View metadata, citation and similar papers at core.ac.uk
brought to you by CORE
provided by Research Papers in Economics

Penn Institute for Economic Research
Department of Economics
University of Pennsylvania
3718 Locust Walk
Philadelphia, PA 19104-6297
pier@econ.upenn.edu
http://www.econ.upenn.edu/pier

PIER Working Paper 08-032

"Inference in a Synchronization Game with Social Interactions Second Version"

by

Áureo de Paula

http://ssrn.com/abstract=1256364

Inference in a Synchronization Game with Social Interactions *

Áureo de Paula †

First Version: October 2004
This version: August 2008

Abstract

This paper studies inference in a continuous time game where an agent's decision to quit an activity depends on the participation of other players. In equilibrium, similar actions can be explained not only by direct influences but also by correlated factors. Our model can be seen as a simultaneous duration model with multiple decision makers and interdependent durations. We study the problem of determining the existence and uniqueness of equilibrium stopping strategies in this setting. This paper provides results and conditions for the detection of these endogenous effects. First, we show that the presence of such effects is a necessary and sufficient condition for simultaneous exits. This allows us to set up a nonparametric test for the presence of such influences which is robust to multiple equilibria. Second, we provide conditions under which parameters in the game are identified. Finally, we apply the model to data on desertion in the Union Army during the American Civil War and find evidence of endogenous influences.

JEL Codes: C10, C70, D70.

*This article was part of my dissertation at Princeton University. I am particularly indebted to José Scheinkman, Bo Honoré and Elie Tamer for their encouragement, guidance and support throughout the development of this project. I would like to thank Fernando Alvarez, Joseph Burton, Vinicius Carrasco, Albina Danilova, Savas Dayanik, Robert Fogel, Rodrigo Guimarães, Ken Hendricks, Chris Julliard, James McPherson, Martin Oehmke, Marcelo Pinheiro, Marcelo Rezende and Matt Weinberg for invaluable discussions. I also benefited from comments by seminar participants at CEMFI, Chicago (Economics and CPE), Chicago Fed, Cornell, Illinois (Urbana-Champaign), LAMES (2004), NYU, Penn, Princeton and PUC-Rio. Financial support from the Fellowship of Woodrow Wilson Scholars, NSF grant SES-0001647 and NIH/NIA grant P01 AG10120 is gratefully acknowledged. All remaining errors are my own.

†Department of Economics, University of Pennsylvania, Philadelphia, PA 19104-6209. E-mail: aureo@sas.upenn.edu

1. Introduction

In this paper we set up a continuous time model to describe a multi-person decision problem of timing coordination. Individual strategies are exit (or entry) times to a certain activity such as when to join a social welfare program, desert from an army or emigrate to a different region. After characterizing the equilibrium for such a situation, we assess the empirical implications of the model in the presence of direct strategic effects of a player's action on other agents' choices. The main finding is that such endogenous effects are necessary and sufficient for simultaneous exits with positive probability in the proposed environment. This has consequences for the statistical treatment of such settings and for inference. We also show that in this model the number of players impacts observed (equilibrium) outcomes only in the presence of endogenous effects. We then devise a test for the existence of endogenous effects taking into account the fact that time is not observed continuously but at discrete intervals. The paper subsequently analyzes circumstances under which parameters of interest are identified. Finally, we illustrate the application of these tools with an analysis of desertion in the Union Army during the American Civil War.

It is difficult to explain why agents behave similarly when they do so. Individuals may act similarly in response to correlated shocks or genuinely in reaction to each other's actions — a legitimate endogenous effect.[1] We analyze a situation in which agents take a binary action and choose the timing for such an action. Crucially, correlated behavior may arise through correlated effects or through a direct impact on others.

One reason to properly account for endogenous effects is that they might have different implications for policy than correlated effects. Endogenous effects may create "social multipliers" and blow up the effect of other factors determining behavior.[2] This may significantly alter the choice of treatments in policy-relevant situations like crime reduction, welfare program participation or immigration. Imagine, for instance, a situation in which agents choose when to join a certain welfare program. A person's timing may be determined by common factors or directly by the timing of other agents' decisions (or both). If the participation of one's reference group — the endogenous effect — is a sufficiently strong determinant for an agent's choice, one could concentrate efforts on a subgroup of the community and hope to affect the remaining members as the focus group joins the targeted activity. If on the other hand the main driver is common shocks that provoke participation, a policy-maker may prefer to identify and directly act on such defining variables.

[1]Manski (1993) provided a clear categorization for the possible causes of similarities in behavior and coined the expression "reflection problem" to characterize the difficulties in separating endogenous and other social and correlated effects.

[2]For a recent exposition on this issue, see Glaeser and Scheinkman (2002).

Manski (1993) notes that the identification of endogenous social effects in a static context is difficult. Nevertheless, the introduction of dynamics allows for identification if group actions influence individuals with a lag. In this paper, even though the distinction between contemporaneous and adjacent periods in a continuous time environment becomes negligible as the two coalesce, the endogenous effects can still be identified since the relevant payoff variables are assumed to evolve smoothly (according to a diffusion process). Hence this article identifies a particular class of environments for which the identification question can be solved. The adequacy of our assumptions will depend on the particular empirical application at hand. In particular, the diffusive behavior of state variables is an important consideration. Examples of empirical analyses (with a focus on individual decision making) where state variables are traditionally modeled via diffusion processes can be found, for instance, in Dixit and Pindyck (1994) and include applications such as investment in offshore oil reserves and participation in securities markets. Building intuition for continuous time settings as limits of discrete time environments, Merton (1990) (Chapter 2) provides conditions under which per period discrete time innovations produce continuous sample paths in continuous time.[3] In fact, traditional discrete time dynamic discrete choice models with normal preference shocks would in the limit (as frequency increases) produce diffusive utility processes (e.g., the works surveyed in Eckstein and Wolpin (1989)). Consequently, at high enough frequency those models are well approximated by continuous time models where the state variable of interest — the utility process — evolves continuously and the adequacy of the discrete time analog would imply the appropriateness of our model. It is important to bear in mind though that our methodology provides no "silver bullet" and its suitability should be determined in accordance with the application.

In a timing framework, statistical inference typically involves survival analysis or duration models. Whereas standard statistical duration models could be employed to identify the existence of hazard dependence among agents (as indeed is done in Costa and Kahn (2003) and Sirakaya (2001) and suggested in Brock and Durlauf (2001)), it is still unclear whether such effects are primarily due to endogenous influences or to correlated unobservables. In contrast, our model clearly separates both channels and lays out the circumstances under which each of these sources is individually identifiable. Another issue that arises in the particular setting we study — timing problems — is that endogenous effects generate simultaneous actions with positive probability in continuous time. This is an outcome that does not occur in standard duration models. Failure to properly account for such phenomena may bias estimation and misguide inference.[4]

[3] In Merton's characterization, continuous sample paths arise even when certain types of "rare events" are allowed.

[4] See, for instance, van den Berg, Lindeboom, and Ridder (1994).

Applications for the above tools comprise all those circumstances that focus on timing coordination and would involve "duration"-type models with multiple agents. One may cite, for instance, participation in a social welfare program, stock market participation (Hong, Kubik, and Stein (2004)), migration (Orrenius (1999)) and even crime recidivism (see, for instance, the empirical investigation by Sirakaya (2001), where social interactions are found to meaningfully affect recidivism among individuals on probation).

This paper contributes to the econometric literature on social interactions. At the same time, it borrows standard tools used in the finance and investment literatures. We review the relevant literature in the following subsection.

1.1. Literature Review

In this paper we provide a model for timing coordination. Early references to such situations can be found for instance in Schelling (1960), which discusses the timing of mob formation. Our paper also relates to the threshold models of collective behavior in Granovetter (1978), for which *"the costs and benefits to the actor of making one or the other choice depend in part on how many others make which choice."* Although that paper focuses on the binary nature of the actions taken, a timing element exists in many of the examples gathered (diffusion of innovations, strikes, leaving social occasions, migration and riot formation). We formalize these ideas using tools of continuous time probability models in which individuals choose an optimal timing strategy to quit (or join) a certain activity. Our theoretical model is also connected to the one developed in Mamer (1987) for a discrete time setting and in a different framework.[5] As a result, our model is in the family of stochastic differential games — continuous time situations in which the history is summarized by a certain state-variable. This literature is more concerned with zero-sum games, whereas we focus on situations involving coordination elements. Our theoretical model can also be related to the continuous time game presented in Hopenhayn and Squintani (2004). In their case the payoff flows evolve discontinuously, whereas in our case the utility flow is continuous with probability one. This distinction diminishes the role of beliefs with respect to the opponent players in our case and turns out to be an important simplifying element in our analysis. As is outlined later in the paper, the continuity of payoff flows is also a crucial identifying assumption once we focus on the empirical content of the model.

