varphi(1-\varphi)w(q_1-q_2) & \dots \\ \beta_2 p(1-p)w(q_2-q_1) & \beta_2(1-2p)(w-m-q_2w+\varphi w(q_2-q_1)) & \dots \\ 0 & \beta_3 q_1(1-q_1)w & \dots \\ 0 & \beta_4 q_2(1-q_2)(w-2\alpha) & \dots \end{pmatrix}.$$

$$J = \begin{pmatrix} \dots & \beta_1\varphi(1-\varphi)(pw-e) & -\beta_1\varphi(1-\varphi)(pw-e) \\ \dots & -\beta_2 w(p-p^2)\varphi & -\beta_2 w(p-p^2)(1-\varphi) \\ \dots & \beta_3(1-2q_1)(pw-e) & 0 \\ \dots & 0 & \beta_4(1-2q_2)[(w-2\alpha)p+\alpha-e] \end{pmatrix}.$$

As a warmup exercise, we compute the 2x2 Jacobian (sub)matrix for 2-d face where $\varphi = 0$ and $p, q_2 \in (0,1)$. At the pure NE $(0,1,\cdot,0)$ it is

$$J = \begin{pmatrix} -\beta_2(w-m) & 0 \\ 0 & -\beta_4(\alpha-(w-e)) \end{pmatrix}$$

and at the other pure NE $(0,0,\cdot,1)$ it is

$$J = \begin{pmatrix} -\beta_2 m & 0 \\ 0 & -\beta_4(\alpha-e) \end{pmatrix}$$

For these diagonal matrices, the diagonal entries are the eigenvalues, and the parametric restrictions guarantee that all of them are negative. Hence these equilibria are both stable “sinks” with respect to dynamics restricted to the face. To assess overall stability, we have to look at the full Jacobian matrix, and the Appendix shows that these include positive eigenvalues. Hence neither NE is a DSE.

The same Jacobian (sub)matrix evaluated at the mixed NE $(0, p^{**}, \cdot, q_2^*)$, where $q_2^* =$

$\frac{w-m}{w}$, is

$$J = \begin{pmatrix} 0 & -\beta_2 p^{**}(1-p^{**})w \\ -\beta_4 q_2^*(1-q_2^*)(2\alpha - w) & 0 \end{pmatrix},$$

whose eigenvalues are real and have opposite signs, since the parametric restrictions imply that $2\alpha - w > 0$. Therefore the equilibrium is an unstable saddle point even with respect to dynamics restricted to the face, and thus is not a DSE.[2]

The systematic way to assess stability is to evaluate the 4x4 Jacobian matrix at each NE. To illustrate, take the last NE listed, $\left(x, \frac{e}{w}, \frac{w-m}{wx}, 0\right)$, where $x \in [\frac{w-m}{w}, 1]$. The Jacobian evaluated at such states is

$$J = \begin{pmatrix} 0 & \beta_1(w-m)(1-x) & 0 & 0 \\ \frac{\beta_2 e(w-m)(w-e)}{w^2 x} & 0 & -\beta_2 e\left(1-\frac{e}{w}\right)x & -\beta_2 e\left(1-\frac{e}{w}\right)(1-x) \\ 0 & \frac{\beta_3(w-m)(m+w(-1+x))}{wx^2} & 0 & 0 \\ 0 & 0 & 0 & \frac{\beta_4(w-2e)\alpha}{w} \end{pmatrix},$$

whose eigenvalues are $\left\{0, \frac{\beta_4(w-2e)\alpha}{w}, \pm\frac{\sqrt{-e(w-e)(w-m)^2(1-x)\beta_1\beta_2 - [m+w(-1+x)]\beta_3}}{w\sqrt{x}}\right\}$. The second eigenvalue is negative in the low-wage region and positive in the high wage region, while the last pair of eigenvalues is pure imaginary since the expression in square brackets is positive for $x \geq \frac{w-m}{w}$. Hence this NE is dynamically unstable in the high wage region but remains a candidate DSE in the low wage region, where we will examine it numerically.

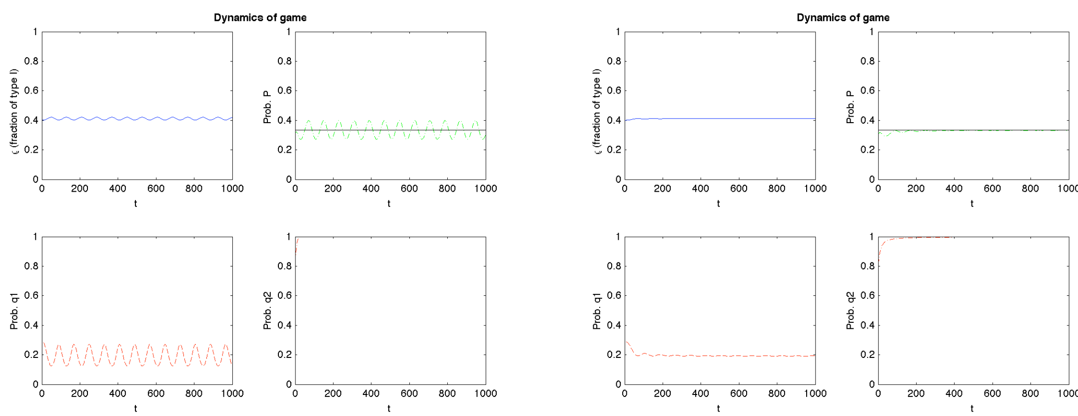
The Appendix examines the other NE using the same techniques. It rules out DSE status for the first five in the list, and confirms that there are no asymptotically stable NE. The only remaining candidate DSE are $\{\left(x, \frac{e}{w}, \frac{w-m}{wx}, 0\right) : x \in [\frac{w-m}{w}, 1]\}$ in the low wage case and

[2]More concretely, one can check that, for small $\varphi > 0$, type I Agent entrants will play a pure strategy that gives them a higher payoff than incumbents' payoff given that Principal plays p^*.

$\{(x, \frac{e}{w}, \frac{wx-m}{wx}, 1) : x \in [\frac{m}{w}, 1]\}$ in the high wage case.

1.2.5 Simulation results

The last step in our algorithm is to investigate convergence behavior of the candidate DSE numerically.



(a) Levels (b) Time average

Figure 1.1: Dynamics in the high wage $w > 2e$ case

Since we do not expect asymptotic stability, we look for convergence in time average of the state variable $S = (\varphi, p, q_1, q_2)$. That is, we shall emphasize numerical approximations of $\lim_{t\to\infty} t^{-1} \int_0^t S(u)du$ more than of $\lim_{t\to\infty} S(t)$.[3]

[3]Stability in time average is also emphasized in the equilibrium concept Time Average of the Shapley Polygon (TASP) proposed by Benaim, Hofbauer and Hopkins (2006).

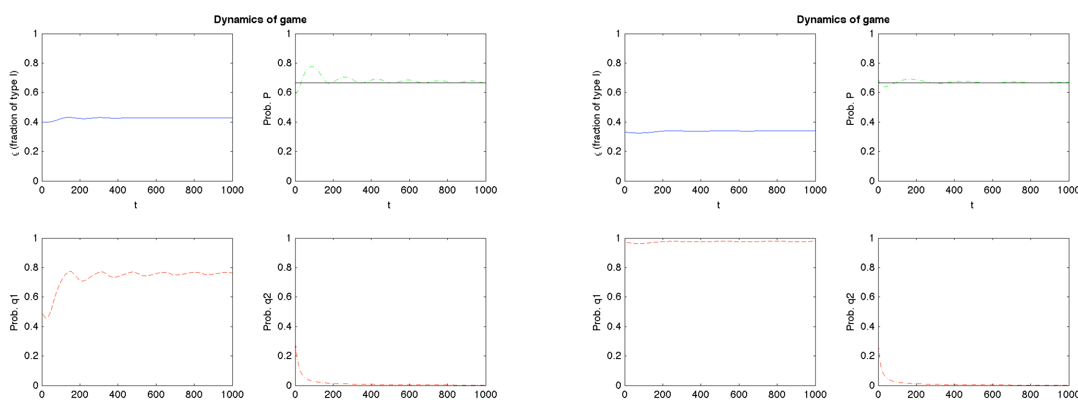
We solve the ODE system numerically using Matlab (codes are available upon request). To illustrate, set speeds of adjustment $\beta_1 = 0.1$ and $\beta_j = 1$ for $j = 2,3,4$ and use baseline parameters $\alpha = 0.5, m = 0.2, e = 0.2, v = w^2 m$, where the high wage is $w = 0.6 > 2e$ and the low wage is $w = 0.3 < 2e$.

For these baseline parameters, the candidate DSE is $(x, 1/3, 1 - \frac{1}{3x}, 1)$, $x \in [\frac{1}{3}, 1]$ for the high wage. Figure 1.1 shows our simulation results for the high incentive wage. We start from initial values $q_1(0) = p(0) = 0.3$, $q_2(0) = 0.8$ and $\varphi(0) = 0.4$. The left panel shows the simulation results in levels and the right panel in time average. Notice that in the left panel dynamics follow a cycle around the interior solution $\varphi^* = 0.41$ and $q_1^* = 0.19$ with direct convergence to $p^* = 1/3$ and $q_2^* = 1.0$. We achieve convergence in time average to that equilibrium in the right panel.

The dynamics are sensitive to the initial values. The farther the initial values from equilibrium, the wider the cycle around it. The time average convergence is not affected by choices of initial values for p^* and q_2^* sufficiently near to the equilibrium; meanwhile the equilibrium values of q_1^* and φ^* vary accordingly to the model predictions. For initial values far enough from the key equilibrium, we do not achieve convergence. For instance, starting with $q_2(0) < 0.1$, the level of monitoring approaches zero, the fraction of type I agents and the mixing probabilities go to 1. This result is not consistent with the target equilibrium.

The other relevant case is the low wage equilibrium, $(x, 2/3, \frac{1}{3x}, 0), x \in [1/3, 1]$ for baseline parameters. Notice that the level of monitoring is higher compared to the high wage equilibrium and type II Agent shrinks. Figure 1.2 shows the numerical results. The left panel starts with $\varphi(0) = 0.4$, $p(0) = 0.6$, $q_1(0) = 0.5$ and $q_2(0) = 0.3$. It shows convergence in time average to the equilibrium $(0.43, 0.67, 0.76, 0.00)$. The right panel starts with $\varphi(0) = 0.33$, $p(0) = 0.7$, $q_1(0) = 0.97$ and $q_2(0) = 0.25$. It shows convergence in time average to the equilibrium $(0.34, 0.67, 0.98, 0.00)$.

The dynamics are similar to the high wage case. Once again, starting from initial values far enough from the key equilibrium, we do not achieve convergence. For instance, assuming initial values similar to the left panel but varying $p(0) < 0.2$, the dynamics are explosive and p, q_1 and q_2 go to the corners meanwhile the fraction of types stays in an interior point. This result is not consistent with the target equilibrium.



(a) Low q_1

(b) High q_1

Figure 1.2: Dynamics in the low wage $w < 2e$ case

We complement Arce's work by endogenously determining the fraction of types and studying the dynamics of the model. As different from his work, we show that both types of Agent coexist independently of the level of the incentive wage for a broad set of parameter values and initial conditions around the equilibrium values. *Ceteris paribus*, the high incentive wages leads to lower rates of monitoring from the Principal and foster the working decision of autonomy-preferring Agents; meanwhile, the low incentive wages causes high monitoring levels and not working for the autonomy-preferring Agents. Our numerical solution shows that the dynamics when both Agents are present follow a cycle around the relevant interior equilibrium.

1.3 The Friedman and Singh (2009) Noisy Trust Game

Analyzing the next model introduces several additional considerations that can be important in games of incomplete information, such as positive tremble rates, evolving preference parameters, and higher dimensional state spaces. We rely more on numerical simulation but are nevertheless able to get a sharp result.

To begin, consider a simple two player game of complete information. The first mover, labelled Self (S), chooses whether to trust (T) or not trust (N). Choice N ends the game with zero payoffs to both players. Choice T gives the move to player Other (O), who can choose either to to cooperate (C) or defect (D). Choice C gives both players unit payoffs, while choice D yields payoffs 2 to Other and -1 to Self. Following D, a vengeful behavioral type Self ($v = v_H > 0$) will take revenge and, at cost v to himself, will inflict harm v/c on Other, given an

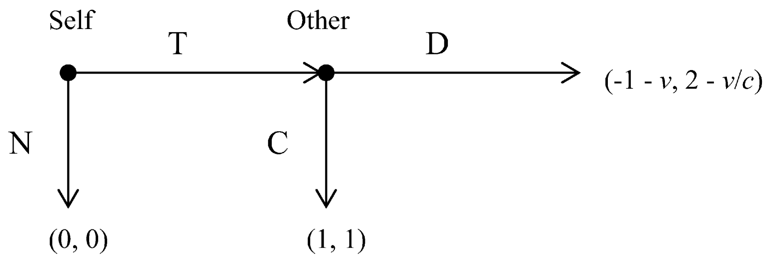


Figure 1.3: Simple trust game.

Unique subgame perfect NE is (N,D) when $v < c$ and is (T,C) when $v > c$.

exogenous marginal cost parameter $c > 0$. The equilibrium payoffs are inefficient at (0,0) when $v = 0$, but are efficient at (1,1) when $v = v_H > c$.

1.3.1 Elements of the model

From this simple game, FS09 construct the noisy trust game illustrated in Figure 1.4. Nature chooses Self's non-vengeful type $v = 0$ with probability $1 - x$, or else chooses a given vengeful type $v = v_H > 0$ with probability x. Nature also independently chooses Other's perception as correct ($s = 0$ for $v = 0$, or $s = 1$ for $v = v_H$) with probability $1 - a$, or incorrect with probability a. The misperception probability is given by:

$$a = A(v_H) = 0.5\exp(-kv_H^2) \tag{1.11}$$

where $k > 0$ represents a precision parameter.

