Title

Incomplete Information, Dynamic Stability and the Evolution of Preferences: Two Examples

Journal

Dynamic Games and Applications, 4(4)

ISSN

2153-0785

Authors

Rabanal, Jean Paul
Friedman, Daniel

Publication Date

2014-12-01

DOI

10.1007/s13235-013-0096-5

Peer reviewed

Incomplete Information, Dynamic Stability and the Evolution of Preferences: Two Examples

Jean Paul Rabanal*
Department of Economics and Finance
City University of Hong Kong

Daniel Friedman
Economics Department
UC Santa Cruz

September 12, 2013

Abstract

We illustrate general techniques for assessing dynamic stability in games of incomplete information by re-analyzing two models of preference evolution, the Arce (2007) employer-worker game and the Friedman and Singh (2009) Noisy Trust game. The techniques include extensions of replicator and gradient dynamics, and for both models they confirm local stability of the key static equilibria. That is, we obtain convergence in time average for initial conditions sufficiently near equilibrium values.

Keywords: Stability, Perfect Bayesian Equilibrium, evolutionary dynamics.
JEL codes: C62, C73

*Corresponding Author. *Email:* jrabanal@cityu.edu.hk *Address:* Department of Economics and Finance, City University of Hong Kong, Kowloon, *Tel:* 852-3442-7600 *Fax:*852-3422-0195

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, imitation or learning); see, for example Weibull (1995), Friedman (1991) and Sandholm (2010) 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 addresses those questions. It does not prove new general results nor offer new models, but it does show how to extend dynamics (specifically, replicator 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.

The dynamical systems that we propose endogenize the evolution of types. In that sense they go beyond recent studies of dynamic stability of games of incomplete information that hold fixed the distribution of types in order to focus on the adaptation of beliefs and/or actions. These recent studies include Ely and Sandholm (2005), which uses best-response dynamics; Cressman (2003, Section 4.7.2) and Amann and Possajennikov (2009) which apply replicator dynamics. The present paper greatly extends the approach of Possajennikov (2005), who fixes a discrete set of four types in Prisoner's Dilemma and Hawk-Dove games in complete and incomplete information environments, and uses replicator dynamics to endogenize the shares of the alternative types.

In the present paper we embed the types in a continuous type space and allow them to evolve (slowly) via gradient dynamics, while replicator dynamics describe the rapid adaptation of individual behavior. Endogenizing the set of types sharpens the model's predictions and may enhance the applicability of evolutionary game theory when information is imperfect. In some applications,

tremble rates and imperfect observability are key features of the game of incomplete information and we therefore show how to incorporate them into a dynamic specification. On the other hand, we do not explicitly model the adaptation of beliefs; for our applied focus they can be regarded as following the distribution of actions as it evolves.

We take no strong position on which underlying processes govern the evolution of types and actions. It is natural for economists to think of entry and exit as governing type evolution, but other social scientists may regard individual preference types as subject to gradual change or conversion under the influence of role models and peers. For a given type of individual, the distribution of actions can adapt via personal experience of success and failure, or social learning, or even entry and exit. The underlying process determines the precise dynamic specification.

We illustrate the techniques using two recent models of preference formation, Arce (2007) and Friedman and Singh (2009). Our idea is that generality is best developed from specific instances and we chose these two models as relatively simple instances of the complexities noted two paragraphs earlier. The models both apply the *indirect evolution* approach to preferences: players of given preference types are matched pairwise to play a game with specified material payoffs. The players adapt actions (on a rapid time scale) so as to maximize expected utility under the given type distribution, while (on slower time scales) the distribution among types and the types themselves evolve according to realized material payoffs. Previous investigations represented indirect 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 the relevant equilibria.

Our work draws on several other strands of literature. Our continuous type space is an instance of continuous trait space, for which the static equilibrium concepts of continously stable strategies (CSS) and neighborhood invader stable (NIS) were first introduced respectively by Eshel (1983) and Apaloo (1997). Later authors, notably Oechssler and Riedel (2002), Cressman (2005), and Doebeli and Hauert (2005), connected these static concepts to dynamic stability under replicator dynamics and Dieckman's canonical dynamics. Friedman and Ostrov (2010, 2013) argue that gradient dynamics are natural when the topology of the continuous action space matters — e.g., when it is easier to evolve (or cheaper to adjust) to a nearby action or trait than to a distant one — and in particular when payoff or fitness is given by the expectation over the current distribution of traits or actions encountered in interactions. Applying the fitness gradient to each action chosen in a large

population, they obtain a partial differential equation governing the distribution of actions and note conditions under which the distribution converges asymptotically to a mass point or to a stationary density. Dieckman's canonical dynamics are a variant of gradient dynamics in which the mean of the current action distribution responds to the payoff gradient while the variance remains constant. We will elaborate on these points (and mention other published work) later when presenting the dynamic analysis.

The presentation begins in Section 2 with Arce's (2007) equilibrium analysis of efficiency wages in a two population game of incomplete information. A population of Workers containing two types (one self-interested and the other autonomy-preferring) is randomly matched pairwise with a population of Employers who cannot observe Worker type. Which preference types and what sort of behavior will survive in the long run? Arce uses static concepts to indentify equilibria and emphasizes the result that increasing the incentive wage can destabilize an efficient equilibrium for some distributions of preference types.

The first step in our analysis is to write out the expected payoff and expected utilities given all state variables, including the population share of each type of Workers, 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 the adjustment of strategy mixtures, the ODE system focuses on utility and applies gradient dynamics, which in this context coincides with standard replicator dynamics. The parameters include speed of adjustment, and we emphasize cases where strategy mixture adjustment is faster than the evolution of types.

For some equilibria, eigenvalue techniques allow us to analytically 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 other theoretical work on models of indirect evolution, e.g. Ok and Vega-Redondo (2001) and Dekel, Ely and Yilankaya (2007). In particular, we find that both kinds of Workers can coexist in states where the aggregate play corresponds to a Nash equilibrium with selfish preferences. Our simulations also provide insight into how key parameters influence transitory behavior.