In simple contexts it is usually difficult to separate endogenous effects from other social forces (Manski (1993)). This difficulty explains our search for structure in the context under analysis. We consider a continuous time model in which utility evolves smoothly. The continuous

[5] In his paper this author is mostly concerned with research and development investment applications in which firms have complementary decisions.

time assumption will be appropriate in circumstances where agents (are allowed to) revise their decisions frequently and observations (by the econometrician) are not coarse. The continuity of sample paths imposes structure on the stochastic shocks in the environment. As stressed earlier, both requirements may or may not be empirically valid approximations. When applicable, they allow us to identify endogenous effects from other correlated effects. Strategic interactions also pose an additional problem that may hinder identification and estimation: that of multiple equilibria.

We use the tools of continuous time optimal stopping problems that appear in the investment and finance literatures (see Dixit and Pindyck (1994)). Whereas studies in this literature do address the interaction of many agents, what distinguishes our model is a clear separation between endogenous and correlated effects.

Our paper is also related to the empirical literature on "duration-type" situations with many interacting agents. One example is Sirakaya (2001), in which the author investigates duration dependence in the timing of crime recidivism. Brock and Durlauf (2001) cite other applications, such as the timing of out-of-wedlock births or first sexual experience. Still, the studies indicated there do not look at the endogenous effect but focus instead on contextual neighborhood variables. In their analysis of group homogeneity and desertion, Costa and Kahn (2003) discuss the possibility of a contagion effect and try to account for it by introducing the fraction of deserters in a military company as a regressor (p. 538). Although this is indicative of endogenous interactions, without a structural representation it is still not clear whether this is due to endogenous or correlated effects.

The structure of the paper is as follows. In the next section we present the general model and establish the existence of equilibrium. Section 3 discusses and characterizes a particular specification for the model and sets the scene for Section 4, in which we discuss the empirical implications of the model. In Section 5 we illustrate the previous discussion with a dataset comprising Union Army recruits during the American Civil War. We obtain evidence that there were endogenous effects involved in the decision to desert the army and estimate the model by simulation methods. The final section concludes.

2. The Model

As a mathematical model of the world, consider a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ in which a given state $\omega \in \Omega$ is chosen according to a probability law $\mathbb{P}$. There are I agents who take part in a certain activity (we will loosely use I to denote the set of agents and its cardinality). A gain function ($u_i : \mathbb{R} \times [0, T] \to \mathbb{R}$) captures the utility an individual derives as he or she exits the activity. If an agent $i \in I$ leaves at a time $\tau^i \in [0, T] (T \in \overline{\mathbb{R}}_{++}$, where $\overline{\mathbb{R}}_{++} = (0, \infty])$, he or she collects a reward of $u_i(x^i_{\tau^i}, \tau^i)$ where x^i_t is an individual-specific state variable (e.g., wealth, utility-relevant inputs) and $x^i_{\tau^i}$ is this variable evaluated at the chosen τ^i. The stopping strategies are represented by $\tau_i : \Omega \to [0, T]$, a (possibly infinite) stopping time with respect to an individual filtration $\mathbb{F}^i = (\mathcal{F}^i_t)_{t\in[0,T]}$[6] representing agent i's flow of information. Although this information flow arises endogenously in the game, we assume throughout that the individual filtration satisfies the usual conditions.[7] We allow the individual information histories to differ across individuals. These individual information sequences will be the basis for an agent's strategy, since the filtration $\mathbb{F}^i = (\mathcal{F}^i_t)_{t\in[0,T]}$ incorporates the assumptions imposed on what each agent knows or not as time evolves. Later on we assume that each agent observes his or her own state variable x^i_t and whether or not other agents in the game have stopped (but not their individual state variables).

We assume that the individual state variable evolves as a process (adapted to the $\mathbb{F}^i = (\mathcal{F}^i_t)_{t\in[0,T]}$ filtration) that may depend directly on the participation of the remaining individuals in the group. This direct influence represents the endogenous effects in our model. Let θ^i_t be the fraction of the population (excluding agent i) that has left before time t: $\theta^i_t = \sum_{s=1,s\neq i}^{I} \mathbb{I}_{\{\tau_s<t\}}/(I-1)$ (with $\mathbb{I}_{\{A\}}$ as the indicator function for the event $A \subset \Omega$). This process will be determined endogenously as individuals choose the stopping times. Throughout we assume that $\theta^i_t \in \mathcal{F}^i_t$: one knows how many players have stopped up to (but excluding) the current instant. Each individual state variable x^i_t is assumed Markovian and is allowed to differ across individuals. The structure for the multi-person problem (payoffs, players, strategy spaces and information assumptions) is as follows.

Definition 1 (Synchronization Game) *A* Synchronization Game *is defined as a tuple* $\langle I, (\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P}, (u_i)_{i\in I}, (x^i)_{i\in I}, (T_i)_{i\in I}\rangle$ *where* I *is the set of agents;* $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$*, a filtered probability space;* $u_i : \mathbb{R} \times \mathbb{R}_+ \to \mathbb{R}$*, an individual gain (utility) function;* x^i*, an individual adapted process having as state space* $\mathbb{R}_+$*; and* T_i*, a set of stopping strategies* $\tau : \Omega \to [0, T]$*.*

Each person i faces the following (individual) optimal stopping problem (where τ generically denotes a stopping time with respect to $(\mathcal{F}^i_t)_{t\in[0,T]}$):

$$\begin{cases} & \mathbb{P}(x^i_t \in \Gamma | \mathcal{F}^i_s), \text{ Markovian.} \\ V_i(x_i) = \sup_{\tau\in T_i} \mathbb{E}_{x_i}[u_i(x_\tau, \tau)] \quad \text{s.t.} & \theta^i_t = \sum_{s=1,s\neq i}^{I} \mathbb{I}_{\{\tau_s<t\}}/(I-1) \\ & x^i_0 = x_i \end{cases} \tag{1}$$

[6]A random variable $\tau : \Omega \to [0, T]$ is a stopping time with respect to $(\mathcal{F}_t)_{t\in[0,T]}$ if, for each $t \in [0, T]$, $\{\omega : \tau(\omega) \leq t\} \in \mathcal{F}_t$. Intuitively they represent stopping strategies that rely solely on past information.

[7]The filtration is right-continuous and $\mathcal{F}_0$ contains all $\mathbb{P}$-negligible sets in $\mathcal{F}$.

In the above definition, $\mathbb{E}_{x_i}[u_i(x^i_\tau,\tau)] = \int_\Omega \mathbb{P}(d\omega)u_i(x^i_{\tau(\omega)}(\omega),\tau(\omega))$ with initial condition given by x_i. We assume that $u_i(x_\infty(\omega),\infty) = \limsup_{t\in[0,T]} u_i(x_t(\omega),t)$. The state variable follows:

$$dx^i_t = \alpha^i(x^i_t,\theta^i_t,t)dt + \sigma^i(x^i_t,\theta^i_t,t)dW^i_t, \qquad x^i_0 \sim F^i_0 \tag{2}$$

where W^i_t is a Wiener process and the drift and dispersion coefficients are positive Borel-measurable functions. The initial distribution F^i_0 is furthermore independent of the Brownian motion W^i_t. There are no restrictions on the contemporaneous correlation between the Wiener processes, which account for the correlated effects.

Notice that the state variable has continuous sample paths ($\mathbb{P}$-a.s.). This allows us to treat individual beliefs about the position of a counterpart's state variable conveniently (in contrast, for instance, to the work by Hopenhayn and Squintani (2004), where sample paths present discontinuities). Since the stochastic utility processes evolve continuously, the probability that a given individual reaches a stopping region between t and $t+\epsilon$ conditional on not having stopped before t vanishes as $\epsilon \to 0$. Consequently, where a counterpart's state variable is located becomes immaterial to the decisions taken within the next infinitesimal period. When empirically suitable, the model leads to the identification of endogenous effects that may otherwise be difficult to pin down (see the discussion in Manski (1993), for example). This is important because the effect of alternative counterfactual interventions such as those briefly alluded to in the introduction may be significantly affected by the presence or absence of endogenous effects.

The following conditions ensure that, given a profile of stopping times for each player, this stochastic differential equation has a (strong) solution:

Assumption 1 (Lipschitz and Growth Conditions) *The coefficients $\alpha^i(x,\theta,t)$ and $\sigma^i(x,\theta,t)$ satisfy the global Lipschitz and linear growth conditions:*

$$\|\alpha^i(x,\theta,t)-\alpha^i(y,\theta,t)\| + \|\sigma^i(x,\theta,t)-\sigma^i(y,\theta,t)\| \le K\|x-y\| \tag{3}$$

$$\|\alpha^i(x,\theta,t)\|^2 + \|\sigma^i(x,\theta,t)\|^2 \le K^2(1+\|x\|^2) \tag{4}$$

for every $t\in[0,T], x,y\in\mathbb{R},\theta\in[0,1]$ and $i\in I$, where K is a positive constant.