Let $p_0 = Pr[T|v = 0]$ denote the probability of trusting when S is non-vengeful, and $p_1 = Pr[T|v = v_H]$ the probability of trusting when S is vengeful. These probabilities are constrained by a tremble rate $e \in [0, 1/2)$, so that $e \le p_0, p_1 \le 1 - e$. Self's (mixed) strategy space

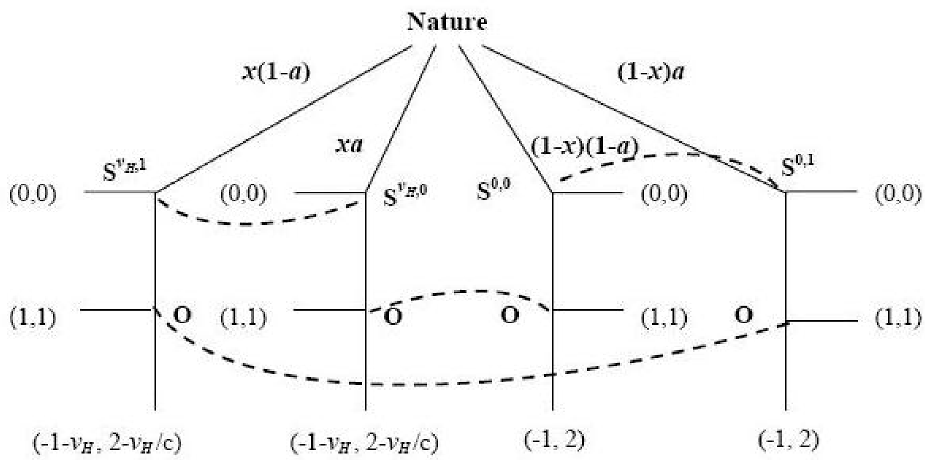


Figure 1.4: The noisy trust game. O denotes Other, S^{ij} denotes Self with vengeance level i and perception j, as determined by Nature's move. The four branch labels are Nature's move probabilities. Source: FS09

thus is $[e, 1-e] \times [e, 1-e]$. Similarly, let $p_2 = Pr[C|s=1]$ and $p_3 = Pr[C|s=0]$ denote the probabilities of cooperating when Other observes a non-vengeful type and a vengeful type, respectively. Other's strategy space is also $[e, 1-e] \times [e, 1-e]$.

The state of the system is a vector $(v, x, p_0, p_1, p_2, p_3) \in [0, \hat{v}] \times [0,1] \times [e, 1-e]^2 \times [e, 1-e]^2$ that specifies Self's actions (p_0 and p_1), Other's actions (p_2 and p_3), the fraction of the vengeful type (x) and its degree of vengefulness ($v = v_H \leq \hat{v}$). This state space is more complicated than that of the Arce model in several respects. Besides the additional mixing variable, we have here a restricted mixture space (to account for trembles, which are conceptually important according to FS09), and an endogenous preference parameter.[4] Topologically, the state space is the 6-d hypercube $[0,1]^6$ with a specific parametrization.

[4]We could also have endogenized α in Arce's model, but that would not have been useful since material payoffs are flat in α except for a discontinuity at a particular threshold that changes type 2 agents' behavior. We will see that evolving v makes good sense in the FS09 model.

The equilibrium concept here is perfect Bayesian equilibrium (PBE). Proposition 1 of FS09 identifies seven families of PBE that depend on game parameters x, a, v_H and e. The pure strategy PBE equilibria lie on the 2-d faces $p_i \in \{e, 1-e\}, i = 0, 1, 2, 3$ of state space, and comprise families called "Separating " when $p_0 = p_3 = e$ and $p_1 = p_2 = 1-e$; "Bad Pooling" when $p_0 = p_1 = p_2 = p_3 = e$; and "Good Pooling " when $p_0 = p_1 = p_2 = 1-e$ and $p_3 = e$. The mixed strategy PBE families lie on higher dimensional faces and are called "Bad Mix" when $p_0 = p_3 = e$ and $p_1, p_2 \in (e, 1-e)$; "Bad Hybrid" when $p_0 = p_3 = e$, $p_1 = 1-e$ and $p_2 \in (e, 1-e)$; "Good Mix" when $p_1 = p_2 = 1-e$ and $p_0, p_3 \in (e, 1-e)$; and "Good Hybrid" when $p_0 = p_1 = p_2 = 1-e$ and $p_3 \in (e, 1-e)$.

1.3.2 Expected payoffs and utilities

The expected payoffs w_s^v and w_s of vengeful and non-vengeful types of Self are:

$$w_s^v = (p_1(1-a)p_2 + p_1 a p_3) * 1 + (p_1(1-a)(1-p_2) + p_1 a(1-p_3)) * (-1-v) \quad (1.12)$$

$$w_s = (p_o(1-a)p_3 + p_o a p_2) * 1 + (p_o(1-a)(1-p_3) + p_o a(1-p_2)) * (-1) \quad (1.13)$$

And the expected payoffs w_o^s and w_o for Other when he perceives a vengeful or a non-vengeful type are:

$$w_o^s = (x(1-a)p_1 p_2 + (1-x)a p_o p_2) * (1) + (x(1-a)p_1(1-p_2)) * (2 - v/c) + ((1-x)a p_o(1-p_2)) * 2 \quad (1.14)$$

$$w_o = (x a p_1 p_3 + (1-x)(1-a)p_o p_3) * (1) + (x a p_1(1-p_3)) * (2 - v/c) + ((1-x)(1-a)p_o(1-p_3)) * 2 \quad (1.15)$$

Equation (1.12) is derived as follows. If vengeful Self does not trust (probability $1-p_1$), she receives a zero payoff. On the other hand, if she trusts (probability p_1), she gets payoff 1 or $-1-v$ depending on Other's decision and perception. Her payoff is 1 when Other perceives correctly (probability $(1-a)$) a vengeful type and cooperates (probability p_2), and also when Other misperceives (probability a) and cooperates (probability p_3). She gets $-1-v$ when Other perceives correctly $(1-a)$ and defects $(1-p_2)$, and also when she misperceives (a) and defects $(1-p_3)$. Similar logic yields the expressions (1.13) and (1.15) for non-vengeful Self's payoff w_s, and the expected payoffs w_o^s and w_o for Other when he perceives a vengeful or a non-vengeful type, respectively.

1.3.3 Dynamic adjustment equations

Recall that the state space is six dimensional, and specifies the fraction of vengeful type (x), the degree of vengefulness (v), and four mixing probabilities (p_i). We therefore specify dynamics as a system of six coupled ordinary differential equations (ODEs), derived from expected payoffs using standard evolutionary principles.

For the share x of vengeful types in the Self population, replicator dynamics postulate that the growth rate $\dot{x}/x$ is proportional (with rate constant β_x) to its own payoff w_s^v relative to the population average.

The remaining state variables involve a continuum of alternatives. Here we rely on gradient dynamics. Thus the degree of vengefulness $v = v_H$ changes at a rate proportional to its gradient $\frac{\partial w_s^v}{\partial v}$. As before, we use hybrid gradient-replicator dynamics for each mixing probability p_i. Its adjustment rate is proportional to its payoff gradient $\frac{\partial w_s^{[v]}}{\partial p_i}$. To shrink the range to $[e, 1-e]$, we include factors $(1-e-p_i)(p_i-e)$, analogous to the binomial factors $(1-x)x$ that keep x in the interval $[0,1]$. Thus our system of six ODEs is:

$$\dot{v} = \beta_v \left(\frac{\partial w_s^v}{\partial v} \right) \tag{1.16}$$

$$\dot{x} = \beta_x (1-x)x(w_s^v - w_s) \tag{1.17}$$

$$\dot{p}_o = \beta_o (1-e-p_o)(p_o-e) \left(\frac{\partial w_s}{\partial p_o} \right) \tag{1.18}$$

$$\dot{p}_1 = \beta_1 (1-e-p_1)(p_1-e) \left(\frac{\partial w_s^v}{\partial p_1} \right) \tag{1.19}$$

$$\dot{p}_2 = \beta_2 (1-e-p_3)(p_2-e) \left(\frac{\partial w_o^s}{\partial p_2} \right) \tag{1.20}$$

$$\dot{p}_3 = \beta_3 (1-e-p_3)(p_3-e) \left(\frac{\partial w_o}{\partial p_3} \right) \tag{1.21}$$

We assume that individual learning enables p_i to adjust more rapidly than does x (which adjusts via entry and exit, or type switching), and that v adjusts least rapidly (via genetic disposition and/or internalized norms). Thus $0 < \beta_v < \beta_x < \beta_0 = \beta_1 = \beta_2 = \beta_3$. To complete the dynamic specification, take the initial state as given and impose the boundary conditions $0 \leq x \leq 1, v \geq 0$ and $e \leq p_i \leq 1-e$.

1.3.4 Dynamic behavior

Which PBE remain when x and v_H can adjust? To answer, FS09 proposes a static refinement called evolutionary perfect Bayesian equilibrium (EPBE). In EPBE, all types in the support of the distribution in each population achieve equal and maximal expected fitness, and no potential entrant (a type outside the support) has higher expected payoff.

Proposition 2 of FS09 shows that only two states survive the EPBE refinement: (i) a unique "Good Hybrid" EPBE in which Self trusts regardless of her type and Other plays a specific mixed strategy when she perceives a non-vengeful type, and (ii) the extreme "Bad Pooling" EPBE in which (apart from trembles) Self never trusts and Other always defects. The good EPBE has state vector

$$S = (x, v, p_o, p_1, p_2, p_3) = (x^*, v^*, 1-e, 1-e, 1-e, p_3^*),$$

where x^*, v^*, p^* are uniquely determined by the exogenous parameters, and the bad EPBE has state vector $(0, v, e, e, e, e)$, where v is arbitrary (and moot, since the vengeful type has population share zero).

Assuming the baseline parameters $k = 0.2$, $c = 0.8$, $e = 0.1$, $\beta_v = 0.01$, $\beta_x = 0.10$ and $\beta_{1,2,3} = 2$, the "Good Hybrid" is $(0.77, 2.78, 0.90, 0.90, 0.90, 0.68)$. Using standard numerical software, we can compute the eigenvalues of Jacobian matrix evaluated at that point. Of the 6 eigenvalues, 3 are negative, 1 is zero and 2 are pure imaginary. Thus we surmise that the good EPBE typically is neutrally stable. To investigate more carefully, we turn to numerical simulations.

1.3.5 Simulation results

Panel A figure 1.5 shows typical numerical solutions for baseline parameters and initial conditions not far from the EPBE. The state indeed cycles around the "good" EPBE with constant amplitude, consistent with Liouville's theorem. Panel B confirms convergence in time average.

Numerical simulations indicate that the good EPBE is indeed locally stable for all parameters values within the set $c \in (0, 1)$, $e \in (0, \hat{e}(k))$and $k \in (0, 0.6)$.[5] In other words, both

[5] See the Appendix "Proof of Proposition 2 and Comparative Statics" in FS09 for the definition of $\hat{e}(k)$

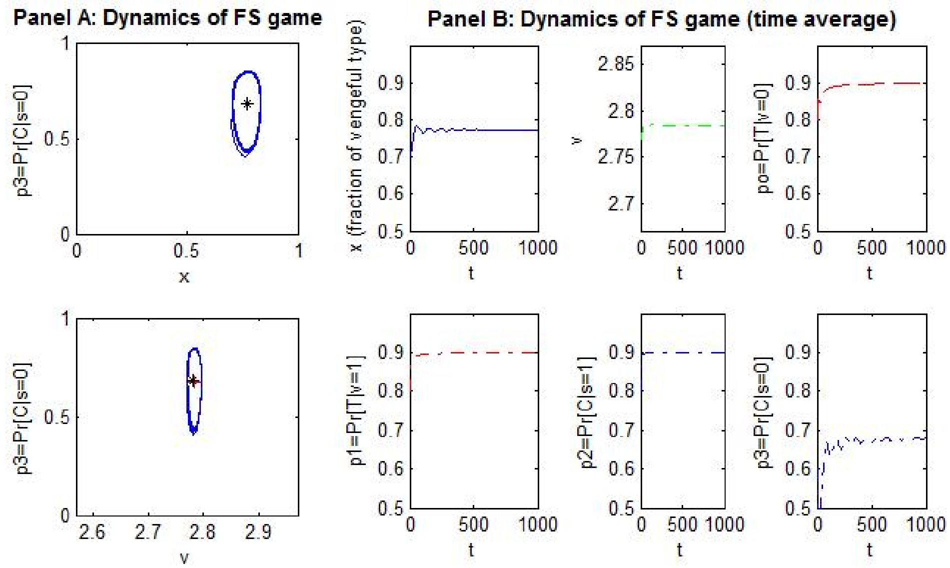


Figure 1.5: Dynamics (Panel A) and Time-Average dynamics (Panel B) of the FS09 game. In Panel A, the "Good" EPBE is indicated by * and the time average by $+$.

types of Self survive in the long run and both trust maximally, while Others' action depends on her perception. If Other perceives a vengeful type she will cooperate with maximal probability $1-e$ and if she perceives a non-vengeful type she will play a specific interior mixed strategy.

There are two caveats. First, we are talking about local stability, so we do not drastically alter the initial state. For baseline parameters we have confirmed convergence from initial states $v^*-0.04 \leq v(0) \leq v^*+0.01$, $x^*-0.2 \leq x(0) \leq x^*+0.15$, and $0.45 \leq p_i(0) \leq 1-e$. Second, we must restrict the adjustment speeds appropriately: $\beta_v < \beta_x$. This restriction is consistent with the idea from FS09 that slow cultural or genetic adjustment controls v, while exit and entry control x.

The "Bad" EPBE is at the corner of the state space, where the mixing probabilities are at the lower bound e and the fraction of vengeful type x goes to zero. Liouville's theorem does not preclude direct convergence to a corner equilibrium. Indeed, we find direct convergence to the "Bad" EPBE given initial values of $x < 0.2$ (not many vengeful types) and $p_2 < 0.2$ (a low probability that Other cooperates).

1.4 Discussion

We have analyzed the dynamic stability of two games of incomplete information in the context of preferences evolution. We complement Arce (2007) results by endogenously determining the fraction of Agent types and studying the dynamics of the state variables. We show that both types of Agents coexist independently of the level of incentive wage. The second example is a noisy trust game due to FS09. In this game, we add dynamics to their static EPBE concept and the numerical results show convergence in time average to the key equilibrium (in which Self trusts regardless of her type and Other cooperates if she perceives a vengeful type and plays a specific mixed strategy if she observes a non-vengeful type).

Perhaps the main contribution of the present paper is to illustrate a toolbox for investigating the dynamic stability of equilibrium in a wide class of games of incomplete information. The toolbox first asks the researcher to write down the expected payoffs and expected utilities at all feasible states. Then it applies standard evolutionary concepts to describe the evolution of types (preference parameters in our examples), their population shares, and the action mixtures. It uses gradient dynamics for a continuous space of types, and uses hybrid gradient-replicator dynamics for the rest of the state vector, the population shares and mixture probabilities. This approach will yield systems of ordinary differential equations (ODEs) in applications like those just analyzed, but it can yield partial differential equations (PDEs) when a continuum of active types is possible.