Section 3 describes the basic trust game and its extension to a game of incomplete information

due to Friedman and Singh (2009, henceforth FS09). They propose 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 individuals adjust their strategy mixes rapidly relative to the rate of change in population shares (which adjust via entry and exit, or type switching). We assume that preference parameter adjusts (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 the equilibrium value. That is, the "good" EPBE of the noisy trust game is locally stable in time average.

The last section summarizes the findings and offers suggestions for applying the techniques more broadly. The appendix includes the mathematical details of the Arce model.

2 The Arce (2007) Employer-Worker 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.

2.1 Elements of the model

Each Employer (row player) in a large population is randomly matched with one Worker (column player). There are two possible types of workers: 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 worker decides only whether to work (W) or shirk (S); the respective mixture probabilities are

denoted q_j and $1-q_j$ for each type $j=1,2$.[1] Thus workers' type-contingent strategy space is $[0,1]\times[0,1]$.

Each Employer decides whether to monitor (M) or not (N); the mixture probabilities are p and $1-p$ respectively. Thus Employer'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 shows the payoffs. The Employer receives gross payoff $v>0$ when the worker works, offset by the wage $w>0$ paid to the worker and by the monitoring cost $m>0$. Note that in this model, the Employer is unable to condition the wage on the level of output. Workers who work incur an effort cost $e>0$. In addition to these material payoffs, a Type 2 worker receives a utility increment $(+\alpha)$ when her Employer does not monitor and a symmetric decrement $(-\alpha)$ with monitoring. In the case that a worker shirks and gets caught, her payoff is equal to zero. The parametric restrictions

$$w>e,\quad w>m,\quad \alpha>e,\quad \alpha+e>w \tag{1}$$

ensure that shirking is not a dominant strategy and eliminate other trivial cases. Arce (2007) sets wage at the value $w=\sqrt{vm}$ that maximizes the Employer's expected payoff; in this case (1) implies restrictions on the gross payoff v.

Table 1: **Employer (row) and Worker (column) Payoffs**. The fraction of Type 1 workers is φ, and mixing probabilities are p for Employer and q_j for workers 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 workers 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 workers are known to be Type 2, there is again a mixed NE in which $p=p^{**}=(\alpha-e)/(2\alpha-w)$, $q_2=q_2^*=(w-m)/w$,

[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.

as well as two pure NE: one at (N,W) or $p=0$, $q_2=1$ and the other at (M,S) or $p=1$, $q_2=0$.

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 Employer's expected payoff ω^P in equation (2) below arises from receiving v when the employee works (probability $\varphi q_1 + (1-\varphi)q_2$), minus the monitoring costs incurred (with probability p) and the wages paid to the worker. Recall from Table 1 that the Employer pays w unless he monitors and the worker shirks, an event of probability $p(\varphi(1-q_1)+(1-\varphi)(1-q_2))$. Thus

$$\omega^P = (\varphi q_1 + (1-\varphi)q_2)\cdot v - p\cdot m - [1-p(\varphi(1-q_1)+(1-\varphi)(1-q_2))]\cdot w \tag{2}$$

Both types of workers 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 Employer does not monitor $(1-p)$. For the self-interested worker (type 1), expected utility coincides with expected material payoff ω^{A1}, which is therefore

$$\omega^{A1} = q_1\cdot(w-e) + (1-q_1)(1-p)\cdot w. \tag{3}$$

Similarly, material payoff for Type 2 worker is

$$\omega^{A2} = q_2\cdot(w-e) + (1-q_2)(1-p)\cdot w, \tag{4}$$

while her expected utility includes the preference parameter α and is

$$\omega_\alpha^{A2} = q_2\cdot[p\cdot(w-e-\alpha) + (1-p)\cdot(w-e+\alpha)] + (1-q_2)(1-p)\cdot w. \tag{5}$$

2.3 Dynamic adjustment equations

Recall that the state space is four dimensional, and specifies the fraction φ of type 1 workers, Employer's mixing probability (p) and workers' mixing probabilities (q_j). We therefore specify dynamics as a system of four coupled ordinary differential equations (ODEs), derived from ex-

pected payoffs using standard evolutionary principles.

Arce (2007, p.718) comments, "This then begs the question, what determines the initial distribution of agent types?" and cites several exogenous factors. For our purposes it is better to complete the model by endogenizing the distribution φ. We invoke the basic principle of evolution that the type with higher material pyoff (= fitness) will increase its share of the population. More specifically,[2] 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 workers is proportional (with rate constant β_φ) to its payoff ω^{A1} relative to the population average ($\bar{\omega}$). Equation (6) and other equations below use the fact that relative payoff can be written as $\omega^{A1} - \bar{\omega} = \omega^{A1} - \varphi\omega^{A1} - (1-\varphi)\omega^{A2} = (1-\varphi)(\omega^{A1} - \omega^{A2})$. Thus $\dot{\varphi}$ is equal to $\varphi(1-\varphi)(\omega^{A1} - \omega^{A2})$ times a positive adjustment speed parameter β_φ.

The remaining equations apply replicator dynamics for the mixture probabilities p, q_1 and q_2. Thus the system of four coupled ODEs is

$$\begin{aligned}
\dot{\varphi} &= \beta_\varphi \varphi(1-\varphi)[\omega^{A1} - \omega^{A2}] && (6)\\
&= \beta_\varphi \varphi(1-\varphi)[(pw-e)(q_1-q_2)] \\
\dot{p} &= \beta p(1-p)\frac{\partial \omega^P}{\partial p} && (7)\\
&= \beta p(1-p)[-m + (\varphi(1-q_1) + (1-\varphi)(1-q_2)) \cdot w] \\
\dot{q}_1 &= \beta q_1(1-q_1)\frac{\partial \omega^{A1}}{\partial q_1} && (8)\\
&= \beta q_1(1-q_1)(pw-e) \\
\dot{q}_2 &= \beta q_2(1-q_2)\frac{\partial \omega^{A2}_\alpha}{\partial q_2} && (9)\\
&= \beta q_2(1-q_2)[(w-2\alpha)p + \alpha - e]
\end{aligned}$$

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

As noted earlier, we assume that the mixing probabilities adjust more rapidly than the type distribution φ, i.e., that $\beta >> \beta_\varphi > 0$. The restrictions (1) apply to parameters w, e, m, α (or to

[2]Here and elsewhere, in modelling the adjustment of shares or mixture probabilities in [0,1] for two (pure) alternatives, there are many smooth monotone (or sign preserving) dynamic specifications to choose among. As catalogued in Weibull (1997) and Sandholm (2010), these include BNN, perturbed best response and various sorts of learning dynamics. The techniques illustrated below can straightforwardly be tailored to such specifications. In our experience with state spaces built from [0,1] factors, the stability results are insensitive to the choice of a specific smooth monotone dynamic, but we offer no guarantee.

v, e, m, α if w is chosen by the square root formula). 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 $0 \leq \varphi \leq 1$.