Notice that $\theta^i_t = \sum_{s=1,s\neq i}^{I} \mathbb{I}_{\{\tau_s<t\}}/(I-1)$ is adapted since θ is the aggregation of indicator functions of events such as $\{\tau < t\}$, where τ is an optional time with respect to the individual filtration. Given the Borel-measurability conditions on the drift and dispersion coefficients, this guarantees that, for fixed x, $(t,\omega)\mapsto \alpha^i(x,\theta^i_t(\omega),t)$ and $\sigma^i(x,\theta^i_t(\omega),t)$ are adapted. The above assumptions guarantee the existence of a strong solution for the stochastic differential equation (2). A sketch for the proof is presented in the Appendix.

2.1. Existence of Equilibrium

The solution concept we seek for this group situation is that of mutual best responses: a standard Nash Equilibrium point. The equilibrium strategies are then a vector of I stopping times such that each individual stopping time is optimal given the stopping rules adopted by the other agents.[8] Denoting by $\tau = (\tau_i)_{i \in I}$ a stopping time profile, let $U_i(\tau) = \mathbb{E}_{x_i}[u_i(x_{\tau_i}, \tau_i)]$ subject to the above transition laws and initial conditions and evaluated at the strategy profile τ. We also adopt the convention of using τ_{-i} as shorthand notation for $(\tau_s)_{s \in I - \{i\}}$. To proceed with the analysis of equilibrium, we make the following assumptions:

Assumption 2 (Exponential Discounting) *Let* $u_i(x,t) = e^{-\gamma_i t} g_i(x), \gamma_i > 0, g_i : \mathbb{R}_+ \to \mathbb{R}, \forall i \in I$. *We refer to* $g_i(\cdot)$ *as the reward function.*

Assumption 3 (Reward Function) *The individual reward functions* $g_i(\cdot), \forall i \in I$ *are such that: (Monotonicity)* $g_i(\cdot)$ *is increasing; (Convexity)* $g_i(\cdot)$ *is convex;* $\mathbb{E}[\sup_{t \in [0,T]} |e^{-\gamma_i t} g_i(x_t^i)|] < \infty$; *(Twice Differentiability)* $g_i(\cdot)$ *is twice differentiable; and (Bounded Derivative) the derivative* $g'(\cdot)$ *is bounded.*

Assumption 4 (Bound on Volatility) *For each* $t < \infty$ *and feasible profile of stopping strategies, the dispersion coefficient is assumed to satisfy:*

$$\mathbb{E}[\int_0^t (e^{-\rho s} \sigma(x_s, \theta_s, s))^2 ds] < \infty.$$

Assumption 5 (Complementarity) *The drift and the dispersion coefficients are assumed to be decreasing on their second argument:* $\partial_\theta \alpha(\cdot,\cdot,\cdot) \leq 0$ *and* $\partial_\theta \sigma(\cdot,\cdot,\cdot) \leq 0$.

The Exponential Discounting Assumption (2) simplifies the manipulation and is standard. The set of assumptions regarding the reward functions, (3), encompasses monotonicity and convexity, which are not very controversial either (convexity is not necessary if σ does not depend on θ, for instance); a boundedness condition, employed to assert the existence of a solution for the optimal stopping problem, and technical assumptions that facilitate the application of existing results in the comparison of solutions for stochastic differential equations. The Bound on Volatility Assumption (4) implies that changes in the profile of stopping decisions affect the objective function only through the drift of the discounted gain function. Finally, the Complementarity Assumption

[8] Since the strategies depend on information generated by the state variables and these are Markovian and since optimization follows Bellman's principle of optimality in dynamic programming — whatever the initial state and decisions are, the remaining decisions must be optimal with regard to the state resulting from the first decision — these are also Markov perfect equilibria.

(5) expresses the idea that higher participation makes the activity more attractive as well as increases the volatility of the returns. This assumption means that one agent's action is a strategic complement to the others' actions. We are now ready to state the following result (the proof is available upon request; all the other proofs are given in the appendix):[9]

Theorem 1 (Existence) *Under Assumptions 1-5, the* Synchronization Game *has a nonempty set of equilibrium points and this set possesses a maximal element.*

Under such general conditions, little can be said regarding uniqueness and other properties. In the next section we make further assumptions on the structure of the game.

3. A Coordination Game

We now specialize the model to extend the analysis. Consider initially a game where agents contemplate the possibility of exit. As before, a state variable x represents the latent utility a player collects when abandoning a certain activity. At exit, he or she pays a cost C. The strategy is then a rule dictating his or her exit decision using the available information at the time. Given a discount rate γ, the objective for the agent is to maximize $\mathbb{E}^x[e^{-\gamma t}(x_t - C)]$.

At an initial stage consider the individual problem where the state variable x follows:

$$dx_t = \begin{cases} \alpha x_t dt + \sigma x_t dW_t & \text{if } t \leq \nu \\ (\alpha - \Delta\alpha) x_t dt + \sigma x_t dW_t & \text{if } t > \nu \end{cases}$$

where $\Delta\alpha \geq 0$ and ν is an exogenously given random time. We assume that the individual observes W_t and whether or not the random time occurred up to (but *excluding*) time t: $\mathcal{F}_t = \sigma(x_s, \mathbb{I}_{\nu<s}, s \leq t)$. The initial condition is drawn from an independent distribution F_0 as in equation (2). The break point for the drift here is exogenously given. At a later stage we will endogenize this stopping time to make it dependent on the decision by the other participants. For there to be a well-defined solution to this problem, we assume that $\gamma > \alpha$. If γ is too low, a patient agent will postpone the switching decision indefinitely and stopping times may not be finite (see, for instance, Dixit and Pindyck (1994), p. 138).

Let $\overline{x}$ be the process corresponding to $\nu(\omega) = \infty, \forall\omega \in \Omega$ (i.e. a geometric Brownian motion with coefficients αx and σx) and $\underline{x}$ be the process corresponding to $\nu(\omega) = 0, \forall\omega \in \Omega$ (i.e. a geometric Brownian motion with coefficients $(\alpha - \Delta\alpha)x$ and σx). We can use dynamic programming to show that the optimal stopping times for these two processes depend on $\overline{z} =$

[9] Mamer (1987) obtains the existence of equilibria in a similar (but more restrictive) game in discrete time through similar techniques.

$z(\alpha, \sigma, C, \gamma)$ and $\underline{z} = z(\alpha - \Delta\alpha, \sigma, C, \gamma)$, where

$$z(\alpha, \sigma, C, \gamma) = \frac{\beta(\alpha, \sigma, \gamma)}{\beta(\alpha, \sigma, \gamma) - 1} C \tag{5}$$

and

$$\beta(\alpha, \sigma, \gamma) = 1/2 - \alpha/\sigma^2 + \sqrt{\left[\alpha/\sigma^2 - 1/2\right]^2 + 2\gamma/\sigma^2} > 1$$

(see Dixit and Pindyck (1994), pp. 140-144). The agent will stop the process as soon as it hits z. For convenience, we omit the parameter dependence of z in the remainder of the section.

Given a random time ν, we propose the stopping rule characterized by the following continuation region:

$$\begin{aligned} &\{x \leq z \equiv z(\alpha, \sigma, C, \gamma; \Delta\alpha, t, \text{other parameters})\} \text{ if } t \leq \nu \\ &\{x \leq \underline{z}\} \text{ if } t > \nu \end{aligned} \tag{6}$$

where the threshold levels z are determined from *value matching* and *smooth pasting* considerations in the optimal stopping problem (see the proof for Proposition 1). The "other parameters" refer to parameters related to the hazard rate associated with ν (as perceived by the agent). If ν arrives at a constant hazard rate λ, for instance, the threshold is constant in time and depends on the arrival rate λ and the decay in the drift $\Delta\alpha$. Once ν arrives, the process starts afresh and one is better by adopting the lower threshold rule. This rule is easily extended to processes with multiple breaks at increasing stopping times. It delivers a stopping strategy by which the agent switches progressively to lower threshold levels as the drift breaks take place. We thus state the result for the more general case:

Proposition 1 *Assume that* $\gamma > \alpha$ *and let* $\log x_t = \alpha t - \Delta\alpha \sum_{k=1}^{n} (t - \nu_k)\mathbb{I}_{t \geq \nu_k} - \frac{\sigma^2}{2} t + \sigma W_t$ *where* $\Delta\alpha \geq 0, \alpha, \sigma > 0, t \in \mathbb{R}_+, n \in \mathbb{N}, W$ *is a standard Brownian motion and* $\{\nu_k\}_{k=1,\ldots,n}$ *is an increasing sequence of stopping times. The optimal continuation region for the stopping problem is given by*

$$\begin{aligned} &\{x \leq z_{k-1}\} \textit{ if } t \leq \nu_k, k = 1, \ldots, n \\ &\{x \leq z_n\} \textit{ if } t > \nu_n \end{aligned}$$

for some threshold levels $z_k \equiv z_k(t)$ *with* $z_k(t) > z_{k+1}(t), \quad \forall t \quad k \in \{1, \ldots, n-1\}$ *and* $z_n \equiv \underline{z}$.