In view of the fact that asymptotic stability can not be expected in key equilibria of games of incomplete information, the toolbox emphasizes convergence in time average and numerical methods. One can check robustness by sampling the economically feasible parameter space. Our toolbox also calls for appropriate restrictions on the adjustment speed parameters. For instance, in the FS09 game, slow cultural or genetic adjustment controls the type variable (the preference parameter v), while the exit and entry allow population shares to adjust at a moderate rate and individual learning allows very rapid adjustment of action mixtures.

We close with two more philosophical remarks. The toolbox presented here did not include some of the more advanced techniques from dynamical systems theory, such as center manifold techniques or bifurcation techniques. In our experience so far, the return on applied researchers' investment seems much higher for the simple numerical methods we emphasize.

Second, in specifying a game of incomplete information, the set of active types and the population shares of those types is exogenously given in most mainstream analyses. That seems to us to push off stage the most interesting part of the story. Hence our toolbox features methods for describing the evolution of these state variables and for characterizing their long-run behavior.

Chapter 2

Strategic default with neighbors: A laboratory experiment

2.1 Introduction

In the United States, following the collapse of home prices in 2007-08, millions of households found themselves underwater. Assets, such as real estate, generally represent a substantial portion of household wealth. Therefore, strategic default can be especially relevant in an attempt to protect household assets. When a sufficient number of households begin to see a decline in their real wealth, the effect on the economy can become quite strong due to changing consumption and investment levels. Given the importance of real estate to household wealth and to overall economic health, it would be of interest to study the driving forces of strategic default. Greater understanding of the causes and consequences of strategic default would allow for more appropriate response and thus more efficient policy-making. In this paper, I hope to shed more light on the factors driving this decision.

Past literature used field data to examine, with limited success, the causes of the strategic default. I propose a rather different approach, an experimental study, that allows us to observe how human subjects react to two important variables that are thought to drive the strategic default but generally cannot be observed through the field data: (i) the value of the deferral option and (ii) the endogenous impact of neighbors.

Deferral options are defined as the right but not obligation, to take action at a pre-determined cost, for a pre-determined period of time. In the context of the mortgage market, a household or an agent has the choice to pay the financial obligation or to default. The optimal policy is to exercise the deferral option —default— when the asset price (house) crosses a threshold whose value depends on (among other things) the asset price volatility. In the laboratory, the experimenter knows the value of the deferral option after choosing the asset price process and loan characteristics. Empirical studies, on the other hand, must rely on crude estimates of these variables. The experimental approach may then be more appropriate to study the intrinsic value of the deferral option in strategic default.

Additionally, researchers have recently began to emphasize the importance of social interactions in strategic default. Field work has found that neighborhoods with high foreclosure rates are more likely to fall behind on their mortgages. However, it is not clear whether this is due to endogenous or correlated effects. In other words, higher foreclosures can be driven by common shocks —correlated effects— within a neighborhood, creating notorious identification problems pointed out by Manski and others. Given these difficulties, a laboratory experiment once again appears to be a useful approach for studying social interactions.

Keeping in mind the aforementioned factors, I designed an experiment that characterizes the strategic default. Human subjects are endowed with an asset that has been acquired with a mortgage. The asset price follows a stochastic process and the subject has the option to default at any point in time. The experimental design is two-by-two. The first treatment variable is asset price volatility —low and high— and the second treatment variable is the presence of the neighbor's effect —independent and dependent. The experiment involves 118 human subjects and each subject is assigned to one of the four different treatment combinations.

The main results are: (i) I demonstrate that people appear to follow the prediction of the strategic default model quite closely, and (ii) I find that incorporating social interactions delays the strategic default beyond what is considered optimal.

The organization of the chapter is as follows: section 2.2 elaborates on the contribution of the paper in the context of existing literature. Section 2.3 lays the theoretical foundation. After describing the basic assumptions, it presents the deferral option model followed by a numerical solution specifying the asset price that triggers default. Section 2.5 contains the testable predictions based on the conjecture that actual behavior will approximate optimal behavior, i.e. subjects will defer their decision until they are sufficiently underwater. Section 2.6 presents the results and section 2.7 concludes with a discussion of results and suggests possible venues of future research.

2.2 Related literature

Recent literature has focused on two hypotheses to explain the factors that drive the decision to default. The first hypothesis, known as a “ruthless” or “strategic default”, is that default occurs when a borrower’s equity falls below the threshold where the cost of paying back the mortgage (liability) just outweighs the benefits of holding on to the asset. Using financial terms, we can refer to this decision as a put option. The borrower decides to give up the asset when the spot asset price is sufficiently lower than the strike price, known as the mortgage obligation.[1] Alternatively, the option to default can be interpreted as a deferral option. Dixit and Pindyck (1994) further elaborate on this interpretation by analyzing investment decisions —call options. In this framework, the agent makes an irreversible investment decision once the project values surpasses the total cost, which includes the material fixed cost plus deferral option value. Deferring the decision has a positive value because exercising the option today implies that it cannot be exercised later.

The second hypothesis complements the first one and is known as the “double trigger” hypothesis. Default occurs following a negative income shock or other unhappy event, such as divorce. Recently, some researchers have emphasized that there may be other non-monetary variables that affect the decision of “walking-away”. I categorize them as part of this hypothesis since they represent additional social or moral costs of default. For example, White (2009) argues that the media and political institutions shape the opinion regarding default. Furthermore,

[1]Default, as part of the early termination of the mortgage contract, affects the valuation of the mortgage. In general, early termination of the mortgage is due to two factors: prepayment and default. The former is treated as an American call option and the latter as an European compound put option. See Kau, Keenan and Taewon (1994) for the valuation of mortgages and Vandell (1995) for a early review on the literature of mortgage default.

Guiso, Sapienza and Zingales (2009) found that 81 percent of respondents answered positively when asked if they think that is morally wrong to walk away from a house when one can afford to pay the monthly mortgage.

Empirical studies have tested both of these explanations by estimating the hazard function, which is defined as the probability of default at a particular time conditional on survival, i.e. that the household has not yet defaulted on the mortgage. The hazard function includes variables that capture the idea behind the strategic default hypothesis, such as loan-to-value (LTV) and another set of variables related to the double trigger hypothesis; e.g. unemployment, credit utilization or credit scores.[2] Notice that testing the strategic behavior by considering the LTV variable does not provide exact information about how strategic the agents are since the option value is unobservable. In an experimental environment, it is feasible to measure the option value and therefore adequately test the financial predictions.

Another recent study by Bhutta, Dokko and Shan (2010) examined the level of equity required (assets minus liabilities) to trigger default, after controlling for liquidity issues. Working with non-prime first-lien home purchase mortgages[3] originating in 2006, with a combined LTV of 100 percent in Arizona, California, Florida and Nevada, the authors find that the median borrower does not strategically default until the equity-to-value ratio is negative 62 percent. As I pointed out earlier, the authors cannot conclude how strategic the agents are since the rate of

[2]There are several empirical papers that deal with the impact of equity and financial variables in the default or prepayment decisions. Depending on the time of analysis, authors consider these two decisions simultaneously by estimating hazard functions with competing risks. See Deng, Quigley and Order (2000), Foote, Gerardi and Willen (2008), Bajari, Chu and Park (2008), Pennington-Cross and Ho (2010) and Gerardi, Shapiro and Willen (2007).

[3]In the financial jargon, the terms non-prime and subprime mortgage loan are used indifferently. *A subprime mortgage loan is a residential mortgage loan that is particularly risky. The elevated risk may stem from the credit history of the borrower, the lack of a large down payment, or a monthly payment that is large relative to the borrower's income* (Foote and Willen 2011).

default includes transactions costs, option value, etc.

Besides the traditional hypotheses discussed, a number of studies have also emphasized contagion effects. These effects can be present via social norms or material incentives. Chan, et. al. (2010) find that mortgage holders, living in the New York City area with high foreclosure rates and high real state owned (REO) activity, have a substantially greater risk of falling behind on the mortgage, as do mortgage holders living in predominantly black neighborhoods. Additionally, [**?**] show that foreclosure at a distance of 0.05 miles lowers the price of a house by approximately one percent in the state of Massachusetts. Using survey data, Guiso, Sapienza and Zingales (2009) find that people who know someone who defaulted strategically are 82 percent more likely to declare their intention to do so.

Identifying endogenous social interaction effects in the field data can be difficult (or impossible). Applying the work of Manski (1993) to mortgage default in neighborhoods, there may be correlated effects that need to be considered. For instance, neighborhoods will have higher default rates due to a common aggregate shock to the economy rather than endogenous interactions within. There are also contextual interactions at play when default rates vary with the socioenomic composition of neighborhoods. Therefore, when empirical studies show that neighborhoods with high foreclosure rates are more likely to fall behind on their mortgages, it is not necessarily clear whether this is due to endogenous or correlated effects. Given these difficulties, laboratory experiments appear to be a perfect environment for studying social interactions.

This paper is not the first to study deferral options in the laboratory. Recent studies have focused on testing call options inspired by Dixit and Pindyck (1994) . In this framework, the project value follows a random process and the subject incurs a fixed cost to seize the irreversible investment opportunity. List and Haigh (2010) work with a two-period model using a simple design. A set of contracts is offered to the subjects, and each contract specifies two alternatives: invest today or tomorrow. Results indicate that undergraduate subjects, as well as professional traders, choose the correct alternatives, according to the theoretical predictions of the option model.

In a different experiment, Oprea, Friedman and Anderson (2009) design a stochastic model with three different option values (high, medium and low) using a finite horizon environment. They effectively demonstrate that undergraduate subjects follow the option model in the low treatment. Subjects in other treatments invest at values below optimum, but with predicted ordering. The authors also find evidence of learning behavior when the option value is either low or medium. The present paper has similar features but also incorporates social interactions. The next section describes the environment and the optimal asset price that triggers default.

2.3 A simple theoretical model

I assume that the subject needs an asset to get a certain level of felicity Θ. She starts the game with an asset H that has been acquired with a loan (L_0) that requires periodic payments and whose LTV is equal to one. Alternatively, the subject can stop paying the financial obligations and become a renter. I focus on the borrower's decision and abstain from solvency issues or ability to service the debt. The objective is to identify the timing of the subject's decision to switch from being an owner to become a renter.

Figure 2.1 displays the environment that the subject faces. H varies randomly and the present value of liabilities is constant (and equal to L_0) due to the characteristics of the loan: interest rate payments are made in each period with the principal paid at the end of the game. The decision to stop (default) involves a fee that can be interpreted as an upfront lease or a transaction cost (see bar under the horizontal line at the end). The agent receives a reward (B) conditional on remaining an asset owner throughout the game (see bar over the horizontal line at the end). B may depend on the decision of other subjects —assumed to be constant for now.

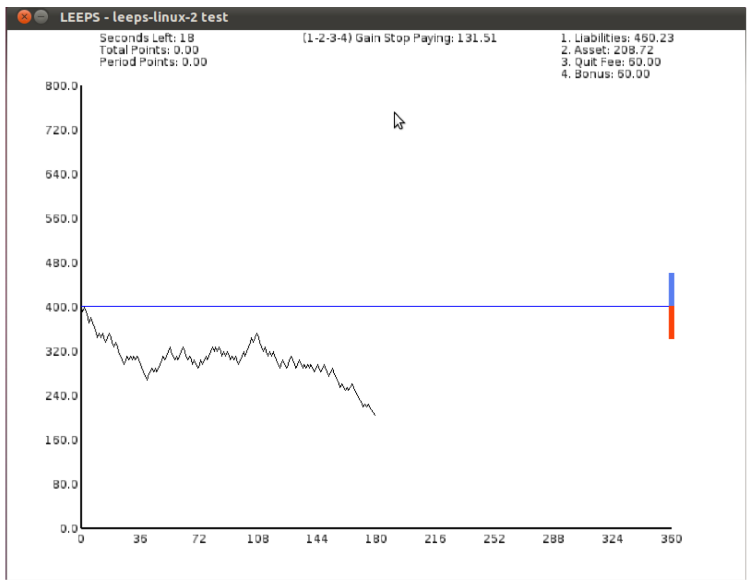


Figure 2.1: The simple model

2.3.1 Payoffs and objective function

The payoffs correspond to the decision: to stop or not to stop. The subject is not allowed to default in the last period. Therefore, the terminal payoff consists of the asset value plus the reward, less the principal value

$$H_T - L_0 + B \tag{2.1}$$

where H_T is the asset price at the end of period, L_0 is the loan principal and B is the reward for paying interest rate obligations.

The present value of the flows that the subject receives as an asset owner is

$$NS(EH,t,L_0,T,r) = \int_0^T (-rL_0 + \Theta)e^{-rt}dt + E_0\left[H_T - L_0 + B\right]e^{-rT} \tag{2.2}$$

where r is the loan interest rate and the risk-free interest rate.

The present value of the flows that the subject receives if she stops paying at time t^*

$$S(EH,t,L_0,T,r|t^*) = \int_0^{t^*} (-rL_0 + \Theta)e^{-rt}dt + \int_{t^*}^T (\Theta - c(t))e^{-rt}dt \tag{2.3}$$

where $c(t)$ is the cost of stopping and becoming a renter. Subtracting (2.3)-(2.2), we obtain $E\pi$, the expected payoff the subject will maximize

$$\max_{\{t^*\}} E\pi(EH,t,L_0,T,r|t^*) = \int_{t^*}^T (rL_0 - c(t))e^{-rt}dt + (L_0 - B - E_{t^*}\left[H_T\right])e^{-rT}$$

Alternatively, the problem can be rewritten as a binary choice. One alternative corresponds to stopping and taking the termination payoff, while the other entails continuation to the next period, where another binary decision will be present. In this setup, the subject will pay a cost C —present value of $c(t)$— to stop and become a renter or the subject will pay the financial debt to remain an owner.

$$rF(H,t) = \max\left\{-rC, -rL_0 + \frac{1}{dt}E[dF]\right\} \tag{2.4}$$

The agent will maximize the value of $F(H,t)$ subject to the end of period condition, the asset price dynamic and the initial conditions. The asset follows a standard Brownian motion where α is the asset price growth and σ is the asset price volatility.

$$F(H,T) = H_T - L_0 + B$$

$$dH = \alpha H dt + \sigma H dz$$

and

$$r, H_0, B, L_0 \text{ and } C \text{ are given.}$$

2.3.2 Neighbors effect

In the presence of social interactions or neighbors effect, the reward B received at the end of the period depends on the number of neighbors that stop paying (n). I assume that the higher the n, the lower the B. This assumption has a number of interpretations. For example, it can be interpreted as the impact of foreclosure on the sale price or the benefit of being a home-owner.