Given only two possible strategies for each type of player (and thus a probability distribution described by one variable) and random matching (and therefore expected payoffs linear in the relevant proportions), the payoff differences across strategies, e.g., $\omega^P|_{[p=1]} - \omega^P|_{[p=0]}$ coincide with the payoff gradients, e.g., $\frac{\partial \omega^P}{\partial p}$. Thus the equations for the mixture probabilities can also be reinterpreted as gradient dynamics supplemented by a binomial variance factor.

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 $t \to \infty$.

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.[3] It is also well known that DSE are a subset of NE and that, for smooth systems of ODEs like (6 - 9), 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 eigenvalues 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 guarantees) failure of asymptotic stability.[4]

[3] In the present case, however, we include an extra DE condition in (10) below that $\dot{\varphi} = 0$. This eliminates from the outset those NE for which the two different surviving types of workers have different material payoffs in equilibrium.

[4] In such cases, it seldom helps to look at second order expansions of the dynamical system, but often third order terms can resolve local stability questions, at the cost of considerable analytic complication. Lyupanov functions

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 (6 - 9) 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

$$\dot{\varphi} = \dot{p} = \dot{q}_1 = \dot{q}_2 = 0. \tag{10}$$

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 (10). 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 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 edge DE in all 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 (10) and (6) 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 $(p, q_1) = (0,1)$ and $(p, q_2) = (1,1)$. Hence there are no edge DE.

are a far more elegant way to establish stability properties, but there is no systematic way of finding such functions. Therefore, as noted below, we favor numerical methods to deal with the problem, and often these methods provide further insights.

The Appendix collects arguments of a similar character that show the full set of NE is[5]

$$\{(0,1,\cdot,0),(0,0,\cdot,1),(1,p^*,\frac{w-m}{w},\cdot),(0,p^{**},\cdot,\frac{w-m}{w}),(\frac{w-m}{w},p^*,1,0)_{[\text{lw}]},$$
$$(\frac{m}{w},p^*,0,1)_{[\text{hw}]},(x,p^*,\frac{-m+wx}{wx},1),(x,p^*,\frac{-m+w}{wx},0)\}. \tag{11}$$

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 $((\frac{\partial \text{RHS eq. } i}{\partial \text{state var. } j}))$, evaluated at each NE, and compute the eigenvalues. The Jacobian of ODE system (6 - 9) is

$$J = \begin{pmatrix} \beta_\varphi(1-2\varphi)(pw-e)(q_1-q_2) & \beta_\varphi\varphi(1-\varphi)w(q_1-q_2) & \beta_\varphi\varphi(1-\varphi)(pw-e) & -\beta_\varphi\varphi(1-\varphi)(pw-e) \\ \beta p(1-p)w(q_2-q_1) & \beta(1-2p)(w-m-q_2w+\varphi w(q_2-q_1)) & -\beta w(p-p^2)\varphi & -\beta w(p-p^2)(1-\varphi) \\ 0 & \beta q_1(1-q_1)w & \beta(1-2q_1)(pw-e) & 0 \\ 0 & \beta q_2(1-q_2)(w-2\alpha) & 0 & \beta(1-2q_2)[(w-2\alpha)p+\alpha-e] \end{pmatrix}.$$

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

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

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

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

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. More intuitively, notice that the pure NE $(0,1,\cdot,0)$ and $(0,0,\cdot,1)$ can be destabilized by an

[5]Recall from footnote 3 that we only include the NE that have equal payoffs for both workers' types when both are present. For instance, the NE $(x,p^{**},0,(w-m)/(w\cdot(1-x))$ for $x \in [0,m/w]$ does not satisfy this equal payoff condition.

invasion of Type 1 worker playing W and S, respectively.

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

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

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. More concretely, any mixed NE in this face will be unstable since it is a mixed equilibrium for the two-population replicator dynamics (see eg. Weibull, 1997, Ch. 5).

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

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

whose eigenvalues are $\left\{0, \frac{\beta(w-2e)\alpha}{w}, \pm\frac{\sqrt{-\beta e(w-e)(w-m)(\beta_\varphi(x-1)(m-w)+\beta(m+w(x-1)))}}{w\sqrt{x}}\right\}$. 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 for $x \geq \frac{w-m}{w}$. Hence this NE is dynamically unstable in the high wage region but remains a DSE candidate in the low wage region.

The Appendix examines the other NE using the same techniques. It rules out DSE status for the first four in the list and confirms that there are no asymptotically stable NE. The only remaining DSE candidates are summarized by the following proposition.

Proposition 1 *The Arce (2007) game with state system $S = (\varphi, p, q_1, q_2) \in [0,1]^4$ and exogenous parameters w, e, m, α, and β_i has two dynamically stable equilibrium (DSE) candidates:*

a. $\{(x, \frac{e}{w}, \frac{wx-m}{wx}, 1) : x \in [\frac{m}{w}, 1]\}$ *in the high wage case ($w > 2e$), and*

b. $\{\left(x, \frac{e}{w}, \frac{w-m}{wx}, 0\right) : x \in [\frac{w-m}{w}, 1]\}$ *in the low wage case* ($w < 2e$).

It is worth noticing that Proposition 1 is related to the conclusions of Ok and Vega-Redondo (2001) and Dekel et al. (2007): in the models of evolution of preferences with incomplete information, only states where the aggregate play (i.e. the strategy of the Employer p and the average strategy of workers $\varphi \cdot q_1 + (1-\varphi) \cdot q_2$) corresponds to a Nash equilibrium with selfish preferences (that of Type 1 workers in Arce's game) can be stable.