Consider now a game with two agents indexed by $i = 1, 2$. They contemplate an exit decision that will cost them $C_i, i = 1, 2$. In return, they collect a value $x_i, i = 1, 2$ but now the latent utility process for one agent is negatively affected once the other agent decides to leave the activity. In analogy to the previous individual setup, the information structure for the game assumes that each player observes his or her own state variable process and whether or not the other agents stopped or not up to but *excluding* the current instant (but not the value of other agents' individual state variables). In particular,

$$dx_t^i = \begin{cases} \alpha^i x_t^i dt + \sigma^i x_t^i dW_t^i & \text{if } t \leq \tau_j \\ (\alpha^i - \Delta\alpha^i) x_t^i dt + \sigma^i x_t^i dW_t^i & \text{if } t > \tau_j \end{cases}$$

where $i, j = 1, 2,\ i \neq j$ and τ^j is the stopping time adopted by the other agent in the game and, as above, $\gamma > \alpha^i$. The contemporaneous correlation between the Brownian motions is left unconstrained and $\Delta\alpha^i$ measures the external effect of the other agent's decision on i. As pointed out previously, this reveals the two major aspects of group behavior under consideration in this study: correlated and endogenous social effects. Individuals might behave similarly in response to associated (unobservable) shocks, which are reflected in the possibility that the increments of W_t^i and W_t^j are correlated. This represents the correlated effects. On the other hand, agents may be directly affected by other agents' actions as well. This would appear as a decrease in the profitability prospects an agent derives by remaining in the game. This is the endogenous effect.

Our parameterization assumes that individual switching costs are constant. This is in line with standard models in the investment literature. Another plausible configuration would allow this cost to depend on the number of individuals in the group I. If exit costs C_i are modeled as continuously evolving state variables (see, for example, Dixit and Pindyck (1994), p. 207) whose drifts are affected by the number of agents who choose to exit, the model is basically akin to the one presented here and x could actually be interpreted as the evolution of a utility process net of exit costs.[10] Our results would be essentially unchanged but with x reinterpreted in this way. Insofar as agents care about net utility for decision making, not the utility and exit costs separately, the detection of endogenous effects would not be affected. The implication of multiplier effects for the take-up of social programs such as those discussed in the Introduction would still occur.

The previous analysis establishes that each agent will use the "high drift" optimal stopping rule characterized by the (moving) threshold $z_i \equiv z_i(t)$ while $\tau^j \geq t$. As soon as $\tau_j < t$, she switches to the "low drift" stopping rule characterized by the threshold $\underline{z}_i$. In this case though, we need to handle the fact that τ^j is not exogenously given but determined within the game. It is illustrative to portray this interaction graphically.

Figure 1 displays the $X_1 \times X_2$ space where the evolution of the vector-valued process (x_1, x_2)

[10] Allowing costs to jump with the number of individuals who exit would introduce discontinuities in the state process and raise concerns previously alluded to.

is represented. Since $\Delta\alpha^i > 0, i = 1, 2$, we should have $z_i(t) > \underline{z}_i, i = 1, 2$. Agents start out under threshold $z_i(t)$. If the other agent stops, the threshold level drops to $\underline{z}_i$. In Figure 1, for instance, the process fluctuates in the rectangle $(0, z_1) \times (0, \underline{z}_2)$ and reaches the barrier z_1 causing agent 1 to stop. Once this happens, agent 2's threshold drops to $\underline{z}_2$, which, once reached, provokes agent 2 to stop.

FIGURE 1 HERE

A more interesting situation is depicted in Figure 2. Here, the vector process sample path attains the upper threshold for agent 1 at $x^2 \geq \underline{z}^2$. The second agent's threshold moves down immediately and both stop simultaneously. So, if an agent's latent utility process is above the subsequently lower threshold when the other one drops out, there will be clustering and they move out concomitantly. This is an interesting feature of the game that is not present in standard statistical models that would handle timing situations such as this: the positive probability for simultaneous events even when time is observed continuously. If not properly accounted for, this can bias results toward erroneous conclusions. [11]

FIGURE 2 HERE

One concern in the analysis of this interaction is how beliefs about the state of one's opponent should affect his or her actions. If an individual knows only whether or not the opponent has quit, how should he or she take into consideration the risk of being preempted? Should the player take the presence of the opponent for granted and delay the decision to quit or must he believe that the other agent is about to quit the game and hence leave the activity immediately? Such considerations point to the importance of beliefs in these environments and are a relevant consideration in Hopenhayn and Squintani (2004), for instance. In the present case, such calculations are of lesser importance since the state variables evolve continuously. This implies that the likelihood that an agent reaches a stopping region between now (t) and an ϵ unit of time in the future ($t + \epsilon$) vanishes as $\epsilon \to 0$. Consequently, the beliefs about the location of the opponents' state variables are of second order to the decisions taken within the next infinitesimal period by a given agent.

The intuition above carries over with more than two agents. Assume as before that an exit costs an agent $C_i, i \in I$ in return for $x^i, i \in I$. The latent utility process is given by:

$$\log x_t^i = \alpha^i t - \Delta\alpha^i \sum_{j:j\neq i} (t - \tau^j)\mathbb{I}_{t>\tau^j}/(I-1) - \frac{\sigma^{i2}}{2}t + \sigma^i W_t^i, \quad i \in I$$

[11] Van den Berg, Lindeboom and Ridder (1994), for instance, point to a negative duration dependence bias in estimates if simultaneity is left unaccounted for.

where τ^j is the stopping time adopted by the agent j. The external effect of other agents on i is $\Delta\alpha^i > 0$ and is considered to be homogeneous across agents, i.e. the amount by which the drift α^i decreases with each stopping decision is the same regardless of who deserts.

A few other definitions are convenient:

z^i_m : $z(\alpha^i, \sigma^i, C^i, \gamma^i, m, \Delta\alpha^i, t)$ where $i, m \in I$

$\mathcal{S}_m$: $\{(x^1, x^2, \ldots, x^I) \in \mathbb{R}^I_+ : \exists i$ such that $x^i \geq z^i_m\}$ where $m \in I$

τ_0 : 0 (meaning $\tau_0(\omega) = 0, \forall\omega$)

A_0 : I_I (identity matrix of order I)

τ_m : $\inf\{t > \tau_{m-1} : A_{m-1}x_t \in A_{m-1}\mathcal{S}_{I+1-\mathbf{1}'A_{m-1}\mathbf{1}}\}$ where $A_{m-1}\mathcal{S}_{I+1-\mathbf{1}'A_{m-1}\mathbf{1}}$ denotes the set formed by operating the matrix A_{m-1} on each element of $\mathcal{S}_{I+1-\mathbf{1}'A_{m-1}\mathbf{1}}$, $\mathbf{1}$ is an $I \times 1$ vector of ones and $m \in I$

A_m : $[a^m_{kl}]_{I\times I}$ where $a^m_{kl} = \mathbb{I}_{x^i_{\tau_m} < z^i_m}$ if $k = l = i$ and $a^m_{kl} = 0$ otherwise and $m \in I$

The stopping times defined above are essentially hitting times. The thresholds z^i_m are defined by the value matching and smooth fit conditions (see the proof of Proposition 2 for details).

The game starts out with no defection and agents hold the highest barrier $z^i_1(t)$ as the initial exit rule. The first exit occurs at τ_1, the hitting time for the stopping region $\mathcal{S}_1$. As the process reaches this set, one or more agents quit. This also shifts the thresholds down as the stopping region moves to $\mathcal{S}_2$. To track the players who drop out at τ_1 we use the matrix A_1, a diagonal matrix with ones for those agents who did not drop out at τ_1 and zeros, otherwise. Analogously for further stopping rounds, defections occur at the stopping times $\tau_.$ and $\mathbf{1}'A_.\mathbf{1}$ records the number of agents that have not stopped after that stage. This goes on until all agents have stopped. The following proposition summarizes our result:

Proposition 2 *The profile $(\tau^*_i)_{i\in I}$ represents a vector of equilibrium strategies for the exit game with I agents:*

$$\tau^*_i = \sum_{k=1}^{I} (\Pi_{j=1}^{k-1}\mathbb{I}_{x^i_{\tau_j} < z^i_j})\mathbb{I}_{x^i_{\tau_k} \geq z^i_k}\tau_k$$

This proposition states that the hitting times constructed previously may be used to represent an equilibrium for this game. For the reasons discussed in the next subsection, this is the equilibrium we focus on in this paper.