The agent will maximize a problem similar to (2.4); however, B will be specified as follows

$$B = \begin{cases} 0 & \text{if } n \geq 1 \\ \bar{B} & \text{if } n = 0 \end{cases}$$

Notice that I also assume that neighbors face the same realization of H_t and they are aware of this fact.

2.4 Numerical solution

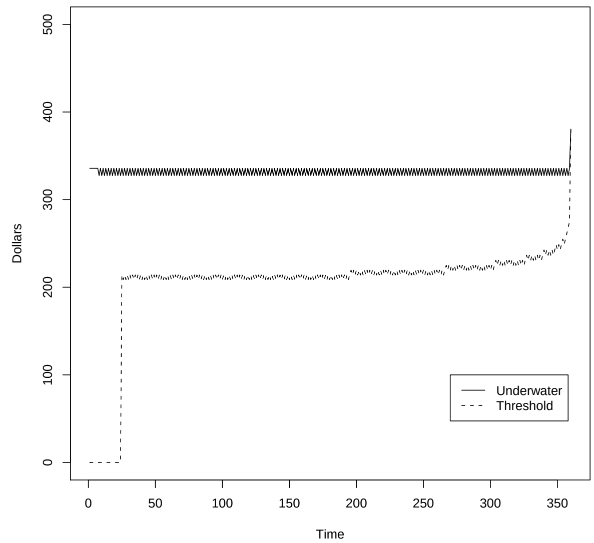


Figure 2.2: Optimal threshold (H^*) and underwater price (H^U)

2.4.1 Optimal default

Equation (2.4) can be solved numerically.[4] The solution begins at the terminal period in the tree and is obtained by solving backwards. At each node, the agent decides to stop paying if the expected early termination payoff is greater than expected future value of continuing. Following Cox, Ross and Rubinstein (1979), the continuous asset process is approximated by a discrete random walk. In this framework, the asset return may take one of two possible values (u, d) with the probability p $(1-p)$ of the asset moving up (down). The parameters are defined

[4]For the numerical solution to financial derivatives see Willmott, Howinson and Dewynne (1999).

as

$$
\begin{aligned}
d &= \frac{1}{u} \\
u &= exp(\sigma\sqrt{\delta t}) \\
p &= (exp(r \cdot \delta t) - d)/(u - d)
\end{aligned}
$$

where δt is the time step. In the discretization, the end of period payoff is

$$F_T^m = EH_T^m - L_0 + B$$

where F_T^m denotes the m-th possible value at time step T, EH_T^m is the expected asset price, L_0 is the principal payment and B is the reward. The optimal value is the maximum between the two possibilities: the early termination payoff and the expected value in the next period.

$$F_n^m = \max(-C, e^{-r\delta t}(-rL_0 + pF_{n+1}^{m+1} + (1-p)F_{n+1}^m)),\ n = 0, 1, \ldots, T$$

Denote H^* as the threshold price, defined as the asset price at which the agent is indifferent between paying and stopping. The asset values above H^* will induce the agent to stay in the game. It is also convenient to define H^U as the underwater price —asset price is equal to the current liabilities minus the cost of stopping. Notice that H^U and H^* are not necessarily equal over time. The threshold price is usually smaller than the underwater price $H^* < H^U$ because the agent places a value on the foregone opportunity of stopping later. The difference between them depends positively on the volatility of the asset price.

Using the following parameters, $T = 360$, $H_o = 400$, $LTV = 100\%$, $B = C = 60$, $r = 0.0006$ and $\sigma = 0.025$, H^U and H^* are depicted in Figure 2.2. The loan interest rate and the discount rate are equal and therefore the underwater line is constant. As H^* approaches H^U the value of waiting decreases.

2.4.2 Simulations

This section provides simulations that illustrate the cost of deviation from the optimal behavior $H^*(t)$. I simulate players that decide to stop paying as soon as the random realization of the asset value crosses the price $H^S(t)$, which is computed as follows

$$H^S(t) = aH^U(t) + (1-a)H^*(t) \text{ where } 0 \leq a \leq 1$$

Alternatively, the simulated players can stop at lower asset values such that

$$H^S(t) = bH^*(t) \text{ where } 0 \leq b \leq 1$$

In other words, the first group of players use a threshold that is a convex combination of H^U and $H^*(t)$. When $a = 1$ the subject decides to stop paying as soon as the random realization of the asset value crosses the underwater price. The value of $a = 0$ corresponds to the optimal strategy. The players $0 \leq a < 1$ can be interpreted as impatient, they decide to stop early. The remainder of players wait longer. In the extreme case $b = 0$, the player does not stop paying.

Figure 2.3 shows the quantile payoffs after 5,000 simulations with parameters similar to the previous figure but with varying volatility ($\sigma_{low} = 0.025$ and $\sigma_{high} = 0.04$). There are a few points worth mentioning before moving on. First, notice the asymmetry in the payoff distribution. The asymmetry is sharper in the high volatility treatment. Second, the impact of substantial departures from the optimal threshold (deviations in some directions, but not all) is

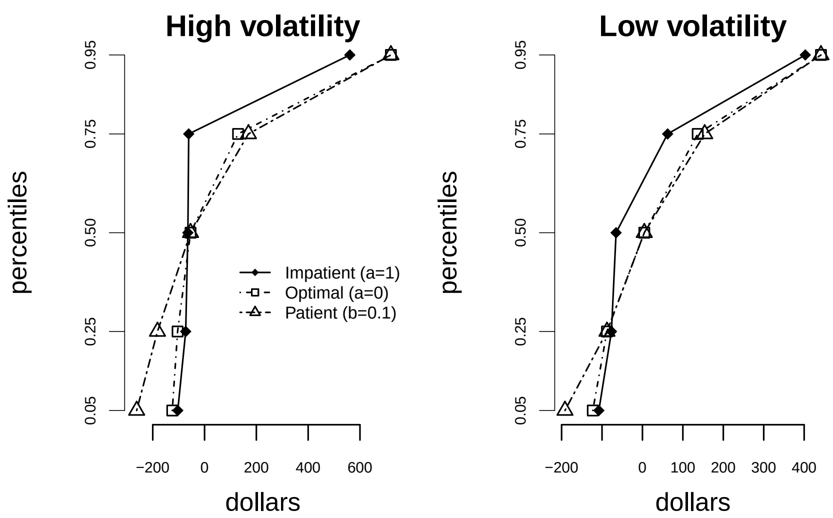


Figure 2.3: Quantile payoffs for the simulated players

not detectable. Under this environment, the subject may wait longer than optimal. Third, it takes time to observe optimal behavior. It is necessary to run a large number of simulations to notice a payoff differential.

2.5 Hypotheses and experimental design

Before presenting the hypotheses, I will first define the treatments used in the experimental design. There are four treatments in total (2×2 design). The first pair corresponds to volatilities (low and high) and the second pair, to the presence of neighbor's effect (dependent and independent).

2.5.1 Hypotheses

Let's denote $h^{j}_{\tau,v}$ as the subject's stopping asset price at time τ where j refers to the treatment with dependent (DEP) or independent (IND) bonus and v refers to the volatility environment: low or high.

Hypothesis 1. *Volatilities order: observed stopping price has the same ordinal rank across treatments as does the theoretical price. Therefore,*

$$h^{IND}_{\tau,high} < h^{IND}_{\tau,low}$$

and

$$h^{DEP}_{\tau,high} < h^{DEP}_{\tau,low}$$

This hypothesis is drawn from the theoretical model. The higher the volatility of the asset price, the lower the asset price that triggers default. In other words, it is optimal to wait longer to default.

Hypothesis 2. *Point prediction: for each treatment, the observed stopping price is equal to the theoretical stopping price.*

This hypothesis focuses on the price level that triggers default. It assumes that the subject stops accordingly to the optimal solution of the model.

Hypothesis 3. *Neighbor's effect: observed stopping price in the independent treatment is equal to or lower than the observed stopping price in the dependent treatment. Therefore,*

$$h^{IND}_{\tau,high} \leq h^{DEP}_{\tau,high}$$

and

$$h^{IND}_{\tau,low} \leq h^{DEP}_{\tau,low}$$

In this case, I allow deviations from the optimal behavior. If a subject stops under the dependent treatment, it is also more likely that her neighbor stops due to the absence of the bonus.

2.5.2 Implementation

Figure 2.1 shows the screen observed by each subject. At the beginning of each period, there are 70 in total, each subject is endowed with a base payment and owns an asset with value H that has been acquired with an LTV equal to 100 percent. The asset evolves according to the discrete binomial approximation of the Brownian motion described above. The liabilities are constant (depicted with a horizontal line) and equal to the initial value of H. Throughout the game, the subject is able to observe whether the bonus is present (bar above the horizontal line) as well as see the value of early termination fee (bar below the horizontal line). The subjects can stop the payments by pressing the space bar. In this case, the payoff will be equal to the base points minus the sum of the quit fee and the interest payments. Despite stopping early, the subject will still be able to observe the black H-line evolving until the end of the period.

The payoffs in the case that subject does not stop are equal to asset value plus the bonus —conditional on the treatment group— minus the financial payments (principal and interests). The payoffs are depicted with a green bar.[5]

[5]Appendix A includes examples of the screen observed by the subject when she stops or does not stop.

In addition to the graphical display, the experimental software shows the numerical values of the gain from stopping, the payoff received in the current period and the cumulative earnings. The gain is computed as the difference between the liabilities and the sum of asset, quit fee and bonus.

The baseline parameters are $T = 360$ (36 seconds), $H_0 = 400,\ r = 0.0006$ and $C = 60 = B$. I work with $\sigma_{low} = 2.5\%$ and $\sigma_{high} = 4.0\%$ that correspond to $(p = .506;\ u = 1.025)$ and $(p = .498;\ u = 1.04)$, respectively.

The subjects participating were 118 undergraduate students at the University of California, Santa Cruz. At the beginning of the experiment, each subject was seated at a visually isolated computer terminal and assigned to a treatment (e.g. low volatility and dependent bonus). Instructions were read aloud and the software was displayed on a screen. The binomial parameters for the chosen treatment were explained and written on a whiteboard. Subjects participated in four practice periods. Each subject then participated in 70 paid periods with no change in treatment. Sessions lasted 80 minutes each.

A total of 34 subjects participated in the low-dependent treatment, 36 in the low-independent treatment, 34 in the high-dependent treatment and 14 in the high-independent treatment. No subject was allowed to participated in more than one session. A subject with cumulative payoff over all periods received cash at the end of the session. Subjects also received a \$ 5 bonus show-up fee. On average subjects received \$ 16.25 in low volatility and \$15.76 in high volatility.

2.6 Results

Table **??** summarizes the experimental results. It provides the following information for the four treatments: the number of subjects, the number of observations is equivalent to the total number of periods (it can also be interpreted as the total number of loans) and the percentage of periods in which subjects stop (default). The sample is divided into three categories: (i) all periods, (ii) periods in which the minimum price is lower than the optimal threshold and (iii) periods where minimum price is lower than 60 percent of the optimal threshold. The optimal threshold is provided by the numerical solution to equation (2.4).

Looking at the complete sample, it is worth noting the low default rate in all treatments. The default rates are under 40 percent and can be explained by the following: (i) default is more likely to be observed when subjects choose a high price threshold, and therefore, the observed decision constitutes an upwardly biased sample, and (ii) given that asset realization is random, in a number of situations the asset value does not cross the underwater line. For these reasons, I decided to work with a subsample of data where the asset crossed 60 percent of the optimal threshold value.[6]

Working with the subsample data, we can clearly observe a higher default rate in the high volatility treatments compared to the low volatility treatments. In particular, the high independent treatment displays a default rate that is greater than 90 percent. Furthermore, the default rate is lower when the social interactions are incorporated. The rate is roughly doubled when the subjects make their decision without the presence of the neighbor's effect.

An alternative way to illustrate the experimental results is to use the cumulative dis-

[6]In the Appendix, I present more subsamples.

Table 2.1: Summary: default decision

	Low Dependent	Low Independent	High Dependent	High Independent
		Sample: All		
Subjects	34	36	34	14
% Default	22	23	24	37
Observations	2380	2198	2380	968
		Sample: min. price < threshold		
Subjects	34	36	34	14
% Default	25	41	39	79
Observations	634	537	578	234
		Sample: min. price < 60% of threshold		
Subjects	34	26	34	14
% Default	32	53	42	94
Observations	34	51	240	62

Note: Obs. represent the total number of loans. Threshold refers to the theoretical prediction of the default asset value.

tribution of default. I work with the Product Limit (PL) estimator, which is used to establish the hazard distribution of default.[7]

Figure 2.4 shows the PL estimator for each treatment using pooled data. In this case, the graph should be read from right to left. Recall that the initial loan value is equal to 400. As the price decreases, it is natural to observe more defaults. However, as mentioned earlier, for the low and high-dependent treatments we do not observe a significant rate of default. Notice that for the high independent treatment, the median default price is close to the theoretical prediction. However in the other treatments, the median price is much lower than the prediction or it cannot be computed due to the severe left-censoring.

To control for the within-subject dependence, I also work with a "by-subject" sample. For the treatment in which the bonus is not affected by neighbor's decision, the observations are

[7]The PL estimator is an appropriate method for censoring data. See Kaplan and Meier (1958).

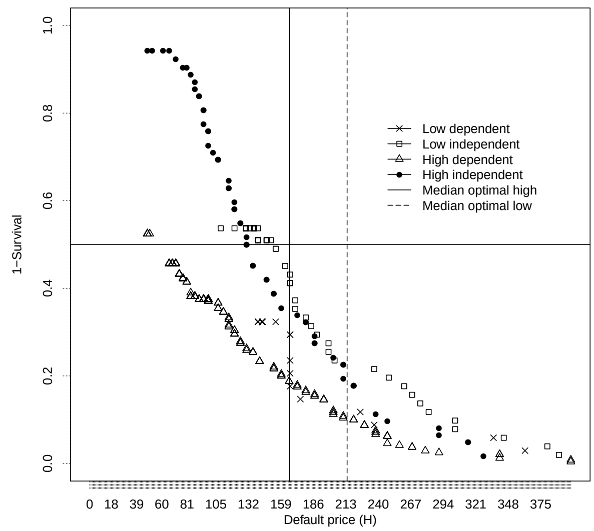


Figure 2.4: Default prices –PL estimator for subsample (pooled data)

completely independent. For the dependent treatment, however, this does not hold. Therefore, I treat the random neighbor-pairs in each session as a single observation. It is appropriate to work with the "by-subject" sample since the pooled data over-samples the subjects that choose a higher threshold value to default — it is more likely to observe their default decision. Table 2.2 presents the means and standard deviations of default prices for the high dependent and high independent treatments. It does not consider the low volatility treatments, due to severe left-censoring. We can see that the mean default price is close to the prediction for the high independent treatment.