2.5 Simulation results

The last step in our algorithm is to investigate convergence behavior of the DSE candidate numerically. 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 $\lim_{t\to\infty} t^{-1} \int_0^t S(u)du$, denoted by $\bar{S}(t)$, more than of $\lim_{t\to\infty} S(t)$.[6]

We solve the ODE system numerically using Mathematica.[7] We set speeds of adjustment $\beta_\varphi = 0.1$ and $\beta = 1$ and use baseline parameters $\alpha = 0.109, m = 0.08, e = 0.1, v = w^2 m$, where the high wage is $w = 0.201 > 2e$ and the low wage is $w = 0.199 < 2e$.

For these baseline parameters, the DSE candidate is $(x, 0.50, 1 - \frac{1}{2.51x}, 1), \quad x \in [0.39, 1]$ for the high wage and therefore $q_1 \in (0, 0.6)$ will depend on the level of x. initial Using initial values $\varphi(0) = 0.4$, $p(0) = 0.52$, $q_1(0) = 0.04$ and $q_2(0) = 0.96$, Figure 1 (left panel) shows the numerical results for the high incentive wage parameters. We obtain a direct convergence for q_2, meanwhile the dynamics for the remaining variables (φ, p and q_1) follow a cycle around the interior solution. We achieve convergence in time average to the relevant equilibrium. At the end of the period, the time average of the state variables are given by $\bar{S}(T) \approx (0.4033, 0.4987, 0.018, 0.9951)$.

The right panel of Figure 1 shows how the cycles amplify when the initial conditions are farther away from the equilibrium values even though we still obtain convergence in time average. In our example, the variability of q_1 decreases over time meanwhile p and φ have a roughly constant amplitude that is much higher than observed in the left panel.

[6] 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).

[7] The codes are available upon request.

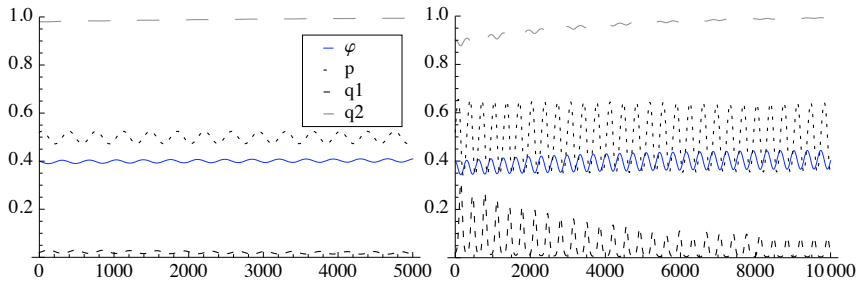

Figure 1: Dynamics in the high wage $w > 2e$ case. Left Panel shows $q_2(0) = 0.96$ and Right Panel shows $q_2(0) = 0.90$

The other relevant case is the low wage equilibrium, $(x, 0.50, \frac{1}{1.67x}, 0), x \in [0.59, 1]$ for baseline parameters. Comparing the low wage case against the high wage case, notice that the mixing probability p is not altered given that we do not change drastically the level of wage. However, the small change of incentives will affect behavior and the minimum fraction of workers' types. The type 1 worker will play close to the upper bound (before they were shirking) meanwhile type 2 worker shirks (before they were working). Also, the minimum level of type 1 workers increases from 0.39 to 0.59. Starting from values very close to the high wage case equilibrium (Panel left in Figure 1), we study the dynamics towards the new low wage equilibrium.

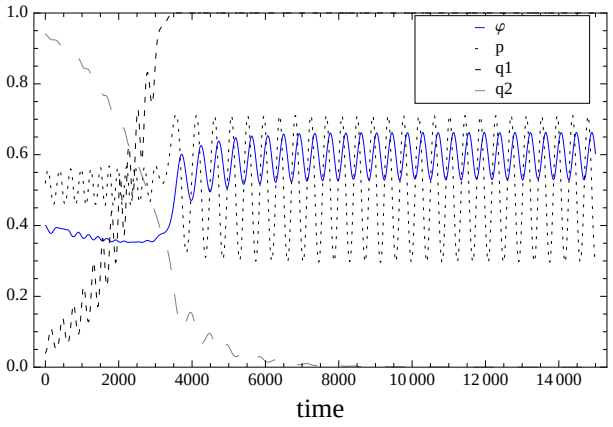

Figure 2: Dynamics in the low wage $w < 2e$ case

Figure 2 shows the simulation results. The workers' mixing probabilities evolve towards the opposite extremes since the initial values are related to the high wage case. The fraction of type 1 workers slightly decreases and then moves upward to the new equilibrium. Again, we observe cycles around the relevant equilibrium value and the time average approaches the equilibrium value

$\bar{S}(T) \approx (0.5822, 0.5052, 0.9664, 0.044)$.

Our analysis thus complements Arce's work by endogenously determining the fraction of types and studying the dynamics of the model. Our value-added includes showing that both types of workers coexist independently of the level of the incentive wage for a broad set of parameter values and initial conditions relatively close to the equilibrium values. Our numerical solution indicates Lyapunov stability, in that the dynamics when both types are present follow a cycle around the relevant interior equilibrium.

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 type Self ($v = v_H > 0$) will take revenge and, at cost v to himself, will inflict harm v/c on Other, given an exogenous marginal cost parameter $c > 0$. The equilibrium payoffs are inefficient at (0,0) when $v = 0$, but are efficient at (1,1) when $v = v_H > c$.

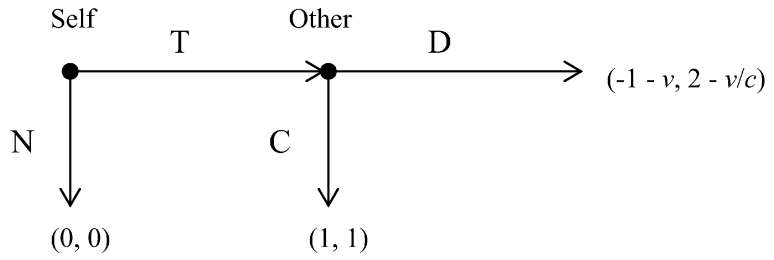

Figure 3: Simple trust game. The unique subgame perfect NE is (N,D) when $v < c$ and is (T,C) when $v > c$

3.1 Elements of the model

From this simple game, FS09 construct the noisy trust game illustrated in Figure 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. We assume that there are only two types, fixed in the short-run but variable in the long-run.[8]

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 depends negatively on the level of vengefulness (v_H):

$$a = A(v_H) = 0.5\exp(-kv_H^2) \tag{12}$$

where $k > 0$ represents a precision parameter explained in FS09.