3.1. On Multiple Equilibria and Equilibrium Selection

The above equilibrium is the only equilibrium that is robust to positive delays in information about the exit of others. As we drive this delay to zero, this is as if each agent observed his or her own state variable process and whether other agents stopped up until but *excluding* the current instant t: stopping decisions are observed with an "infinitesimal" delay.

Most of the equilibria considered in the previous discussion rely on a strong degree of synchronization among agents. If nevertheless one agent's exit is perceived with a delay, dropping out may not elicit other players' exit. The profitability of remaining in the game, represented by the drift coefficient, would be unchanged. Then, exit would only be optimal on the equilibrium portrayed in the previous subsection. This "synchronization risk" is inherent in many similar situations (see Abreu and Brunnermeier (1997), Brunnermeier and Morgan (2004) and Morris (1995)) and equilibria that survive such synchronization issues are naturally more compelling. To formalize this intuition, consider the case of two players ($I = 2$) and a vector-valued random variable $(\epsilon_i)_{i \in I}$ representing the agents' perception delay. Let an individual's perceived utility be given by:

$$dx_t^i = \begin{cases} \alpha^i x_t^i dt + \sigma^i x_t^i dW_t^i & \text{if } t \leq \nu \\ (\alpha^i - \Delta\alpha^i) x_t^i dt + \sigma^i x_t^i dW_t^i & \text{if } t > \nu \end{cases}$$

where $i, j = 1, 2,\ i \neq j$ and $\nu = \tau^j + \epsilon_i$ is the stopping time adopted by the other agent in the game with an ϵ_i delay and, as before, $\gamma > \alpha^i$. The following statement then holds:

Proposition 3 *Assume that* $\mathbb{P}(\{\epsilon_i > 0\}) = 1, i = 1, 2$ *and* $\mathbb{P}(\{\epsilon_1 = \epsilon_2\}) = 0$. *Also, let*

$$\mathcal{S}(t) = \{(x^1, x^2) : \exists i \text{ such that } x^i \geq z^i(t)\}$$

and $\tau_{\mathcal{S}} = \inf\{t > 0 : (x_t^1, x_t^2) \in \mathcal{S}(t)\}$ *denote the hitting time for this set. The stopping strategies below represent the unique equilibrium profile for this game:*

$$\tau_i^* = \tau_{\mathcal{S}} \mathbb{I}_{x^i_{\tau_{\mathcal{S}}} = z^i} + \inf\{t > \tau_{\mathcal{S}} + \epsilon_i : x_t^i > \underline{z}^i\} \mathbb{I}_{x^i_{\tau_{\mathcal{S}}} \neq z^i}, \qquad i = 1, 2.$$

As $\epsilon_i \xrightarrow{P} 0$, the above strategies converge to the strategy depicted in Proposition 2. For this reason, we restrict our attention to the unique equilibrium that is robust to such perturbations and corresponds to the one displayed in the last subsection.

Other information structures would nonetheless be susceptible to multiple equilibria,[12] though the occurrence of joint exit with positive probability is robust to the existence of multiple equilibria and the issue of equilibrium selection. Given this, some of our results in this and the subsequent section are robust even to the existence of multiple equilibria. Next we discuss the empirical implications of the model.

[12]A discussion of this can be found in a longer version of the paper available on the author's website.

4. Empirical Implications

In this section we investigate the empirical implications of the model. The unit of observation is a game[13] and N such units are recorded in a sample. We restrict attention to the unique equilibrium depicted in the previous section, which is robust to perturbations in the timing at which agents become aware of the actions of other players.

Our analysis of the model has so far allowed for asymmetry in individual parameters so that for instance $\Delta\alpha$ and σ may be indexed by individual, i. To discuss its empirical relevance, we impose additional assumptions, and we consider a symmetric version of the game where $\Delta\alpha, \sigma, C$ and ρ are homogeneous across players and α may differ insofar as it depends on individual specific covariates, but is otherwise identical across individuals.[14] Each agent's latent utility hence follows:

$$\log x_t^i = \alpha^i t - \Delta\alpha \sum_{j:j\neq i} (t-\tau^j)\frac{\mathbb{I}_{t>\tau^j}}{I-1} - \frac{\sigma^2}{2}t + \sigma W_t^i + \log x_0^i, \quad i \in I$$

where τ^j is the stopping time adopted by player j. The cross-variation process for the Brownian motions is given by $\langle W^i, W^j \rangle_t = \rho t, i \neq j$ and the initial condition $\mathbf{x}_0 = (x_0^i)_{i\in I}$ follows a probability law F_0^i. It is assumed throughout that $|\rho| < 1$.

The individual initial drift coefficient is potentially a function of an l-dimensional vector of individual covariates $w_{i(1\times l)}$, which is independent of the Brownian motion. More specifically, $\alpha^i = \alpha(w_i)$. To benefit readability, we suppress the argument and denote the drift by α^i. Let $F_{\mathbf{w}}$ denote the distribution of $\mathbf{w} = (w_i)_{i\in I}$. In what follows all of the statements are conditional on $\mathbf{w} = (w_i)_{i\in I}$. The parameter $\Delta\alpha$ measures the external effect of the other agents decisions on i and introduces endogenous social effects. The coefficient ρ represents correlated social effects. In addition to the above parameters, each agent pays a cost C to leave and discounts the future at the exponential rate γ. Finally, $z^i, i \in I$ denotes the threshold presented in the previous section.

4.1. Characterization

The next proposition states that (under continuous time observability) simultaneous departures occur only in the presence of endogenous effects.[15]

Proposition 4 $\mathbb{P}[\tau^i = \tau^j, i \neq j, i, j \in I] > 0$ *if and only if there are endogenous effects* ($\Delta\alpha > 0$).

[13] In our empirical application a game is a military company. In other applications, it would be a household or a geographic market or some other arena of interaction for the agents under analysis.

[14] This does not necessarily mean that the asymmetric game is not identified. If the parameters depend on i, we conjecture that the model is not identified unless one can for example sample the same players repeatedly.

[15] This proposition does not depend on the fact that $\Delta\alpha_i = \Delta\alpha, \forall i$. It nonetheless relies on the assumption that $\Delta\alpha_i > 0, \forall i$.

Intuitively, because of the diffusive nature of the utility process correlated disturbances occur in a much smoother manner if compared to the impact of an agent's exit, which is immediately assimilated via a decrease in the growth rate of x and a revision of the stopping rule. This is a useful feature of this model and holds in many empirical situations in which the model applies. Moreover, this empirical implication does not rely on the uniqueness of the equilibrium. Notice that in traditional econometric models for duration analysis the probability of simultaneous exit is zero and such incompatibility may provoke biased estimates and contaminate conclusions.

This result relies on the continuity of the sample paths for the stipulated process. If discontinuities are allowed, this may not hold any longer.[16] The problem would nonetheless be lessened if one knew the timing of such shocks since clustering in other moments is then seen as evidence in favor of endogenous effects.

Another implication is that the number of players should affect equilibrium stopping outcomes only in the presence of endogenous effects. This is stated in the next proposition.

Proposition 5 *If the number of players I affects the marginal distribution of equilibrium stopping times in the game, then there are endogenous effects ($\Delta\alpha > 0$).*

Notice that the direction in which the equilibrium stopping times are affected is not clear. On the one hand, more players will cause each one's exit to have a smaller impact on an agent's latent utility process; on the other hand, exits will tend to occur earlier. Also, as mentioned earlier, if costs are also allowed to depend on group size, the proposition above should be framed in terms of utility processes net of costs.

4.2. Nonparametric Test for Endogenous Interactions

If time were recorded continuously, Proposition 4 would suggest that observing simultaneous exits would be enough to detect endogenous effects. When time is marked at discrete intervals, though, exit times would be lumped together regardless of the existence of endogenous influences. In this subsection we explore the possibility of testing for the existence of social interactions, taking into consideration that time is not sampled continuously.

Let $n = 1, \ldots, N$ index independent realizations of the game and denote by I_n the number of players in realization n. Time is observed at discrete intervals of stepsize Δ_N. Given a discretization $\{t_0, t_1, \ldots\}$ such that $t_{i+1} - t_i = \Delta_N, \forall i$, we denote the probability of a simultaneous exit by any pair of players

$$\mathbb{P}_{\Delta_N}(\{\text{ simultaneous exit }\}) = p(\Delta_N)$$

[16] One way to introduce such discontinuities is to insert an exogenous jump component dQ^i in equation (2). In this case, beliefs would play a more significant role.

and allow the discretization to depend on the sample size.