As a complement to the non-parametric analysis, I estimate a Tobit regression using the subsample pooled data (Table 2.3, using clustered standard errors by subject for the independent treatment and by group for the dependent treatment. The dependent variable is the

Table 2.2: Default prices: Mean and standard deviation

	High Dependent	High Independent
Pooled	105.2 3.0	130.4 5.3
By subject	109.8 9.9	150.7 13.4
Median prediction	166	166

default price and the regressors are dummies that capture the treatment effects. The estimation shows that the only significant treatment effect observed is the presence of neighbor's effect. The lack of statistical significance of the asset volatility treatment could be explained by the severe left censoring. This behavior can be explained along with the numerical simulations that demonstrate that waiting longer to default does not imply significantly lower payoffs.

Table 2.3: Parametric estimation (Tobit model)

Sample: subsample (387 obs)
Dependent variable: Default price (h)

Coefficient	Value (s.d)
Constant	65.97***
	(6.33)
Low	28.73
	(11.90)
Independent	80.82***
	(14.12)
Low*Indep	-23.92
	(21.50)

Clustered standard errors by subject and group considered
*** Significant at 5% significance

Below, I present the non-parametric tests of the point predictions and neighbor's effect, summarize the parametric results and follow with a brief discussion.

I test the point prediction using the Wilcoxon T-test on the "by-subject" sample, I fail to reject that the median stopping price is equal to the theoretical prediction for the high independent treatment at 5 percent significant level. For the other treatments, the observed stopping price lie well below the predicted price.

Result 1. *Point predictions: The data supports only part of the point predictions. In the high independent treatment, the observed default price center near the predicted level. In other treatments, the observed stopping prices are significantly lower than predicted.*

I also use the "by-subject" sample to test the hypothesis of equal distributions of the high dependent and independent treatment using the Mann Whitney. The null hypothesis is rejected at 5 percent significance level. Likewise, using the parametric regression, the treatment effect of dependent/independent is relevant with a positive coefficient at a 5 percent significance. Therefore, the subjects in the dependent treatment, thus experiencing the neighbor's effect, wait longer to default.

Result 2. *Neighbor's effect: Subjects delay significantly their decision to default in the neighbor (dependent) treatment.*

2.7 Discussion

In this paper, I have analyzed the default decision in a mortgage using an experimental approach and found that the subjects act according to the theoretical predictions under a high asset price volatility and lack of social interactions. Adding the social interactions via a direct effect on the subject's payoff leads to a delay in the strategic default. This behavior persists even when the asset price goes sufficiently low.

The theoretical predictions are drawn from a simple model that captures the subject's decision between holding an asset acquired with a mortgage and defaulting. The asset follows a stochastic process and the subject receives its value at the end of the game, in addition to a bonus for paying interest. When the subjects defaults, she loses the asset, pays a quit fee and is then free from any future financial obligations. Under the neighbor's effect treatment, social interactions are tied to the bonus.

A possible explanation of the low default rate under social interactions is that subjects want to avoid harming their neighbors. Moreover, the stigma costs could become more relevant under this treatment even tough there is a blind pair-match between subjects.

I conjecture that my results can be applied to policies that seek to diminish the number of foreclosures. For instance, my results suggest the importance of identifying high-risk households and offering them timely, affordable alternatives to pay their financial obligations. This could help maintain community cohesiveness which has a significant role in mitigating the overall foreclosure rate.

More work is needed to fully distinguish the factors that drive default. In particular, little is known about the real costs of default. Having a better approximation of this variable, would allow us to better understand how the default option is exercised. If this cost is indeed small, then it must be social norms that are driving the willingness to wait longer to default.

Chapter 3

Cooperative behavior in one-shot games with assortative matching (with Olga A. Rabanal)

3.1 Introduction

When a domestic firm decides to form a partnership or joint venture with a foreign entity, it can never be certain of the outcome of a resulting match due to pre-commiment problems. The efficient outcome is determined by the investment levels from both sides. However, there are incentives to shirk and "free-ride" from the decisions of the counter-party. These scenarios actually resemble a one-shot Prisoners' Dilemma (PD) type game where the Nash Equilibrium (NE) is not socially optimal. Therefore, in order to reach a more optimal or mutually beneficial solution (a Pareto optimal outcome), it may necessary to include proper incentives that could help overcome these inefficiencies; e.g. communication, reputation, repetition and punishment.

In this paper, we introduce a mechanism that may enhance cooperation through group formation, or assortative matching, thus leading to a Pareto optimal outcome. The mechanism could be implemented as a middleman whose task is to match parties according to type. Following our joint-venture example, the middleman is a third party with certain knowledge about both sides –investment levels, and who is able to match firms based on type (we can interpret investment level as firm type –high and low). This ability to unite similar types can become especially important in markets where gathering information about other participants is costly.

Our results indicate that under assortative matching we can elicit highly cooperative behavior and thus achieve a pareto superior outcome as compared to the control group results. Furthermore, in line with other experimental evidence, the results suggest a heterogeneity of types that is quite persistent. According to related experimental evidence, there may be (at least) two types of behavior: (i) payoff maximizing (free-riding) and (ii) conditional cooperation. Remarkably, our results show high concentration of strategies near the upper bound implying cooperation as the dominant strategy for the majority of the participants.

The main result of our paper highlights that cooperation is indeed possible in one-shot PD under proper incentives. The rest of the paper is organized as follows: section 3.2 provides an overview of related literature, while section 3.3 describes the proposed game in further detail. In section3.4 the experimental design is presented, followed by our hypotheses. Lastly, section 3.5 provides a brief discussion of results.

3.2 Related literature

Numerous studies have analyzed the stability of cooperative equilibria in Prisoners' Dilemma (PD) type games. Generally, the findings suggest that cooperation is not fully sustainable unless the game is infinitely repeated or otherwise incorporates incentive compatible features that induce cooperation on part of the players (Cooper et al. 1996). In this paper, we hope to introduce a mechanism that enhances cooperation through group formation, or assortative matching, thus leading a Pareto optimal outcome.

To begin with, we would like to consider past literature on cooperative play. For example, Kreps et al. (1982) suggests that a small belief that an opponent will cooperate is enough to support the notion of cooperative play despite the presumption of self-interest on the part of the players —conditional cooperation.

In 1996, Cooper et al. decided to test two leading theories of cooperative play: altruism and reputation building. In particular, they looked to PD games to analyze the outcomes of one-shot and finitely repeated games. Their results suggest that while reputation models are unable to fully explain observed cooperation in one-shot games (above 20 % over 400 observations), the presence of altruists without reputation building is insufficient to explain the higher cooperation rates in finitely repeated games. Clearly, when dealing with one shot games, reputation cannot be a factor, yet according to these results, the alternative explanation of altruism also appears to be insufficient.

Using a different approach, Friedman and Oprea (2012) study a continuous PD, while also examining cooperation levels in one-shot specification. Their results suggest that mean mutual contributions are nearly zero following the 12th period (out of 32 total) in one-shot games. Therefore, given these past findings, we can safely conclude that cooperation remains low.

As mentioned earlier, our ability to enhance cooperation among players and thus achieve a Pareto superior outcome may depend on group formation strategy. Chaudhuri (2010) and Charness and Kuhn (2011) provide an extensive survey regarding group formation in public-goods games. Given that in our game, the cooperation takes the form of contribution (or effort/wage in the gift-exchange literature), we look to public goods literature as a guide for has been studied to promote such behaviour.[1]

In effect, group formation can take either exogenous or endogenous form. In the first case, the experimenter applies a pre-determined rule which may or may not be known to the participants, while in the second case the participants are able to form or leave groups at their own will. The motivation behind group formation is to mitigate the free-rider problem commonly found in public goods games, and to unite similar player types. Experimental evidence shows that in the presence of conditional cooperators, contributions to the public good can be sustained (but rarely achieve the social optimum).

Ehrhart and Keser (1999) were among the first to consider a public-goods game with endogenous group-formation. In their setup, nine participants were randomly placed into three initial groups. Each person was then told the sizes and average contributions for each group, and could unilaterally decide, at a fixed cost, to switch groups or to form a new (one person)

[1] In the standard public good games, each member of a group, i, voluntarily contributes m_i 'tokens' out of her endowment, b_i to a common account. Each member's payoff is then just $(b_i - m_i) + a/N \cdot (\sum_i m_i)$, where N is group size and a is the efficiency gain from public provision.

group. While this approach lead to some improvement, it had its limitations. Without exclusion or an entry restriction, this approach lead to free-riders chasing more cooperative players.

Additionally, a handful of studies, Page, Putterman, and Unel (2005), and Fehr and Gachter (2000, 2002), have looked at punishment and group formation as a way to fix the free-rider problem and achieve a pareto superior outcome. While both can lead to superior outcomes, a mechanism that is based on group formation is less costly, and thus preferable to punishment. Generally, the type of group formation studied has been endogenous, and feedback was closely related to reputation building, which is absent in this paper. The structure of our game is such that player matches are anonymous, and are based strictly on a pre-specified matching algorithm.

In terms of exogenous group formation, Page, Putterman, and Unel (2005) study the interaction of four person groups (initially randomly assigned) over 20 rounds of a public goods game. Following every three periods, there is a regrouping, where each player rates all other participants based on the public information about average contribution. The four people with the highest sum of mutual ratings amongst all groups go on to form the first group. The same procedure determines the second and third groups, with the leftovers forming the last group. The authors find that the average contribution improves from 38 to 70 percent. Similarly, Gunnthorsdottir, Houser, McCabe, (2007) secretly sort people into high and low contribution groups and find similar results, with bifurcations in the contribution level.

While group formation has mainly permeated public goods literature, we argue that it can also be applicable in other domains. The case we present can be more intuitively modeled as a worker-manager or an investor-investee relationship. In the sections to follow, we use a modified PD game that has been applied to gift-exchange games and incorporate a middleman whose job is to "sort" and match subjects. In effect, this middleman represents a variation of exogenous group formation that is pre-specified by a matching rule, which remains constant throughout the game and is common-knowledge to all players.

The treatment group under exogenous formation will be subject to assortative matching, where contributors are ranked and then matched in an ascending order. Thus, those that contribute the most, will be matched with similar types, in effect "rewarding" cooperative behavior. Because the matches are anonymous, this becomes similar to a one-shot PD.

In our control group, the matching rule is random, and therefore independent of strategy. In this case, there is no extra incentive to cooperate. In fact, the results do seem to support the notion that the existence of such "middleman" can lead to a pareto superior solution, when compared to the control group. The next section provides the setup of the game in greater detail.

3.3 The game

The game is played by two populations. Agents from each population are able to simultaneously select their respective strategies prior to being matched with a counterpart (from the other population). In the first environment, a middleman, after observing the players' strategies, matches pairs according to respective contributions. In order to form a match, the middleman must first sort contributions from each population in a descending order. After sorting, the first match is formed by selecting the highest contributing player from each population. Subsequently, the second match is created by selecting the second highest contributors from each population and so on, until all players are matched with a counterpart. In the second environment, the matching process is random and therefore independent of players' contributions.

For a more intuitive approach, we can think of the two populations as home (H) and foreign (F) firms that seek to form a joint venture. Neither side has much information regarding possible partners overseas yet must pre-commit a certain level of investment x_H and x_F, respectively. Both populations, H and F, will have the same number of firms. After simultaneously selecting their investment levels between 0 and 100 ($x_. \in \{0, 100\}$), the firms' profits will be calculated according to the following functions

$$\begin{aligned} \pi_H &= 100 - x_H + 5 \cdot x_F \\ \pi_F &= 100 - x_F + 5 \cdot x_H \end{aligned} \tag{3.1}$$

where π_H and π_F refer to the home and foreign profits, respectively; and x_H and x_F refer to their investment decisions. Notice that the profits are symmetric and depend negatively on own investment and positively on the investment of the other firm. The NE of this game occurs when $(x_H = 0, x_F = 0)$ with payoffs of $(100, 100)$. The pareto optimal solution is, on the other hand, $(x_H = 100, x_F = 100)$ with payoffs of $(500, 500)$.[2]

The firm profits depend on the matching algorithm, assortative or random. Under assortative matching, the investment levels are sorted in a decreasing order and the pairs are formed by matching similar ranks. In other words, if $i = 1, \ldots, n$ represents the rank within the population, then (i_H, i_F) will constitute a pair. Under random matching, the rank is independent of investment levels. Therefore, the latter case is similar to a lottery where a firm's choice

[2]The payoff structure is borrowed from the literature on "bilateral gift exchange games" where the game is played sequentially. Play begins with the manager choosing a wage. The employee then observes this wage and responds by choosing a costly effort level.

cannot affect who the counter-party is.

3.4 Experimental design and hypotheses

3.4.1 Experimental design

The experiment employed UCSC undergraduate students who were split into two groups of eight subjects: treatment and control. Each subject was assigned to one cell. Thus the students were subjected to either a game under assortative matching (treatment) or they played the same game under random matching rule (control). Each session had 30 periods, which were further split into two practice rounds and 28 paying periods, over the course of which earnings were accumulated. Each period approximated 120 seconds, depending on whether how quickly the subject chose to move on to the next round. The subjects were never told the total number of periods, but rather that they would play some indeterminate number of games. In total, 48 subjects were in the treatment group and 16 subjects were assigned to the control group. All subjects were recruited using the ORSEE system.

Each subject faced a screen with the following information, regardless of the group assignment: (i) period number, (ii) cumulative payoffs in present time, (iii) respective payoff functions and (iv) a box to enter the strategy and a button to confirm the selected strategy.*The appendix includes the instructions and the screenshots in z-tree.* After confirming the desired strategy, the subjects were paired up either according to an assortative or a random matching rule that stayed the same throughout the session. Under assortative matching rule, the strategies were matched in descending order according to rank. Therefore, the players who contributed the most were paired together, followed by the next highest and so on. In the case of an odd number of equal strategies, the pairs were formed randomly. Under random matching, the players were matched randomly regardless of strategy.[3]

Furthermore, the subjects received feedback via strategy history displayed on the bottom of the screen. Therefore, during each period they were able to review their own and counterparty past strategies and payoffs prior to making a decision.