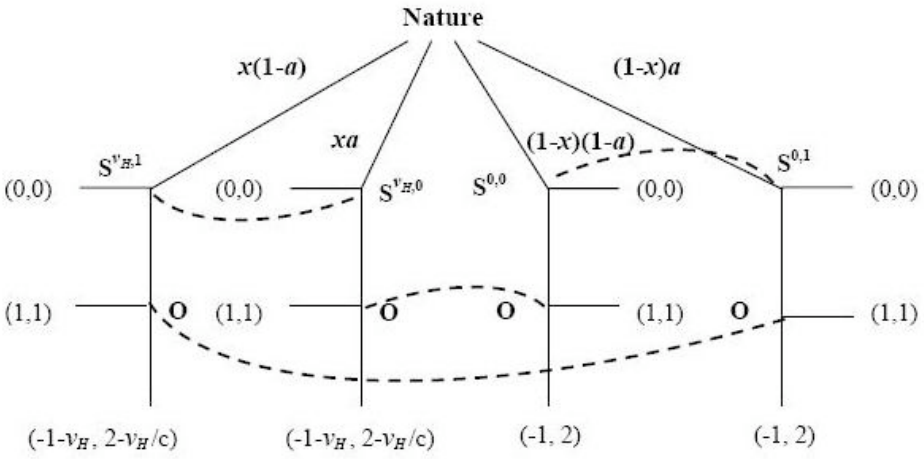

Figure 4: The noisy trust game. O denotes Other, $S^{i,j}$ denotes Self with vengeance level i and perception j, as determined by Nature's move. The four branch labels are Nature's move probabilities. Source: FS09

Let $p_1 = Pr[T|v = v_H]$ denote the probability of trusting when Self is vengeful, and $p_2 = Pr[T|v = 0]$ the probability of trusting when Self is non-vengeful. These probabilities are constrained by a tremble rate $e \in (0, 1/2)$, so that $e \leq p_0, p_1 \leq 1 - e$. Self's (mixed) strategy space is thus $[e, 1-e] \times [e, 1-e]$. Similarly, let $q_1 = Pr[C|s = 1]$ and $q_2 = Pr[C|s = 0]$ denote the probabilities of cooperating when Other observes a vengeful type and a non-vengeful type, respectively.

[8] FS09 argue informally that fitness landscape dynamics will yield degenerate distributions with support on at most two discrete points, one fixed at zero and another at some value $v_H > c > 0$ that can vary over time.

Other's strategy space is $[e, 1-e] \times [e, 1-e]$.

The state of the system is a vector $(v, x, p_1, p_2, q_1, q_2) \in [0, \hat{v}] \times [0,1] \times [e, 1-e]^2 \times [e, 1-e]^2$ that specifies Self's actions (p_1 and p_2), Other's actions (q_1 and q_2), the fraction of the vengeful type (x) and the degree of vengefulness ($v = v_H \leq \hat{v}$). The current state space is more complicated than that of the Arce model in several respects. Besides the additional mixing variable, we have a restricted mixture space (to account for trembles, which are conceptually important according to FS09) and an endogenous preference parameter.[9] Topologically, the state space is the 6-d hypercube $[0,1]^6$ with a specific parametrization.

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. More specifically, we will find different equilibria depending on the fraction of vengeful types. The pure strategy PBE equilibria lie on the 2-d faces $p_i \in \{e, 1-e\}$ and $q_i \in \{e, 1-e\}$. In terms of the log ratio $L(y) = \log(1-y)/\log(y)$ they are:

- "Separating," or $p_2 = q_2 = e$ and $p_1 = q_1 = 1-e$ when $L(c/v_H) + L(e) - L(a) \leq L(x) \leq L(c/v_H) + L(e) + L(a)$;
- "Bad Pooling," or $p_2 = p_1 = q_1 = q_2 = e$ when $L(x) \geq L(c/v_H) + L(a)$; and
- "Good Pooling," or $p_2 = p_1 = q_1 = 1-e$ and $q_2 = e$ when $L(x) \leq L(c/v_H) + L(a)$.

The mixed strategy PBE families lie on higher dimensional faces:

- "Bad Mix," or $p_2 = q_2 = e$ and $p_1, q_1 \in (e, 1-e)$ when $L(c/v_H) + L(a) \leq L(x) \leq L(c/v_H) + L(e) + L(a)$;
- "Bad Hybrid," or $p_2 = q_2 = e$, $p_1 = 1-e$ and $q_1 \in (e, 1-e)$ when $L(x) = L(c/v_H) + L(e) + L(a)$;
- "Good Mix," or $p_1 = q_1 = 1-e$ and $p_2, q_2 \in (e, 1-e)$ when $L(c/v_H) - L(a) \leq L(x) \leq L(c/v_H) + L(e) - L(a)$; and
- "Good Hybrid," or $p_2 = p_1 = q_1 = 1-e$ and $q_2 \in (e, 1-e)$ when $L(x) = L(c/v_H) - L(a)$.

[9] 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 Workers' behavior. We will see that evolving v makes good sense in the FS09 model.

The dynamic analysis will help us identify which equilibria persist in the long-run. As in the previous application, we start by constructing the expected payoffs and utilities.