Imagine that there are no endogenous interactions. In this case, for a small enough discretization, doubling the observation interval would roughly double the probability of recording exits as simultaneous. If these endogenous effects are present, since even at continuous time sampling there would still be clustering, doubling the discretization does not increase the probability of joint exit by as much. In the limit, if all exits are indeed simultaneous in continuous time, varying the grid of observation would have no effect on the probability of observing simultaneous dropouts. We use this intuition to develop a test for the null hypothesis of no endogenous effect through variation in the interval of observation.

In our model, when there are no endogenous effects, the function $p(\cdot)$ is differentiable with $p'(\cdot) > 0$. Also $p(0) = 0$ when there are no endogenous effects; whereas $p(0) > 0$, otherwise. Denote by

$$\mathbf{y}_{n,\Delta_N} = \binom{I_n}{2}^{-1} \sum_{\{i,j\}\in\pi_n} \mathbb{I}_{\{\tau^i_{\Delta_N} = \tau^j_{\Delta_N}\}},$$

where π_n is the set of all player pairs in game n and $\tau^i_{\Delta_N}$ is the exit time observed when the discretization grid size is Δ_N. If the game has only two players, $\mathbf{y}_{n,\Delta_N}$ records whether there was simultaneous exit under a discretization of size Δ_N. It can be established that $\mathbb{E}(\mathbf{y}_{n,\Delta_N}) = p(\Delta_N)$. For $I_n = 2$, $\text{var}(\mathbf{y}_{n,\Delta_N}) = p(\Delta_N)(1-p(\Delta_N))$. For the general case, we denote $\text{var}(\mathbf{y}_{n,\Delta_N}) = v(\Delta_N)$. It is easily seen that $p(0) > 0 \Rightarrow v(0) > 0$ and $p(0) = 0 \Rightarrow v(0) = 0$. Given the observation of N i.i.d. copies of such games, consider $\overline{\mathbf{y}}_{N,\Delta_N} = N^{-1}\sum_{n=1}^{N} \mathbf{y}_{n,\Delta_N}$. Then:

Theorem 2 *Assume*

1. *$p(\cdot)$ is differentiable and $p'_+(0) > 0$ if $p(0) = 0$;*
2. *$\Delta_{N,i} = a_i N^{-\epsilon}, i = 1, 2, 3$ with $a_1 < a_2, a_3 (a_2 \neq a_3)$, and $1/3 < \epsilon < 1$;*
3. *The games observed are i.i.d..*

Then, under the (null) hypothesis that there are no endogenous effects ($H_0 : p(0) = 0$),

$$\sqrt{N}\sigma_N^{-\frac{1}{2}} \left[\frac{\overline{\mathbf{y}}_{\Delta_{N,2}}}{\overline{\mathbf{y}}_{\Delta_{N,1}}} - \frac{a_2}{a_1} - \xi \left(\frac{\overline{\mathbf{y}}_{\Delta_{N,3}}}{\overline{\mathbf{y}}_{\Delta_{N,1}}} - \frac{a_3}{a_1} \right) \right] \xrightarrow{\mathcal{L}} \mathcal{N}(0,1) \tag{7}$$

where

$$\sigma_N = \left[\frac{\xi p(\Delta_{N,3}) - p(\Delta_{N,2})}{p(\Delta_{N,1})^2} \quad \frac{1}{p(\Delta_{N,1})} \quad -\frac{\xi}{p(\Delta_{N,1})} \right] var \begin{bmatrix} \mathbf{y}_{n,\Delta_{N,1}} \\ \mathbf{y}_{n,\Delta_{N,2}} \\ \mathbf{y}_{n,\Delta_{N,3}} \end{bmatrix} \begin{bmatrix} \frac{\xi p(\Delta_{N,3}) - p(\Delta_{N,2})}{p(\Delta_{N,1})^2} \\ \frac{1}{p(\Delta_{N,1})} \\ -\frac{\xi}{p(\Delta_{N,1})} \end{bmatrix}$$

and

$$\xi = \frac{a_2(a_2 - a_1)}{a_3(a_3 - a_1)}$$

whereas if there are endogenous effects ($p(0) > 0$), the statistic in (7) is not bounded in probability.

Above, $p'_+(\cdot)$ denotes the right-derivative of function $p(\cdot)$. The smaller and the closer $\Delta_{N,1}$ and $\Delta_{N,2}$ are, the higher the precision for the ratio is.[17] Also, in estimating the asymptotic variance, one could use as consistent estimators the sample counterparts:

$$\widehat{p(\Delta_{N,j})} = \overline{\mathbf{y}}_{\Delta_{N,j}} \quad j = 1, 2$$

and the sample variance covariance matrix across the games.

The above result allows us to test the null hypothesis that there are no endogenous effects ($H_0 : p(0) = 0$) and no clustering results in continuous time when observations are recorded discretely with the suggested statistic. As a corollary for the theorem, the test will be consistent. Its performance may nevertheless be affected (especially in small samples) by other parameters such as the correlation coefficient ρ.

Since it relies on Proposition 4, the above result is also robust to the existence of multiple equilibria (as long as the equilibrium played is the same across the games sampled). In the next subsections we explore some representation and identifiability properties under the assumption that the equilibrium played is the one characterized in the previous section.

4.3. Identification

One question that arises naturally is the possibility of disentangling correlated and endogenous effects in the data. The econometrician observes the equilibrium exit strategies $(\tau_1, \ldots, \tau_I)$ for a certain number of realizations of the game. Let τ denote some outcome variables observed by the researcher and $\mathbf{w}$, some observable covariates. A parameter ψ (of arbitrary finite dimension) lies in a certain set Ψ and governs the probability distribution $P(\cdot|\mathbf{w};\psi)$ of the outcome variables. The following defines identification.

Definition 2 (Identification) *The parameter $\psi \in \Psi$ is identified relative to $\hat{\psi}$ if ($\hat{\psi} \notin \Psi$) or ($P(\cdot|\mathbf{w};\psi) = P(\cdot|\mathbf{w};\hat{\psi}), F_{\mathbf{w}}$-a.e. $\Rightarrow \psi = \hat{\psi}$).*

[17] As it relies on Assumption 2, the power of the test may be affected by the coarseness of the data in a non-negligible manner. This is an issue as well for related techniques in continuous time finance and in the empirical game estimation literature. Our empirical application employs data at a daily frequency, which is appropriate for the phenomenon investigated.

The first stand on identifiability for the model above is a negative one: the full parameter vector is not identified. To see this, notice that with no social interactions or correlated effects ($\Delta\alpha = 0$ and $\rho = 0$), the individual Brownian motions are independent and each agent's latent utility process evolves as a geometric Brownian motion with drift α^i, diffusion coefficient σ and initial position x_i. As a consequence, the exit times τ_i^* are independent (possibly defective) inverse Gaussian random variables (see Chhikara and Folks (1989)). This distribution is characterized by two parameters for which the mean and harmonic mean are maximum likelihood estimators and minimal sufficient statistics. Since we would still have more than two parameters $(\alpha, \sigma, C, \gamma)$, the model remains unidentified.

Under certain circumstances, though, some positive assertions about the parametric identification for this model can be made even when $\rho, \Delta\alpha \neq 0$. Identifiability may be achieved if one is able to introduce "enough variability" through the use of covariates. Recall that we assumed $\alpha^i = \alpha(w_i)$, where w_i is a set of covariates. These covariates may encompass individual-specific characteristics, such as income or assets, as well as publicly known information, which would be a common covariate to all agents, or observable group characteristics (which would account for contextual effects in Manski's categorization). Let $g(t; \psi, \mathbf{w})$ denote the probability density function for the first desertion time under the parameters $\psi = (\alpha, \Delta\alpha, \sigma, \rho, \gamma, C)$ and conditioned on the observable covariates $\mathbf{w}$. The following statement establishes sufficient conditions for the identification of ψ. It basically states that relative identification is achieved if, by perturbing the covariates, one perturbs the Kullback-Leibler information criterion.[18]

Theorem 3 *Let w_i contain at least one continuous random covariate, $\alpha(\cdot)$ be C^1 with respect to such variable and, for some i and some continuous covariate l,*

$$\partial_{w_{il}} \int \log \left[\frac{g(t; \psi, \mathbf{w})}{g(t; \hat{\psi}, \mathbf{w})} \right] g(t; \psi, \mathbf{w}) dt \neq 0 \tag{8}$$

then ψ is identified relative to $\hat{\psi}$.

To check condition (8) one should obtain the density $g(\cdot)$. To illustrate the above theorem, consider a very simple example in which $g(\cdot)$ is an inverse Gaussian probability density function. In particular,

$$g(t; a, v) = \frac{a}{\sqrt{2\pi t^3}} \exp\left(-\frac{(a - vt)^2}{2t} \right).$$

[18] Another potential avenue for identification would be through the results presented in McManus (1992). With a sufficiently high number of players (corresponding to endogenous variables) relative to the parameters of the model, the structure can be seen to be (generically) identified.