3.4.2 Hypotheses

Given that the subjects were split into two groups, one treatment and one control, we expect two outcomes, as detailed below.

Hypothesis 1. *Under random matching, we expect the strategies to approach the lower bound of zero. Therefore, NE* $\in (0,0)$.

This is the predicted Nash Equilibrium. At each node in the sequence, when deciding on the strategy and given the payoff function, each player has a reason to lower his contribution until it is zero. This result should be particularly strong under random matching rule as it does not provide any incentives for cooperative behavior.

Hypothesis 2. *Under assortative matching, we expect the contributions to approach the upper bound of 100. Therefore, the solution is Pareto optimal.*

Cooperation is expected given the structure of incentives under assortative matching. Knowing that the pairing depends on the individual's contributions, the agents should be much

[3]To approximate random matching, we generated random numbers for each agent, which were then sorted and ranked. These ranked numbers were then paired in a descending order. The randomly generated numbers were completely unrelated to effort levels chosen.

more likely to choose a strategy close to the upper bound since it will be increase the probability to be paired with a similar type (high-strategy).

3.5 Results

3.5.1 Assortative matching

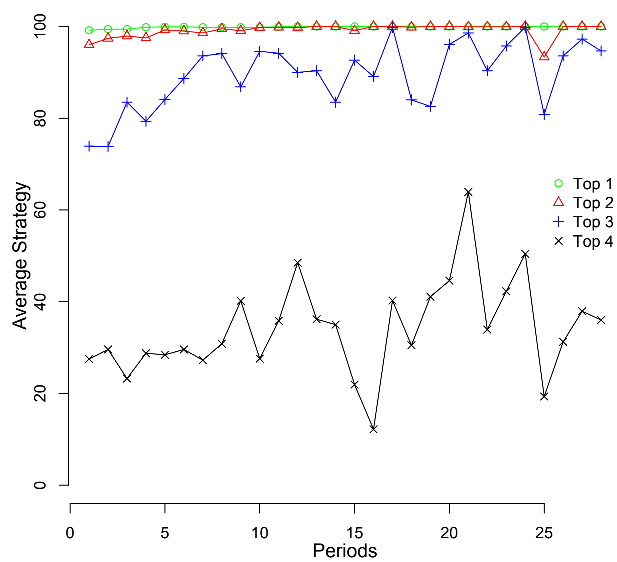


Figure 3.1: Group (pair) strategy per period (average-pooled data)

The experimental results show heterogeneity between subjects. Two groups (out of four) select a strategy equal to 100 (on average) during most of the session (28 periods). These top two groups, therefore, obtain cooperation and achieve the Pareto optimal solution. The third group also had a high-level of strategy. The level of contribution increases steeply following the first few periods but never completely approaches the upper bound. It appears that this group attempts to free-ride from the top groups while avoiding assignment with low contributors. In fact, Figure 3.1 shows a clear distinction between the third and fourth groups.[4] This last group achieves the lowest level of efficiency among all four groups. The strategies selected are remarkably lower when compared to other groups. Despite low level of efficiency, the fourth group never completely approaches the lower bound of zero, the NE of one-shot PD.

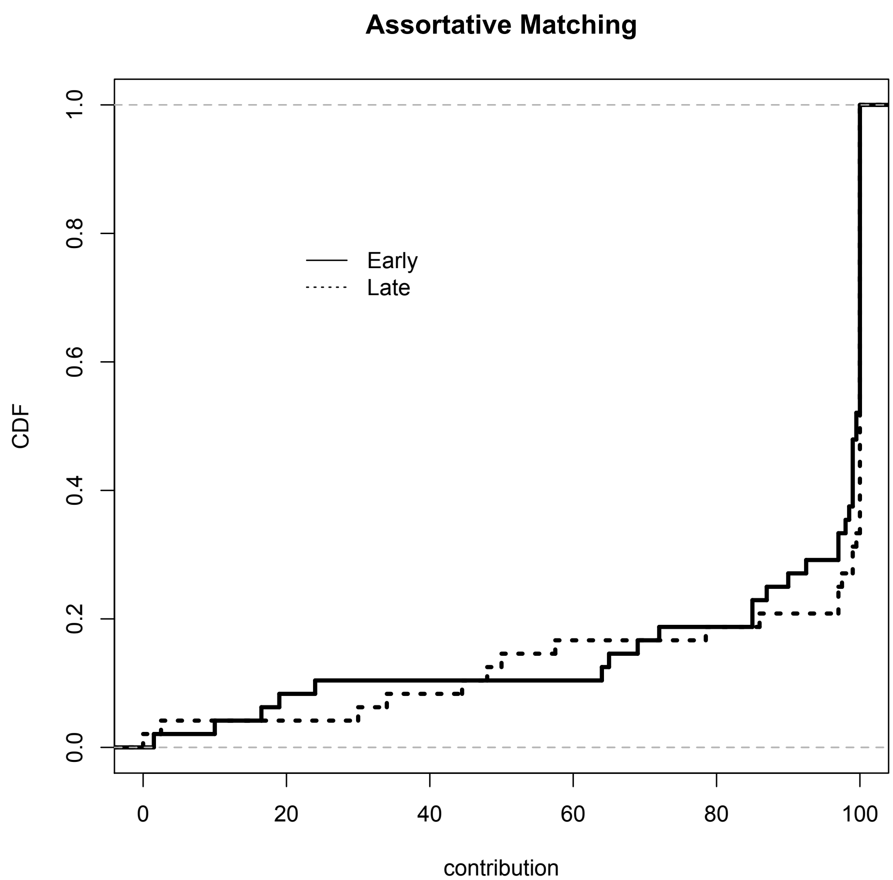


Figure 3.2: CDF of individual contribution (subject data)

[4]Notice that although we use averages, the median will give us the same information given that this is a pairwise comparison.

Additionally, the strategies display high levels of variation. While the top two groups are nearly stable, the bottom two fluctuates considerably in their strategy over the length of the session. Therefore, we find it more appropriate to work with the median as oppose to the mean. Figure 3.2 depicts the CDF of the individual decisions for two blocks: early (period 1-14) and late (period 15-28). Remarkably, the median strategy is about 100 in both blocks. In other words, half of the subjects stay close to the upper bound.

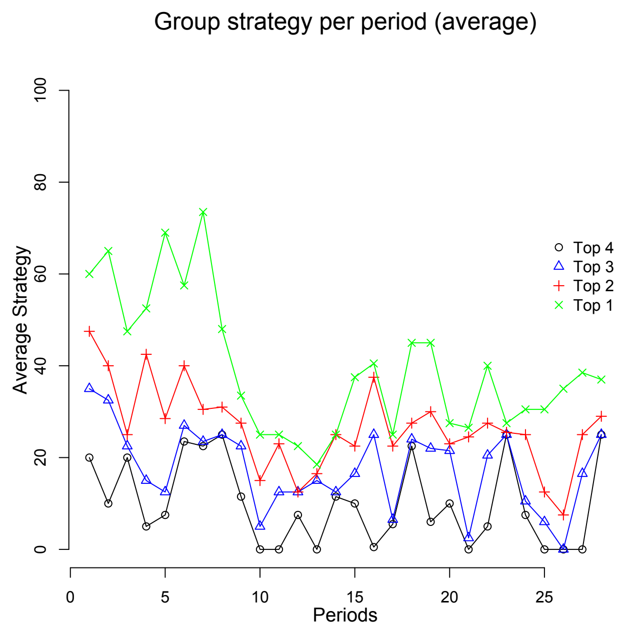


Figure 3.3: Group (pair) strategy per period under random matching (average-pooled data)

3.5.2 Random matching

Past literature indicates lower cooperation levels in one-shot PD type of games. Similarly, our data corroborates the findings of past studies and shows declining cooperation rates.

Furthermore, the strategies never fully approach the lower bound of zero. Therefore, our results are unable to achieve complete unraveling. Figure 3.3 illustrates the pairwise comparison as was done under assortative matching.[5]

[5]To clarify, this does not represent actual pairings but rather what the pairing would be under the assortative matching.

Bibliography

[1] Arce, D., 2007, ”Is agency theory self-acting?,” *Economic Inquiry*, 45 (4), pp. 708–720.

[2] Bajari, P., C. S. Chu, and M. Park, 2008, An Empirical Model of Subprime Mortgage Default From 2000 to 2007, National Bureau of Economic Research. Working Paper 14625.

[3] Benaim, M., J. Hofbauer and E. Hopkins, 2009, ”Learning in games with unstable equilibria,” *Journal of Economic Theory*, 144(4), pp. 1694–1709.

[4] Bhutta, N., J. Dokko, and H. Shan, 2010, The depth of negative equity and mortgage default decisions, Finance and Economics Discussion Series, Board of Governors of the Federal Reserve System (U.S.)(2010-35).

[5] Campbell, J. Y., S. Giglio, and P. Pathak, 2009, Forced Sales and House Prices, National Bureau of Economic Research. Working Paper. 14866.

[6] Chan, S., M. Gedal, V. Been, and A. Haughwout, 2010, The Role of Neighborhood Characteristics in Mortgage Default Risk: Evidence from New York City, Furman Center for Real State and Urban Policy. NYU. Working Paper.

[7] Charness, G. and P. Kuhn, 2011, *Lab Labor: What Can Labor Economists Learn from the Lab?*, in the Handbook of Labor Economics, Elsevier.

[8] Chaudhuri, A., 2010, "Sustaining cooperation in laboratory public goods experiments: a selective survey of the literature." *Experimental Economics*, 14, 47-83.

[9] Cox, J., S. Ross, and M. Rubinstein, 1979, Option Pricing: A Simplied Approach, *Journal of Financial Economics*, 7.

[10] Cooper, D. and J. Lightle, 2012, "The Gift of Advice: Communication in a Bilateral Gift Exchange Game.", mimeo.

[11] Cooper, R., D. DeJong, R. Forsythe, and T. W. Ross, 1996, "Cooperation without Reputation: Experimental Evidence from Prisoner's Dilemma Games." *Games and Economic Behavior*, 12(2): 187-218.

[12] Cressman, R., 2003, *Evolutionary dynamics and extensive form games*, MIT Press.

[13] Dekel, E., J.C. Ely and O. Yilankaya, 2007, "Evolution of Preferences," *Review of Economic Studies*, 74(3), pp. 685–704.

[14] Deng, Y., J. M. Quigley, and R. V. Order, 2000, Mortgage Terminations, Heterogeneity and the Exercise of Mortgage Options, *Econometrica*, 68(2), 275–308.

[15] Dixit, A. K., and R. S. Pindyck, 1994, Investment Under Uncertainty. Princeton University Press.

[16] Fehr, E. and S. Gachter, 2000, “Cooperation and punishment.” *American Economic Review*, 90, pp. 980-994.

[17] Fehr, E. and S. Gachter, 2002, “Altruistic punishment in humans.” *Nature*, 415, pp. 137-140.

[18] Foote, C. L., K. Gerardi, and P. S. Willen, 2008, Negative equity and foreclosure: Theory and evidence, *Journal of Urban Economics*, 64(2), 234–245.

[19] Foote, C. L., and P. S. Willen, 2011, subprime mortgage crisis, the, in The New Palgrave Dictionary of Economics, ed. by S. N. Durlauf, and L. E. Blume. Palgrave Macmillan.

[20] Friedman, D., 1991, “Evolutionary games in economics.” *Econometrica*, 69, pp. 637-666.

[21] Friedman, D. and R. Oprea, 2012, “A continuous Dilemma.” *American Economic Review*, 102, 337-363.

[22] Friedman, D. and D. Ostrov, 2010, ”Gradient Dynamics in Population Games: Some Basic Results,” *Journal of Mathematical Economics*, 46(5), pp. 691–70 .

[23] Friedman, D. and N. Singh, 2009, “Equilibrium vengeance.” *Games and Economic Behavior*, 66, pp. 813-829.

[24] Fudenberg, D. and D. Levine, 1998, *The Theory of Learning in Games*, MIT Press.

[25] Gerardi, K., A. H. Shapiro, and P. S. Willen, 2007, Subprime outcomes: risky mortgages, homeownership experiences, and foreclosures, Federal Reserve Bank of Boston. Working Papers. 07-15.

[26] Greiner,B, 2004, *An Online Recruitment System for Economic Experiments.* in Kurt Kremer, Volker Macho (Eds.), Forschung und wissenschaftliches Rechnen 2003. GWDG Bericht 63, Göttingen : Ges. für Wiss. Datenverarbeitung, 79-93.

[27] Guiso, L., P. Sapienza, and L. Zingales, 2009, Moral and Social Constraints to Strategic Default on Mortgages, Working Paper 15145, National Bureau of Economic Research.

[28] Güth, W. and M. Yaari, 1992, "An evolutionary approach to explaining reciprocal behavior." In: Witt, U. (Ed.), *Explaining Process and Change — Approaches to Evolutionary Economics.* The University of Michigan Press, Ann Arbor.

[29] Hofbauer, J. and K. Sigmund, 1988, *The Theory of Evolution and Dynamical Systems*, Cambridge University Press.

[30] Hirsch, M. and S. Smale, 1974, *Differential Equations, dynamical system and linear algebra*, Academic Press.

[31] Kaplan, E., and P. Meier, 1958, Nonparametric estimation from incomplete observations, *Journal of American Statistical Association*, (53), 457–481.

[32] Kau, J. B., D. Keenan, and K. Taewon, 1994, Default Probabilities for Mortgages, Default Probabilities for Mortgages, 35(3), 278–296.

[33] Kaufman, S., 1993, *The Origins of Order: Self-Organization and Selection in Evolution*, NY: Oxford U Press.

[34] Kreps. D., P. Milgrom, J. Roberts and R. Wilson, 1982, "Rational cooperation in a finetely repeated prisoners' dilemma.", *Journal of Economic Theory*, 27, 245-252.

[35] Lande, R., 1976, Natural Selection and Random Genetic Drift in Pheontypic Evolution, *Evolution*, 30 no. 2, 314–334.

[36] List, J. A., and M. S. Haigh, 2010, Investment under uncertainty: Testing the option model with professional traders, *The Review of Economics and Statistics*, 92(4), 974–984.

[37] Manski, C. F., 1993, Identication of Endogenous Social Effects: the Reection Problem, *Review of Economic Studies*, 60(3), 531–42.

[38] Oprea, R., D. Friedman, and S. T. Anderson, 2009, Learning to Wait: A Laboratory Investigation, *Review of Economic Studies*, 76(3), 1103–1124.