3.2 Expected payoffs and utilities

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

$$w_s^v = (p_1(1-a)q_1 + p_1 a q_2)\cdot 1 + (p_1(1-a)(1-q_1) + p_1 a(1-q_2))\cdot(-1-v) \quad (13)$$

$$w_s = (p_2(1-a)q_2 + p_2 a q_1)\cdot 1 + (p_2(1-a)(1-q_2) + p_2 a(1-q_1))\cdot(-1) \quad (14)$$

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

$$w_o^s = (x(1-a)p_1 q_1 + (1-x)a p_2 q_1)\cdot(1) + (x(1-a)p_1(1-q_1))\cdot(2-v/c) + ((1-x)a p_2(1-q_1))\cdot 2 \quad (15)$$

$$w_o = (xa p_1 q_2 + (1-x)(1-a)p_2 q_2)\cdot(1) + (xa p_1(1-q_2))\cdot(2-v/c) + ((1-x)(1-a)p_2(1-q_2))\cdot 2 \quad (16)$$

Equation (13) 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 correctly perceives (probability $(1-a)$) a vengeful type and cooperates (probability q_1), and also when Other misperceives (probability a) and cooperates (probability q_2). She gets $-1-v$ when Other perceives the vengeful type correctly $(1-a)$ and defects $(1-q_1)$; and when she misperceives (a) and defects $(1-q_2)$. Similar logic yields the expressions for non-vengeful Self's payoff w_s as well as Other's possible expected payoffs w_o^s and w_o.

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 and q_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.[10] Thus the degree of vengefulness $v = v_H$ for all vengeful players changes at a rate proportional to its gradient $\frac{\partial w_s^v}{\partial v}$.

As before, we use replicator dynamics for each mixing probability p_i and q_i. Its adjustment rate is proportional to its fitness difference, which coincides with 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:

$$\dot{v} = \beta_v \left(\frac{\partial w_s^v}{\partial v} \right) \tag{17}$$

$$\dot{x} = \beta_x (1-x) x (w_s^v - w_s) \tag{18}$$

$$\dot{p}_1 = \beta (1-e-p_1)(p_1-e) \left(\frac{\partial w_s^v}{\partial p_1} \right) \tag{19}$$

$$\dot{p}_2 = \beta (1-e-p_2)(p_2-e) \left(\frac{\partial w_s}{\partial p_2} \right) \tag{20}$$

$$\dot{q}_1 = \beta (1-e-q_1)(q_1-e) \left(\frac{\partial w_o^s}{\partial q_1} \right) \tag{21}$$

$$\dot{q}_2 = \beta (1-e-q_2)(q_2-e) \left(\frac{\partial w_o}{\partial q_2} \right) \tag{22}$$

We assume as usual that p_i and q_i adjust more rapidly than does x, and that v adjusts least rapidly (perhaps via genetic disposition and/or internalized norms). Thus $0 < \beta_v < \beta_x < \beta$. To complete the dynamic specification, take the initial state as given and impose the boundary conditions $0 \le x \le 1, 0 \le v, e \le p_i \le 1-e$ and $e \le q_i \le 1-e$.

[10] The evolution of continuous biological traits is commonly modeled via gradient dynamcs (e.g., Wright (1949), Lande (1976) and Kauffman (1993)) or by Dieckmann's restricted version mentioned in the introduction. Continuous strategy sets are seen less often in economics, but there is a cluster of papers beginning with Oechssler and Riedel (2001) that applies the continuous extension of the replicator equation. However, economists going back at least to Sonnenschein (1982) have also applied gradient dynamics. Friedman and Ostrov (2010)) argue at length that gradient dynamics are more appropriate when larger changes per unit time are more difficult or expensive, while continuous-state replicator dynamics are more appropriate when adjustment is via deaths and births not spatially connected.

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:

- A "Good Hybrid" EPBE: $S = (x, v, p_1, p_2, q_1, q_2) = (x^*, v^*, 1-e, 1-e, 1-e, q_2^*)$, in which Self trusts regardless of her type and Other plays a specific mixed strategy when she perceives a non-vengeful type,[11] for the parameter values $c \in (0,1)$, $e \in (0, \hat{e}(k))$ and $k \in (0, 0.6)$,[12] and

- the "Bad Pooling" EPBE: $S = (0, v, e, e, e, e)$, for all $c > 0$, and all behavioral errors rate $e \in (0, 1/2)$, in which (apart from trembles) Self never trusts and Other always defects and v is arbitrary (and moot, since the vengeful type has population share zero).

Assuming the baseline parameters $k = 0.4$, $c = 0.5$, $e = 0.05$, $\beta_v = 0.001$, $\beta_x = 0.10$ and $\beta = 2$, the "Good Hybrid" is $(0.69, 1.67, 0.95, 0.95, 0.95, 0.87)$. 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.

3.5 Simulation results

Figure 5 shows typical numerical solutions for baseline parameters and initial conditions not far from the EPBE. The state for v, x and q_2 indeed cycles around the "good" EPBE with constant amplitude, consistent with Liouville's theorem. The right panel of Figure 5 confirms convergence in time average. In the simulations, the remaining mixing probabilities p_1, p_2 and q_1 adjust quite rapidly to the upper extreme $(1-e)$; meanwhile v hardly moves since it starts with an initial value close to the equilibrium and its adjustment rate is, by assumption, very slow.

[11] FS09 presents the conditions that x^*, v^* and q_2^* should satisfy. In our dynamic system, we can obtain similar conditions considering that the gradient has to be zero for the dynamic equation of v and q_2 and that both types get the same payoff.

[12] $\hat{e}(k)$ is given by $R(k)/(2-2a+2R(k))$ where $R(k) = (kv(1+v/2)-1)a$

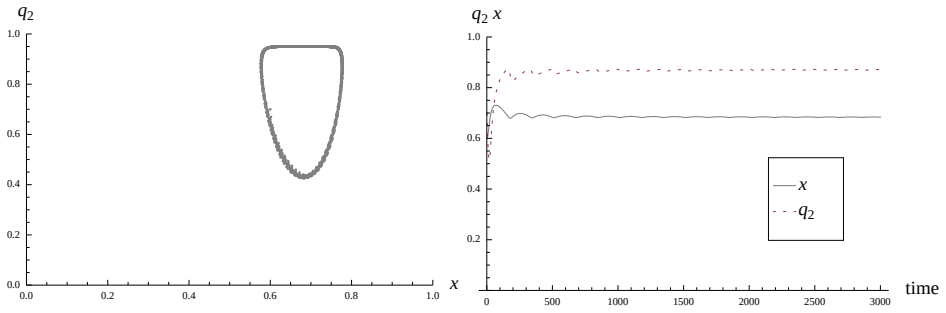

Figure 5: Dynamics (Panel left) and Time-Average dynamics (Panel right) of x and q_2

The dynamic convergence to the good EPBE is not due to an arbitrary selection of parameters. From a wide range of parameters consistent with the definition of the good EPBE and initial values sufficiently near the equilibrium values, we obtain convergence in time average. For baseline parameters we have confirmed convergence from initial states $v^* - 0.04 \leq v(0) \leq v^* + 0.01$, $x^* - 0.54 \leq x(0) \leq x^* + 0.10$, and for individual mixing probabilities $e \leq p_i(0) \leq 1 - e$ and $e \leq q_i(0) \leq 1 - e$. If we simultaneously change all p_i probabilities and q_i probabilities, to achieve the relevant equilibrium we need that $q_i(0) > 0.7$.