As pointed out in the preceding discussion, this will be the probability density function for stopping times in our model when there are no social effects and hence $\Delta\alpha = \rho = 0$. In this scenario, $v = \alpha(w_i)/\sigma - \sigma/2$ and $a = \log(z/x_0)/\sigma$, with z as in equation (5). For two sets of parameters (a, v) and $(\hat{a}, \hat{v})$. After some algebra, the Kullback-Leibler criterion for this $g(\cdot)$ is seen to be:

$$\int \log\left[\frac{g(t; a, v)}{g(t; \hat{a}, \hat{v})}\right] g(t; a, v)dt = \log\frac{a}{\hat{a}} - \frac{1}{2} + \frac{\hat{a}^2}{2}\left(\frac{v}{a} + \frac{1}{a^2}\right) - \hat{v}\left(\hat{a} - \frac{\hat{v}a}{2v}\right)$$

The derivative of this expression with respect to w_i can be obtained and evaluated for different pairs of parameters in order to verify the sufficient condition in the theorem. For example, if $\alpha(w_i) = 0$ and $\hat{\alpha}(w_i) = \hat{\alpha}w_i(\hat{\alpha} \neq 0)$, after some algebra we obtain that this derivative is:

$$\frac{\hat{\alpha}w_i}{\hat{\sigma}^2(\hat{\beta}(w_i) - 0.5 + \hat{\alpha}w_i/\hat{\sigma}^2)(\hat{\beta}(w_i) - 1)}\left(\frac{v\hat{a}}{a} + \frac{\hat{a}}{a^2} - \frac{1}{\hat{a}} - \hat{v}(w_i)\right) - \frac{\hat{\alpha}w_i}{\hat{\sigma}}\left(\hat{a}(w_i) - \frac{a\hat{v}(w_i)}{v}\right)$$

where $\hat{\beta}$ is the expression in (3) (using as parameters $\hat{\alpha}w_i, \hat{\sigma}, \hat{\gamma}$) and we highlight the fact that $\hat{\beta}, \hat{v}$ and $\hat{a}$ all depend on w_i. For a given set of parameters the above can be evaluated. For instance, if $C = \hat{C} = 0.1$, $x_0 = \hat{x}_0 = 1$, $\gamma = \hat{\gamma} = 0.05$, $\sigma = \hat{\sigma} = 0.1$ and $w_i = 1$, the expression is greater than 25. As previously noted, $\hat{\alpha} = 0.04$ and $\alpha = 0$. These two parameter vectors are relatively identified provided the support of w_i includes 1. It should then be noted that given two parameter vectors the above expression (or a similar one for $\alpha \neq 0$) will be solved only for certain values of w_i and provided its support is rich enough, a nonzero value for this derivative can easily be obtained for a large set of w_i values and one can use this to ascertain more encompassing global identification results.

In general, one possible route to obtain $g(\cdot)$ is to use the close association between the theory of stochastic processes and the study of differential equations. Assuming that the equilibrium played is the one selected in the previous section, the equilibrium strategies can then be expressed as hitting times to certain sets and it is possible to characterize the survivor function for the first hitting time $G(\cdot)$ through the associated partial differential equations (and $g(\cdot) = -dG(\cdot)/dt$). When $\Delta\alpha = 0$ (but ρ is left unrestricted) and there are only two players, $G(\cdot)$ is (see Theorem 3.5.2 in Rebholz (1994) or He, Keierstead, and Rebholz (1998), Remark 2.2(ii)):

$$G(t) = \mathbb{P}(\tau_1^* \wedge \tau_2^* \geq t) = e^{a_1 \log(z^1/x_1) + a_2 \log(z^2/x_2) + bt} f(r', \theta', t) \tag{9}$$

where

$$f(r', \theta', t) = \frac{2}{\lambda' t} \sum_{n=1}^{\infty} \sin(\frac{n\pi\theta'}{\lambda'}) e^{-\frac{r'^2}{2t}} \int_0^{\lambda'} \sin(\frac{n\pi\theta}{\lambda'}) g_n(\theta) d\theta$$

with

$$g_n(\theta) = \int_0^\infty re^{-\frac{r^2}{2t}} e^{-b_1 r\cos(\theta-\lambda)-b_2 r\sin(\theta-\lambda)} I_{\frac{n\pi}{\lambda}}(\frac{rr'}{t})dr$$

and

$$\begin{aligned}
\tan\lambda' &= -\tfrac{\sqrt{1-\rho^2}}{\rho} \\
\lambda &= \lambda' - \tfrac{\pi}{2} \\
r' &= \tfrac{1}{\sqrt{1-\rho^2}}\left(\tfrac{\log(z^1/x_1)^2-2\rho\log(z^1/x_1)\log(z^2/x_2)+\log(z^2/x_2)^2}{\sigma^2}\right)^{\frac{1}{2}} \\
\theta' &= \tfrac{log(z^1/x_1)}{\sigma r'} \\
a_1 &= \tfrac{(\alpha^1-\sigma^2/2)-\rho(\alpha^2-\sigma^2/2)}{(1-\rho^2)\sigma^2} \\
a_2 &= \tfrac{(\alpha^2-\sigma^2/2)-\rho(\alpha^1-\sigma^2/2)}{(1-\rho^2)\sigma^2} \\
b &= \tfrac{\sigma^2}{2}(a_1^2+2\rho a_1 a_2+a_2^2)-(\alpha^1-\sigma^2/2)a_1-(\alpha^2-\sigma^2/2)a_2 \\
b_1 &= (a_1+a_2\rho)\sigma \\
b_2 &= a_2\sigma\sqrt{1-\rho^2}
\end{aligned}$$

with $z^i \equiv z(\alpha^i, \sigma, C, \gamma)$ and where I_v is the modified Bessel function. Iyengar (1985), which also derives an expression for the above function, hints that the above is generalizable for higher dimensions in our specific situation. The density and the Kullback-Leibler criterion can be obtained from the above expression for the evaluation of the condition in the theorem for different parameter values.[19]

5. Empirical Illustration: Desertion in the Union Army

Using a dataset comprising detailed individual records for soldiers of the Union Army in the American Civil War, we now intend to illustrate the previous discussion on stopping decisions and timing coordination. Desertion is the event we are interested in. Historians estimate that desertion afflicted a bit less than 10% of the Union troops (circa 200,000 soldiers).

Whereas one could think of the decision to desert as an isolated one, historical studies and anecdotal evidence support the existence of endogenous effects. Evidence of simultaneous desertion (on both sides) is pervasive in Lonn (1928): "*Usually the recorded statements of specific instances of desertion whether from Union or Confederate reports, show the slipping-away of individuals or of small groups, varying from five to sixteen or twenty.*" (pp. 152-3) The author goes on to point out instances where Union soldiers would desert by the hundreds at the same time.

From these facts, it is valid to infer that a soldier's decision to desert probably had a

[19] In a proposition available upon request we present the partial differential equation associated with the survivor function for more general cases. We conjecture that, even in the presence of discontinuities in the state processes, a similar result may be attained relying on partial differential equations for the characterization of equilibrium exit distributions.

direct impact on the behavior of others in his company. If no one deserts, the social sanctions attached to exit tend to be high; if there is mass exit, such sanctions tend to be minimized and the effectiveness of the military company, tends to decrease. Furthermore, such decisions entailed costs — the probability of being caught and facing a military court.[20] These two aspects are in accordance with the model we investigated previously. Another feature of these data that is particularly helpful is the fact that recruits tended to be with a company from its inception and hence there was very little flow of soldiers into or out of the unit.

5.1. Data and Preliminary Analysis

The data used consist of 35,567 recruits in the Union Army during the American Civil War. This dataset was collected by the Center for Population Economics at the University of Chicago under the auspices of the National Institute of Health (P01 AG10120). It is publicly available at `http://www.cpe.uchicago.edu`. The men are distributed across 303 military companies from all states in the Union with the exception of Rhode Island. These companies were randomly drawn using a one-stage cluster sampling procedure, and all recruits for each selected company, except for commissioned officers, black recruits, and some other branches of military service, were entered into the sample. These soldiers represent 1.27% of the total military contingent in the Union and a significant portion of the 1,696 infantry regiments in that army. According to the Center for Population Economics they seem to be representative of the contemporary white male population who served in the Union Army.

A number of variables is available for each recruit. These include dates of enlistment, muster-in, and discharge as well as information on promotion, AWOL (absent without leave), desertion, and furlough.[21] More detailed military information from the recruit records is available and is complemented by background information and post-war history originally from the census. We focus on the main military variables.