[39] Pennington-Cross, A., and G. Ho, 2010, The Termination of Subprime Hybrid and Fixed-Rate Mortgages, Real Estate Economics, 38(3), 399–4

[40] Ok, R. and F. Vega-Redondo, 2001, "On the evolution of individualistic preferences: An incomplete information scenario", *Journal of Economic Theory*, 97, pp. 231–254.

[41] Page, T., L. Putterman and U. Bulent, 2005, "Voluntary Association in Public Goods Experiments: Reciprocity, Mimicry and Efficiency." *The Economic Journal*, 115, 1032-1053.

[42] Sonnenschein, H., 1982, "Price dynamics based on the adjustment of firms. " *American Economic Review*, 72 (5), pp. 1088-1096.

[43] Sandholm, W., 2011, *Population games and evolutionary dynamics*, MIT Press.

[44] Taylor, P.D. and L.B. Jonker, 1978, "Evolutionary stable strategies and game dynamics." *Mathematical Biosciences*, 40, pp. 145-156.

[45] Urs, F., 2007, z-Tree: Zurich Toolbox for Ready-made Economic Experiments, *Experimental Economics*, 10(2), 171-178

[46] Vandell, K., 1995, How Ruthless is Mortgage Default? A Review and Synthesis of the Evidence, *Journal of Housing Research*, 6(2), 245–246.

[47] Weibull, W., 1997, *Evolutionary game theory*, MIT Press.

[48] White, B., 2009, Underwater and Not Walking Away: Shame, Fear and the Social Manage- Underwater and Not Walking Away: Shame, Fear and the Social Management of the Housing Crisis, Arizona Legal Studies Discussion Paper No. 09-35.

[49] Wilmott, P., S. Howison, and J. Dewynne, 1999, The mathematics of nancial derivatives. Cambridge.

[50] Wright, S., 1949, "Adaption and Selection," in L. Jepsen, G.G. Simpson, and E. Mayr eds., *Genetics, Paleontology, and Evolution.* Princeton, N.J.: Princeton University Press.

Appendix A

Supplemental material to Chapter 1

a.Finding DE and NE in the Arce (2007) Model

Recall that section 1.2.4 already identified all corner and edge DE and the subset that are NE.

On the 2-d faces that lie inside the 3-d faces $\varphi \in \{0,1\}$, section 1.2.1 noted that the only additional NE are the mixes $(\varphi, p, q_1, q_2) = (1, \frac{e}{w}, \frac{w-m}{w}, \cdot)$ and $(0, \frac{\alpha-e}{2\alpha-w}, \cdot, \frac{w-m}{w})$. The remaining 2-d faces involve $\varphi \in (0,1)$ and a strict mix of only one of the state variables p, q_1, q_2. The last two cases entail one of the q_j pure and the other strictly mixed, but (1.5) then implies that p is strictly mixed, contradicting the definition of this 2-d face. The remaining 2-d possibility involves $\varphi, p \in (0,1)$, which by (1.6) implies that $\varphi^* = \frac{m/w+q_2-1}{q_2-q_1}$. Ruling out[1] $q_2 - q_1 = 0$, we see from (1.5) that $p^* = e/w$. Consequently the only new candidate equilibria are $(\varphi^*, p^*, 1, 0)$ and $(\varphi^*, p^*, 0, 1)$. The dynamics of q_2 depends on the sign of $\frac{\alpha(w-2e)}{w}$ after plugging p^* in (1.8). The case $w - 2e > 0$ is called high incentive wages, and yields the $q_2^* = 1$

[1]It is hard to keep φ^* finite in that case, and by (1.7-1.8), it also entails non-generic parameters ($w = 2e$).

equilibrium, while low incentive wages, the case $w-2e<0$, yields the equilibrium above with $q_2^*=0$.

We have already found all NE in the 3-d faces $\varphi \in \{0,1\}$. The 3-d faces $p \in \{0,1\}$ have no NE, since q_j is strictly mixing for states in such faces, and therefore $p = p^*$ by (1.8), contradicting $p \in \{0,1\}$. Similarly, the faces $q_1 \in \{0,1\}$ contain no new NE since a strictly mixed strategy for q_2 implies $p = p^{**} = (\alpha - e)/(2\alpha - w)$ which contradicts the solution of p^* in (1.5). On the faces $q_2 \in \{0,1\}$ we pick up two new NE, $(\varphi^*, \frac{e}{w}, \frac{-m+w\varphi^*}{w\varphi^*}, 1)$ and $(\varphi^*, \frac{e}{w}, \frac{-m+w}{w\varphi^*}, 0)$; the argument parallels that for the 2-d face where $\varphi, p \in (0,1)$. Keeping the third component $q_1 \in [0,1]$ implies the restriction $\varphi \in [\frac{m}{w}, 1]$ for the first new NE and $\varphi \in [\frac{w-m}{w}, 1]$ for the second.

Finally, the interior points are unstable since we already know that the dynamics of q_2 depends on the sign of $\frac{\alpha(w-2e)}{w}$ which forces it to 1 (or zero) in the case of high (or low) wage.

b. Evaluating the Jacobian matrix at the NE

The text analyzed stability for the first three NE and the last NE listed in (1.10). In this section, we find Jacobian matrices and eigenvalues for the remaining NE.

The Jacobian matrix for (1.5 - 1.8) evaluated at the equilibrium $(\varphi, p, q_1, q_2) = (0, 1, \cdot, 0)$ is

$$J = \begin{pmatrix} \beta_1 q_1 (w-e) & 0 & 0 & 0 \\ 0 & -\beta_2 (w-m) & 0 & 0 \\ 0 & \beta_3 (1-q_1) q_1 w & \beta_3 (1-2q_1)(w-e) & 0 \\ 0 & 0 & 0 & -\beta_4 (\alpha - (w-e)) \end{pmatrix},$$

whose eigenvalues are $\{-\beta_4(\alpha-(w-e)), -\beta_2(w-m), \beta_1 q_1 (w-e), \beta_3(1-2q_1)(w-e)\}$. As noted in the text, the first two are always negative in our parameter space. The third is positive except when $q_1 = 0$, in which case the last eigenvalue is positive. Hence this NE is definitely not a DSE.

The Jacobian matrix evaluated at $(0,0,\cdot,1)$ is

$$J = \begin{pmatrix} \beta_1 e(1-q_1) & 0 & 0 & 0 \\ 0 & -\beta_2 m & 0 & 0 \\ 0 & \beta_3(1-q_1)qw & \beta_3 e(-1+2q_1) & 0 \\ 0 & 0 & 0 & -\beta_4(\alpha - e) \end{pmatrix},$$

whose eigenvalues are $\{-\beta_4(\alpha - e), -\beta_2 m, \beta_1 e(1-q_1), \beta_3 e(-1+2q_1)\}$. The third is positive except when $q_1 = 1$, in which case the last eigenvalue is positive. Hence this NE also is definitely not a DSE.

The Jacobian at $(x,1,1,1)$ is

$$J = \begin{pmatrix} 0 & 0 & \beta_1(w-e)(1-\varphi)\varphi & \beta_1(w-e)(1-\varphi)\varphi \\ 0 & \beta_2 m & 0 & 0 \\ 0 & 0 & -\beta_3(w-e) & 0 \\ 0 & 0 & 0 & -\beta_4(\alpha-(w-e)) \end{pmatrix},$$

whose eigenvalues are $\{0, -\beta_4(\alpha-(w-e)), \beta_2 m, -\beta_3(w-e)\}$. The second and third are positive, so this equilibrium is not a DSE.

The Jacobian at $(1, e/w, (w-m)/w, \cdot)$ is

$$J = \begin{pmatrix} 0 & 0 & 0 & 0 \\ \frac{\beta_2 e(w-e)(m+(-1+q_2)w)}{w^2} & 0 & -\beta_2 e\left(1-\frac{e}{w}\right) & 0 \\ 0 & \frac{\beta_3 m(w-m)}{w} & 0 & 0 \\ 0 & \beta_4(1-q_2)q_2(w-2\alpha) & 0 & \frac{\beta_4(-1+2q_2)(2e-w)\alpha}{w} \end{pmatrix},$$

with eigenvalues $\left\{0, \pm\frac{\sqrt{-\beta_2\beta_3 em(w-m)(w-e)}}{w}, \frac{\beta_4(-1+2q_2)(2e-w)\alpha}{w}\right\}$. The last of these is positive in the low wage case when $q_2 > 0.5$ and in the high wage case when $q_2 < 0.5$; in either case it can be destabilized by an invasion of type 2 agents and so is not a DSE.[2]

The Jacobian at $(0, (\alpha-e)/(2\alpha-w), \cdot, (w-m)/w)$ is

$$J = \begin{pmatrix} -\frac{\beta_1(w-2e)(m-(1-q_1)w)\alpha}{w(w-2\alpha)} & 0 & 0 & 0 \\ \frac{-\beta_2(m-(1-q_1)w)(\alpha-(w-e))(\alpha-e)}{(w-2\alpha)^2} & 0 & 0 & \frac{-\beta_2 w(\alpha-(w-e))(\alpha-e)}{(w-2\alpha)^2} \\ 0 & \beta_3(1-q_1)q_1 w & \frac{-\beta_3(-1+2q)(w-2e)\alpha}{2\alpha-w} & 0 \\ 0 & -\frac{\beta_4 m(w-m)(2\alpha-w)}{w^2} & 0 & 0 \end{pmatrix},$$

with eigenvalues $\left\{\pm\sqrt{\frac{\beta_4\beta_2 m(w-m)(\alpha-e)(\alpha-(w-e))}{w(2\alpha-w)}}, \frac{\beta_3(1-2q_1)(w-2e)\alpha}{2\alpha-w}, \frac{-\beta_1(w-2e)(m-(1-q_1)w)\alpha}{w(2\alpha-w)}\right\}$. The first pair is real with opposite signs, so this NE is not a DSE.

The Jacobian at $(\frac{w-m}{w}, \frac{e}{w}, 1, 0)$ is

$$J = \begin{pmatrix} 0 & \frac{\beta_1 m(w-m)}{w} & 0 & 0 \\ -\beta_2 e\left(1-\frac{e}{w}\right) & 0 & \frac{-\beta_2 e(w-e)(w-m)}{w^2} & \frac{-\beta_2 em(w-e)}{w^2} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{\beta_4(w-2e)\alpha}{w} \end{pmatrix},$$

[2] Indeed, performing simulations, q_2^* goes to 1 when $w-2e>0$ and q_2^* goes to 0 when $w-2e<0$ for initial values $q_2(0)\in(0,1)$, and $\varphi(0)$, $p(0)$ and $q_1(0)$ relatively closed to the key equilibrium. See section 1.2.5 for the simulation results of the equilibria that survive in the long run.

with eigenvalues $\left\{0, \frac{\beta_4(w-2e)\alpha}{w}, \pm\sqrt{\frac{-\beta_1\beta_2 em(w-m)(w-e)}{w}}\right\}$. The second is negative in the relevant case of low wages, $w-2e<0$, and the last pair is pure imaginary. Hence this NE remains a candidate DSE, requiring further investigation. It can be seen to be an extreme case of the NE family listed last in (1.10) and already analyzed in the text.

At $(\varphi^*, e/w, 0, 1)$, we have $\varphi^* = m/w$ and the Jacobian is

$$J = \begin{pmatrix} 0 & -\beta_1 m\left(1-\frac{m}{w}\right) & 0 & 0 \\ \beta_2 e\left(1-\frac{e}{w}\right) & 0 & \frac{-\beta_2 em(w-e)}{w^2} & \frac{\beta_2 e(w-e)(w-m)}{w^2} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{-\beta_4(w-2e)\alpha}{w} \end{pmatrix},$$

with eigenvalues $\left\{0, \frac{-\beta_4(w-2e)\alpha}{w}, \pm\sqrt{\frac{-\beta_1\beta_2 em(w-m)(w-e)}{w}}\right\}$. The second is negative in the relevant case of high wages, $w-2e>0$, so this NE also remains a candidate DSE. It is an extreme case of the next NE family.

The Jacobian at $\left(\varphi^*, \frac{e}{w}, \frac{-m+w\varphi^*}{w\varphi^*}, 1\right)$ is

$$J = \begin{pmatrix} 0 & -\beta_1 m(1-\varphi^*) & 0 & 0 \\ \frac{\beta_2 em(w-e)}{w^2\varphi^*} & 0 & -\beta_2 e\left(1-\frac{e}{w}\right)\varphi^* & -\beta_2 e\left(1-\frac{e}{w}\right)(1-\varphi^*) \\ 0 & \frac{\beta_3 m(-m+w\varphi^*)}{w\varphi^{*2}} & 0 & 0 \\ 0 & 0 & 0 & \frac{-\beta_4(w-2e)\alpha}{w} \end{pmatrix},$$

with eigenvalues $\left\{0, \frac{-\beta_4(w-2e)\alpha}{w}, \pm\frac{\sqrt{\beta_2 em(w-e)(m(-1+\varphi^*)\beta_1+(m-w\varphi^*)\beta_3)}}{w\sqrt{\varphi^*}}\right\}$. The second is negative in the high wage case, and the last pair is pure imaginary since $\frac{-m+w\varphi^*}{w\varphi^*} \geq 0$, so the entire family with $\varphi^* \in [\frac{m}{w}, 1]$ is a candidate DSE in the high wage case.

Appendix B

Supplemental material to Chapter 2

a. Instructions - UW

Welcome! This is an economics experiment. If you pay close attention to these instructions, you can earn a significant sum of money. It will be paid to you in cash at the end of the last period.

Please remain silent and do not look at other participants' screens. If you have any questions, or need assistance of any kind, please raise your hand and we will come to you. If you disrupt the experiment by talking, laughing, etc., you may be asked to leave without compensation. We expect and appreciate your cooperation today.

The Basic Idea

Each period you will receive an asset that has been acquired with a loan. During the period, you will make interest payments on the loan and watch the asset value change randomly with time. You can decide to stop paying at any point, in which case the asset value no longer matters to you. If you keep the asset until the end of the period, you will receive its final value less the loan amount.

Assets over time The session today is divided in a number of periods, each lasting 36 seconds. During the period, you will observe fluctuations in the asset value. At each point, the asset value can go UP or DOWN with a given probability. The increment value and the probability of going up/down are written on the white board. The asset value is displayed as the black line on your screen (see Figure 1).

Liabilities over time In the beginning of each period, you acquire the asset with a loan. The characteristics of the loan are such that you only pay interest payments until the end of the period. At the end of period, you also pay the principal. Your screen also shows the sum of future payments as a blue line. When the black line is above the blue line, it indicates that the asset value currently exceeds the future payments (liabilities). When it is below the blue line, those liabilities currently exceed the asset value.