The dynamic analysis has two caveats. First, notice that in several cases, we do not drastically alter the initial state. Thus, our analysis focuses on local stability. Second, we must restrict the adjusment 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, from our previous analysis, when we start from $x(0) < 0.15$ (not many vengeful types) and $q_i(0) < 0.35$ (a low probability that Other cooperates) the bad EPBE persists in the long-run.

4 Discussion

We have analyzed the dynamic stability of two games of incomplete information in the context of the evolution of preferences. We complement Arce’s (2007) results by endogenously determining the fraction of worker types and studying the dynamics of the state variables. We show that both

types of workers coexist independently of the level of incentive wage. The second example is a noisy trust game due to FS09. Here we add dynamics to their static EPBE concept and numerically illustrate a 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 for all feasible states. Then it applies standard evolutionary concepts to describe the evolution of types (preference parameters in our examples), their population shares, and action mixtures. It uses gradient dynamics for a continuous space of types, and uses replicator dynamics for the rest of the state vector, the population shares and mixture probabilities. (In passing, we note that replicator dynamics for the mixture probabilities are equivalent to gradient dynamics modified by a binomial variance factor.) The result is a system of ordinary differential equations (ODEs) in applications like those just analyzed. (When a continuum of active types is possible, the result can include a partial differential equation.)

In view of the fact that asymptotic stability cannot be expected in key equilibria of games of incomplete information, the toolbox emphasizes convergence in time average and includes 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.

Two remarks on modeling philosophy may be helpful to applied economists.

- The toolbox presented here omits some of the more advanced techniques from dynamical systems theory, such as center manifold techniques or bifurcation techniques, because we expect these will yield a lower return on applied researchers' investment. It also omits Lyapunov functions, since we can offer no systematic way of finding them.

- Mainstream analysis of games of incomplete information typically exogenously specifies the set of active types and the population shares of those types. That seems to us to push off stage

the most interesting part of the story. Hence our toolbox emphasizes methods for describing the evolution of these state variables and for characterizing their long-run behavior. As illustrated in the FS09 model, endogenizing the set of types can resolve the multiplicity of equilibria and lead to sharper predictions in applications.

Our exposition brings into focus several open theoretical questions. For example, are the static refinements CSS or NIS sufficient or necessary for Lyapunov or asymptotic stability under various specifications of dynamics? To what extent do our toolbox techniques survive for multidimensional type spaces or action spaces? We hope that our presentation encourages evolutionary theorists to investigate these and other open questions for games of incomplete information.

Appendix: Mathematical Details

Finding DE and NE in the Arce (2007) Model

Recall that section 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 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 (6) 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 (7) implies that $\varphi^* = \frac{m/w+q_2-1}{q_2-q_1}$. Ruling out $q_2 - q_1 = 0$,[13] we see from (6) 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 (9). The case $w - 2e > 0$ is called high incentive wages, and yields the $q_2^* = 1$ 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 (9), 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 (6). 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.

Evaluating the Jacobian matrix at the NE

The text analyzed stability for the first three NE and the last NE listed in (11). In this section, we find Jacobian matrices and eigenvalues for the remaining NE.

[13] Notice that if both mixing probabilities are pure and $q_2 - q_1 = 0$, the best reply p is also pure and thus the dynamics is not on a 2-d face.

The Jacobian matrix for (6 - 9) evaluated at the equilibrium $(\varphi, p, q_1, q_2) = (0, 1, \cdot, 0)$ is

$$J = \begin{pmatrix} \beta_\varphi q_1(w-e) & 0 & 0 & 0 \\ 0 & -\beta(w-m) & 0 & 0 \\ 0 & \beta(1-q_1)q_1 w & \beta(1-2q_1)(w-e) & 0 \\ 0 & 0 & 0 & -\beta(\alpha-(w-e)) \end{pmatrix},$$

whose eigenvalues are $\{-\beta(\alpha-(w-e)), -\beta(w-m), \beta_\varphi q_1(w-e), \beta(1-2q_1)(w-e)\}$. 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

$$J = \begin{pmatrix} \beta_\varphi e(1-q_1) & 0 & 0 & 0 \\ 0 & -\beta m & 0 & 0 \\ 0 & \beta(1-q_1)q_1 w & \beta e(-1+2q_1) & 0 \\ 0 & 0 & 0 & -\beta(\alpha-e) \end{pmatrix},$$

whose eigenvalues are $\{-\beta(\alpha-e), -\beta m, \beta_\varphi e(1-q_1), \beta e(-1+2q_1)\}$. 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

$$J = \begin{pmatrix} 0 & 0 & \beta_\varphi(w-e)(1-\varphi)\varphi & -\beta_\varphi(w-e)(1-\varphi)\varphi \\ 0 & \beta m & 0 & 0 \\ 0 & 0 & -\beta(w-e) & 0 \\ 0 & 0 & 0 & \beta(\alpha-(w-e)) \end{pmatrix},$$

whose eigenvalues are $\{0, \beta(\alpha-(w-e)), \beta m, -\beta(w-e)\}$. The second and third are positive, so this equilibrium is not a DSE. Notice that this result also follows from the fact that $q_2 = 1$ is not a best-reply for $p = 1$.

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

$$J=\begin{pmatrix} -\frac{\beta_\varphi(w-2e)(m-(1-q_1)w)\alpha}{w(w-2\alpha)} & 0 & 0 & 0 \\ \frac{-\beta(m-(1-q_1)w)(\alpha-(w-e))(\alpha-e)}{(w-2\alpha)^2} & 0 & 0 & \frac{-\beta w(\alpha-(w-e))(\alpha-e)}{(w-2\alpha)^2} \\ 0 & \beta(1-q_1)q_1 w & \frac{-\beta(-1+2q_1)(w-2e)\alpha}{2\alpha-w} & 0 \\ 0 & -\frac{\beta m(w-m)(2\alpha-w)}{w^2} & 0 & 0 \end{pmatrix},$$

with eigenvalues $\left\{\pm\sqrt{\frac{\beta^2 m(w-m)(\alpha-e)(\alpha-(w-e))}{w(2\alpha-w)}},\frac{\beta(1-2q_1)(w-2e)\alpha}{2\alpha-w},\frac{-\beta(w-2e)(m-(1-q_1)w)\alpha}{w(w-2\alpha)}\right\}$. The first pair is real with opposite signs, so this NE is not a DSE. This result is along with the notion that a mixed equilibrium is unstable in the two-population replicator dynamics, see Weibull (1997, Ch. 5).

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

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

with eigenvalues $\left\{0,\pm\frac{\sqrt{-\beta^2 em(w-m)(w-e)}}{w},\frac{\beta(-1+2q_2)(2e-w)\alpha}{w}\right\}$. The second eigenvalue is imaginary meanwhile the third can be negative as long as the wage corresponds to the low wage case ($w<2e$) and $q_2<1/2$ or the wage is set in the high wage case and $q_2>1/2$. Hence this NE remains a candidate DSE, requiring further investigation.

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

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

with eigenvalues $\left\{0,\frac{\beta(w-2e)\alpha}{w},\pm\sqrt{\frac{-\beta_\varphi\beta em(w-m)(w-e)}{w}}\right\}$. 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 (11) and already analyzed in the text.

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

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

with eigenvalues $\left\{0, \frac{-\beta(w-2e)\alpha}{w}, \pm\sqrt{\frac{-\beta_\varphi \beta em(w-m)(w-e)}{w}}\right\}$. 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

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

with eigenvalues $\left\{0, \frac{-\beta(w-2e)\alpha}{w}, \pm\frac{\sqrt{\beta em(w-e)\left(m(-1+\varphi^*)\beta_\varphi+(m-w\varphi^*)\beta\right)}}{w\sqrt{\varphi^*}}\right\}$. 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.

5 Acknowledgements

We thank the co-editor Frank Riedel, two anonymous referees, Dann Arce and Bill Sandholm for their comments and suggestions that significantly improved our paper.

References

[1] Amann E, Possajennikov A (2009) On the stability of evolutionary dynamics in games with incomplete information. Mathematical Social Sciences 58: 310-321

[2] Arce D (2007) Is agency theory self-activating? Economic Inquiry 45 (4): 708-720.

[3] Apaloo J (1997) Revisiting strategic models of evolution: the concept of neighborhood invader strategies. Theoretical Population Biology 52:71-77

[4] Benaim M, Hofbauer J, Hopkins E (2009) Learning in games with unstable equilibria. Journal of Economic Theory 144(4):1694-1709

[5] Cressman R (2003) Evolutionary dynamics and extensive form games. MIT Press

[6] Cressman R (2005) Stability of the replicator equation with continuous strategy space. Mathematical Social Sciences 50:127-147

[7] Dekel E, Ely JC, Yilankaya O (2007) Evolution of Preferences. Review of Economic Studies 74(3):685-704

[8] Doebeli M, Hauert C (2005) Models of cooperation based on the Prisoner's Dilemma and the Snowdrift game. Ecology letters 8:748-766

[9] Ely JC, Sandholm W (2005) Evolution in Bayesian Games I: Theory. Games and Economic Behavior 53:83-109

[10] Eshel I (1983) Evolutionary and continuous stability. Journal of Theoretical Biology 103:99-111

[11] Friedman D (1991) Evolutionary games in economics. Econometrica 69: 637-666

[12] Friedman D, Ostrov D (2010) Gradient Dynamics in Population Games: Some Basic Results. Journal of Mathematical Economics 46(5): 691-700

[13] Friedman D, Ostrov D (2013) Evolutionary Dynamics over Continuous Action Spaces for Population Games that Arise from Symmetric Two-player Games. Journal of Economic Theory 148(2):743-777.

[14] Friedman D, Singh N (2009) Equilibrium vengeance. Games and Economic Behavior 66:813-829

[15] Fudenberg D, Levine D (1998) The Theory of Learning in Games. MIT Press

[16] Güth W, Yaari M (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

[17] Hirsch M, Smale S (1974) Differential Equations, dynamical system and linear algebra. Academic Press

[18] Hofbauer J, Sigmund K (1988) The Theory of Evolution and Dynamical Systems. Cambridge University Press

[19] Kauffman S (1993) The Origins of Order: Self-Organization and Selection in Evolution. NY: Oxford U Press

[20] Lande R (1976) Natural Selection and Random Genetic Drift in Phenotypic Evolution. Evolution 30(2):314-334

[21] Oechssler J, Riedel F (2001) Evolutionary dynamics on infinite strategy spaces. Economic Theory 17:147-162

[22] Oechssler J, Riedel F (2002) On the dynamic foundation of evolutionary stability in continuous models. Journal of Economic Theory 107:223-252

[23] Ok R, Vega-Redondo F (2001) On the evolution of individualistic preferences: An incomplete information scenario. Journal of Economic Theory 97:231-254

[24] Possajennikov A (2005) Cooperation and Competition: Learning of Strategies and Evolution of Preferences in Prisoner's Dilemma and Haw-Dove games. International Game Theory Review 7(4):443-459

[25] Sonnenschein H (1982) Price dynamics based on the adjustment of firms. American Economic Review 72(5):1088-1096

[26] Sandholm W (2010) Population games and evolutionary dynamics. MIT Press

[27] Taylor PD, Jonker LB (1978) Evolutionary stable strategies and game dynamics. Mathematical Biosciences 40:145-156

[28] Weibull W (1997) Evolutionary game theory MIT Press

[29] Wright S (1949) Adaption and Selection. In Jepsen L, Simpson GG, Mayr E (eds) Genetics, Paleontology, and Evolution. Princeton, N.J.: Princeton University Press