Stopping You can decide to stop paying the loan at any time by pressing the space bar. When you press, you have to pay a quit fee and you save the sum of future interest payments and principal, but you no longer own the asset. If you want to see what the value of stopping is now, you can look at the text at the center top of the screen. It shows the current gain if you stop paying — the liabilities saved minus the quit fee and the asset value. The quit fee is shown as a red bar under the blue line, as in Figure 1.

Anytime that you stop by pressing the space bar you will observe a vertical bar that appears at the time you decide to stop.

Payoffs Your payoff at the end of the period depends on whether you stopped paying. If you didn't stop paying, you get Base pay + final asset value — all interest payments — loan amount. If you did stop paying you get Base pay — interest payments already made — quit fee. In some periods you may receive a bonus, an addition to your base pay, if you never stop the interest payments. This bonus is shown as a bar on top of the blue line. This bonus may change during the period, depending on the decisions of other players in your group. The bonus shrinks whenever someone in your group stops paying. Your screen shows the payoff you get on the top right corner and it also represent in the green bar that appears either when you decide to stop or at the end of the period. If the bonus depends on the decision of other players, you will see the fraction of players that stop paying below the "Gain Stop Paying".

Earnings You will be paid at the end of the experiment for all points earned over all periods. You can see your total points on the top left of the screen. The conversion rate from payoff points to U.S. dollars is written on the whiteboard.

Frequently Asked Questions

Q1. Is this some kind of psychology experiment with an agenda you haven't told us? Answer. No. It is an economics experiment. If we do anything deceptive or don't pay you cash as described then you can complain to the campus Human Subjects Committee and we will be in serious trouble. These instructions are meant to clarify how you earn money, and our interest is in seeing how people make decisions.

Q2. Is the gain of stopping equal to the payoff I get? Answer. No. The gain of stop paying tells us what is the benefit to stop paying right now against the alternative of paying until the end of the period. The payoff you receive considers the interest payments that you have already paid and the quit fee if you decide to stop.

Figures

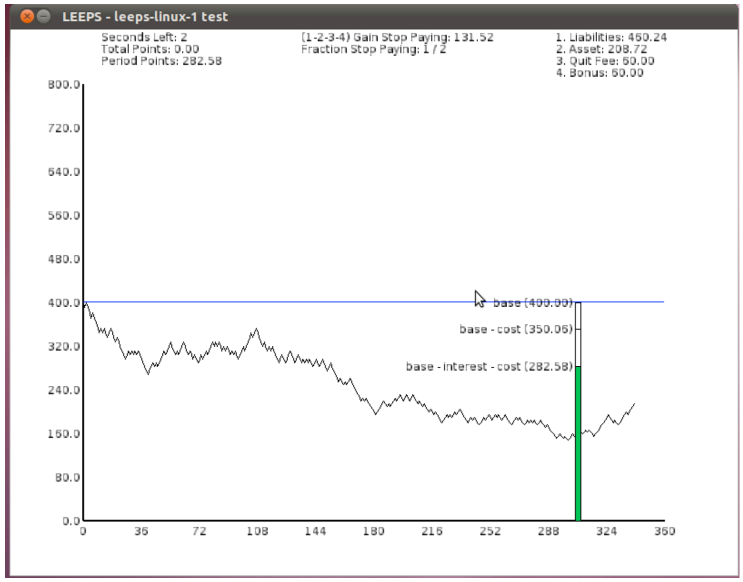


Figure B.1: Screen Plot - Stopping early

b. Different asset price subsamples

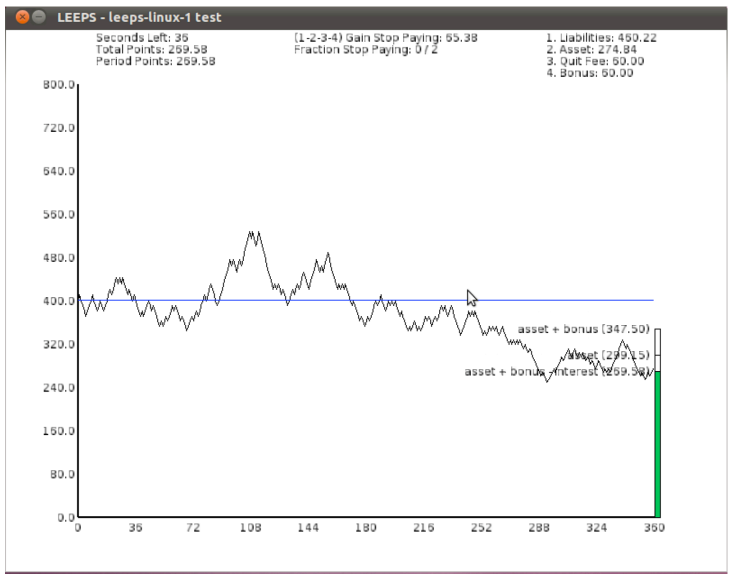


Figure B.2: Screen Plot - No Stopping

Table B.1: Stopping decision

	Low Dep.	Low Indep.	High Dep.	High Indep.
Sample: All				
Subjects	34	36	34	14
% Default	22	23	24	37
Observations	2380	2198	2380	968
Sample: min. price < threshold				
Subjects	34	36	34	14
% Default	25	41	39	79
Observations	634	537	578	234
Sample: min. price < 90% of threshold				
Subjects	34	36	34	14
% Default	26	45	39	79
Observations	416	321	496	170
Sample: min. price < 80% of threshold				
Subjects	34	36	34	14
% Default	28	47	39	88
Observations	194	219	402	110
Sample: min. price < 70% of threshold				
Subjects	34	33	34	14
% Default	29	51	41	90
Observations	90	131	308	98
Sample: min. price < 60% of threshold				
Subjects	34	26	34	14
% Default	32	53	42	94
Observations	34	51	240	62

Note: Obs. is the total number of loans. Threshold refers to the prediction of the default asset value.

Appendix C

Supplemental material to Chapter 3

Assortative Matching - Instructions

Welcome! This is an economics experiment. If you pay close attention to these instructions, you can earn a significant sum of money, which will be paid to you in cash at the end of the last period.

Please remain silent and do not look at other participants screens. If you have any questions, or need assistance of any kind, please raise your hand and we will come to you. If you disrupt the experiment by talking, laughing, etc., you may be asked to leave and may not be paid. We expect and appreciate your cooperation today.

The Basic Idea The experiment will be divided into a number of periods and in each period you will be matched with another player according to your strategy. In each period you and your counterpart will secretly select strategies. At the end of the period the combination of your and your counterpart strategy will determine your earnings for the period.

Your earnings are computed by the following function:

Earnings = 100 - YOUR STRATEGY + 5 * OTHER

Your earnings are symmetric with your counterpart's. In particular, if you and your counterpart both choose the same strategy, then you will both earn the same amount.

Step by step After everyone decides on a strategy between 0 and 100, the computer will sort the strategies and match pairs according to their rank. For example, the top 2 strategies will be match together, then the next 2 and so on.

The earnings are computed with the earning function described above. This function will not change over the course of the experiment.

The screen display

Figure C.1 shows the computer display you will use to make decisions. At the top of screen is the time left as well as the number of the current period. There are also two panels. The left panel gives you information about the earnings function. The right panel allows you to enter your strategy. After you write the input desired, you have to press CONFIRM.

Figure C.2 shows the computer display with the matches and earnings. The screen will list all the matches formed during the period as well as the strategies.

Random Matching - Instructions

Welcome! This is an economics experiment. If you pay close attention to these instructions, you can earn a significant sum of money, which will be paid to you in cash at the end of the last period.

Please remain silent and do not look at other participants screens. If you have any questions, or need assistance of any kind, please raise your hand and we will come to you. If you disrupt the experiment by talking, laughing, etc., you may be asked to leave and may not be paid. We expect and appreciate your cooperation today.

The Basic Idea The experiment will be divided into a number of periods and in each period you will be randomly matched with another player. In each period you and your counterpart will secretly select strategies and at the end of the period the combination of your and your counterpart strategy will determine your earnings for the period. Your earnings are computed by the following function:

Earnings = 100 - YOUR STRATEGY + 5 * OTHER

Your earnings are symmetric with your counterpart's. In particular, if you and your counterpart both choose the same strategy, then you will both earn the same amount.

Step by step After everyone decides on a strategy between 0 and 100, the computer will randomly match you with a counterpart.

The earnings are computed with the earning function described above. This function will not change over the course of the experiment.

The screen display

Figure C.1 shows the computer display you will use to make decisions. At the top of screen is the time left as well as the number of the current period. There are also two panels. The left panel gives you information about the earnings function. The right panel allows you to enter your strategy. After you write the input desired, you have to press CONFIRM.

Figure C.2 shows the computer display with the matches and earnings. The screen will list all the matches formed during the period as well as the strategies.

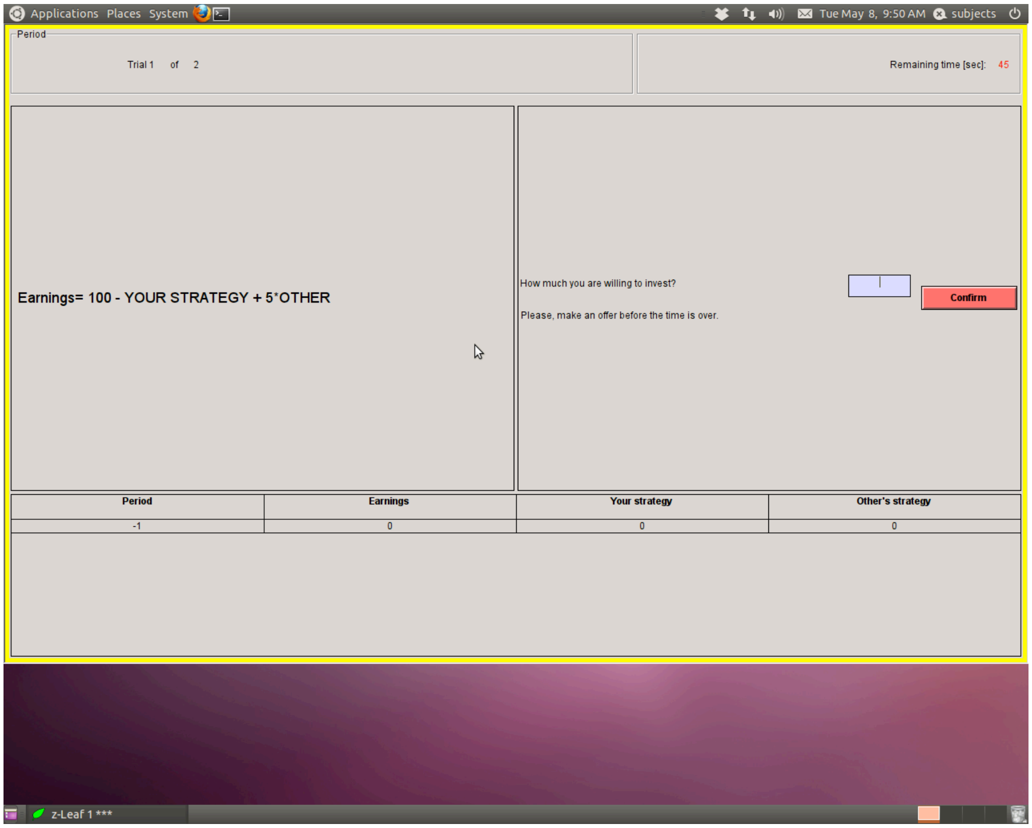


Figure C.1: Screen 1- Computer display

Earnings Your earnings will be given in points. Your points will accumulate over the course of the experiment. The screen will always display your Earnings for each period on the bottom. It will also list all strategies chosen during that period. You will be paid cash for points earned at a rate written on the white board at the front of the room.

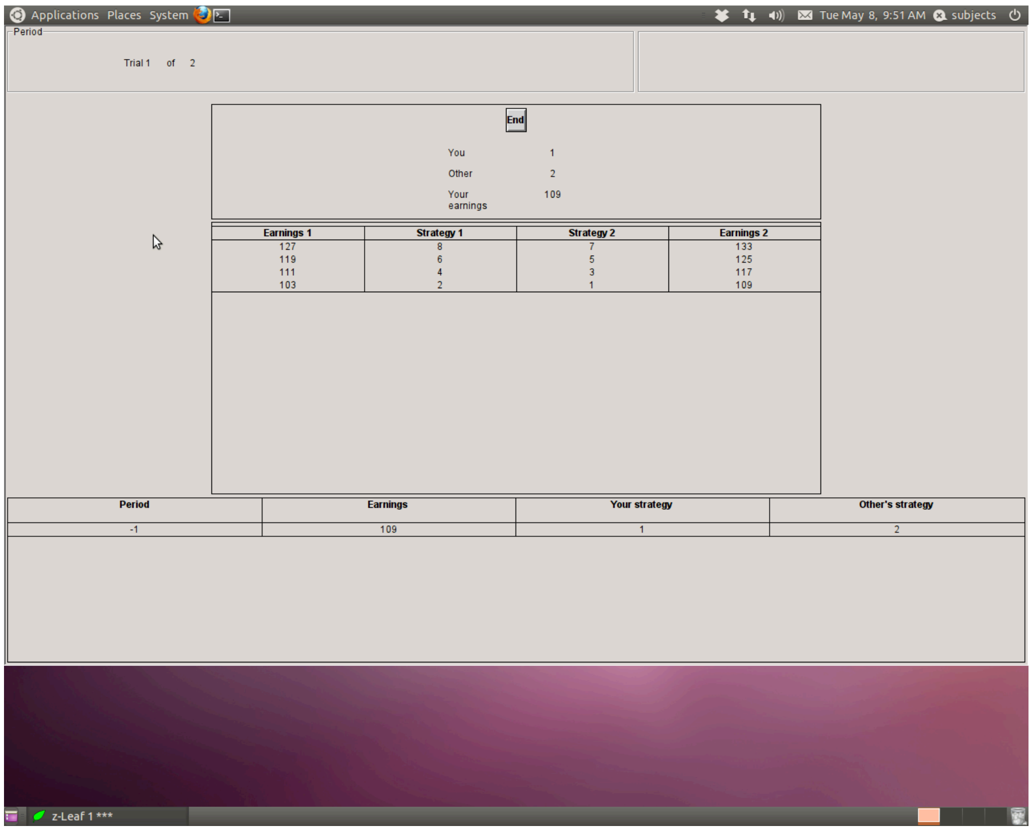


Figure C.2: Screen 2- Computer display