According to Lonn (1928), desertion was markedly higher among foreigners, substitutes and "bounty-jumpers." Substitutes and "bounty-jumpers" appeared as the government started induc-

[20]Even though the Military Code in effect at the beginning of the war mandated sanctions as harsh as the death penalty, such punishments required the approval of the President or (later during the war) the commanding general. According to Costa and Kahn (2003), out of an estimated 200,000 deserters, 80,000 were caught, of which only 147 were executed. Especially in the early years, punishments were notoriously mild, consisting of dismissal with loss of pay and, toward the end, imprisonment for the duration of the war.

[21]Desertion and other military events were recorded by the company officers. Some mis-measurement of desertion is to be expected, and we ignore this possibility. Records are nonetheless reported to have become more accurate toward the end of the war, especially after the institution of the office of provost marshall general in September 1862 (see Lonn (1928)).

ing enlistment through enrollment bounties — which created the figure of the "bounty-jumper", who would enlist, collect the reward and desert just to repeat the scheme in another state or county — and the possibility for draftees to hire substitutes. In the data it is possible to identify foreigners and substitutes. To assess the effect of "bounty-jumpers" on desertion, we try the bounty amount paid to each recruit as a proxy variable.[22] Other variables are also included, such as marital status, age and height as well as dummy variables for state and year of enlistment. The ideal dataset would contain continuous time records for desertion. Here, an event is marked with daily precision[23] and time to desertion is measured from the earliest muster-in date for recruits in a given company.[24]

One of the implications of our model is that company size will affect the equilibrium exit strategies only in the presence of endogenous effects (see the discussion of Proposition 5 nonetheless). Table 1 presents evidence for this. The regressions investigate the effect of certain variables on the mean (log of the) time to desertion at the individual level.[25] Company size is a significant and robust determinant for the timing of desertion. In addition, we tried other specifications with different combinations of independent variables. The company size variable remains significant in all of those. Anecdotal evidence and history texts point to a very unsystematic enlistment process, typically held at the local level by community leaders, which provides some justification for assuming that the effect of company size does not represent an omitted factor other than the numbers in the group.

TABLE 1 HERE

To further investigate the presence of endogenous effects in our data, we compute the statistics in Theorem 2 for various discretization levels. Even though the proposed test could suffer from the arbitrary choice of these parameters, all of them yield results that reject the null hypothesis of no endogenous effects, as displayed in Table 2 below. The results are for desertions that did not occur during battles, lest these represent common shocks that discontinuously affect the utility flow. The conclusions are unchanged if one includes desertions that occurred during battles.

[22]The bounty amount was not adjusted for inflation, but whenever it was used year dummies were also present which would capture nationwide inflation levels.

[23]Some deserters did not have precise dates and were thus discarded.

[24]Non-parametric estimation of the hazard rate suggests negative duration dependence at earlier dates and mildly positive to no duration dependence later in the soldier's army life.

[25]The regressions can be related to an accelerated failure time model for the time to desertion. Similar versions were also run at the company level with essentially the same conclusions.

TABLE 2 HERE

In the next subsection we proceed with the analysis by structurally estimating the model considered in the paper.

5.2. Estimation

In this subsection we use a simulated minimum distance estimator for the relevant parameters in the model proposed in Section 3. We normalize the discount rate ($\gamma = 5\%$ per year),[26] the exit cost ($C = 1$) and the initial condition ($x_0 = 0.1$). The normalizations are necessary as we do not use covariates in this exercise and identification would be jeopardized without any restriction. The four parameters estimated are α, ρ, σ and $\Delta\alpha$. Intuitively the first three are pinned down by the distribution of first exit, which is essentially the minimum of correlated Inverse Gaussian random variables. In the case of two players, this distribution is given by (**??**) which is a nonlinear function of these three parameters. The parameter $\Delta\alpha$ on the other hand controls the probability of simultaneous exits and can be solved for once the other parameters are obtained. Our estimator $\hat{\psi}$ then minimizes the following distance:

$$\|G_N(\psi)\| = \|N^{-1}\sum_{n=1}^{N} m(\tau_n) - (NR)^{-1}\sum_{r=1}^{NR} m(\tau_r(\psi))\|$$

where $m : \mathbb{R}_+ \longrightarrow \mathbb{R}^k$ and the second sum is taken over the simulated observations generated under parameter ψ. The stopping times recorded are only those prior to a certain horizon T, which in the context stands for the individual term of service in the army. We use $R = 1$. To simulate the phenomenon, we have to discretize the sample paths in the simulations. The discretization error in the simulations can be ignored in large samples for our estimation as long as the discretization grids are $o(\sqrt{N})$, which we guarantee via an Euler scheme (see Glasserman (2004)). Consistency and asymptotic normality are then a straightforward application of the results in Pakes and Pollard (1989). One important condition for this application is that $\Delta\alpha > 0$. We assume that to be reasonable, given the test statistics obtained in Table 2.

An important simplification is that we assume the thresholds to be constant: $z_k = z(\alpha - \Delta\alpha(k-1)/(I-1), \sigma, \gamma, C)$.[27] The moments matched were mean, harmonic mean, average number of desertions in each desertion episode and percentage of soldiers leaving before two years.[28] In the

[26]For comparison, commercial paper rates in United States during the war fluctuated between 4% and 8% (NBER Macrohistory Database).

[27]Similar approximations can be found on the treatment of finite horizon options and seem to work satisfactorily. Examples are Huang, Subrahmanyam, and Yu (1996) and Ju (1998).

[28]Very similar results were obtained if one of the percentages was substituted for the average number of deserters at each desertion episode.

absence of endogenous or correlated effects, the mean and harmonic mean are MLE and sufficient statistics for the inverse Gaussian distribution. The other two moments are used as $\Delta\alpha$ and ρ affect the clustering and speed at which individuals drop out in the group. The following table displays the results and normalizations used in the estimation:

TABLE 3 HERE

The parameters are precisely estimated. The results indicate a substantial endogenous effect. Using as parameters for this distribution the point estimates above, one obtains a probability of 13.90% for leaving the game before 150 days in the absence of endogenous effects ($\Delta\alpha = 0$). If on the other hand one-fourth of the company deserts immediately after the beginning of the war for instance, the endogenous effect coefficient estimate implies a probability of leaving the army before 150 days of 31.78%.[29]

6. Conclusion

The problem studied here is of great importance in many settings. Decisions about participation in a social welfare program, bank runs, migration, marriage and divorce are only a few of the possibilities. Disentangling endogenous and correlated effects is thus fundamental not only to illuminate economic research but also to enlighten policy. The setup delineated in this paper allows us to better understand the nature of endogenous and correlated effects. Whereas this problem is unfeasible in simpler settings (see Manski (1993)), the separation is not clear in other approaches that deal with similar situations (as in Brock and Durlauf (2001)).

For the environment considered in this paper, we have learned that endogenous interactions may be an important component in multi-person timing situations. They can generate simultaneous actions with positive probability and thus interfere with the usual statistical inference through standard duration models. A few characterizations were possible, and a test for the presence of endogenous influences was delivered. Finally, structural estimation points to a significant effect on the outcome of our particular example.

[29] We have also estimated the model imposing termination exogenously through death. The results are not much different.

Appendix: Proofs

Sketch of Proof for the Existence of a Strong Solution

The proof that there exists a strong solution for equation 2 follows from a slight modification of the proof provided in Karatzas and Shreve (1991), p. 289. The key is to note that the iterative construction of a solution follows through if we replace $b(s,x)$ and $\sigma(s,x)$ by $b(s,x,\omega)$ and $\sigma(s,x,\omega)$ in the definition of $X^{(k)}$. If, for fixed x, $(s,\omega)\mapsto b(s,x,\omega)$ and $(s,\omega)\mapsto \sigma(s,x,\omega)$ are adapted processes, the resulting process is still adapted. The remainder of the proof is identical. (See also Port and Stone (1991), Theorem V.7)

Proof of Proposition 1

Let the breaks in the drift arrive randomly at the stopping times ν_k with corresponding arrival rates $\lambda_k(t;\omega)$. In other words, let $k\in\{0,1,\ldots,n\}$ describe the regime in which the drift coefficient is $\alpha - k\Delta\alpha$ and the hazard rate at t for moving from state k to state $k+1$ is given by $\lambda_k(t,\omega)$. Since $0=\nu_0\le\nu_1\le\cdots\le\nu_k$, $\lambda_k(t;\omega)=0$ if $t<\nu_{k-1}(\omega)$. The value function for this problem is then given by:

$$J(x,k,t)=\sup_{\tau\ge t}\mathbb{E}[e^{\gamma\tau}(x_\tau - C)|x_t=x,k_t=k]$$

where $k_t\in\{0,1,\ldots,n\}$ marks the regime one is at. Heuristically one has the following Bellman equation:

$$\begin{aligned}J(x,k,t) &= \max\Big\{x-C,(1+\gamma dt)^{-1}\{\lambda_k(t)dt\mathbb{E}[J(x+dx,k+1,t)|x]+\\ &\qquad (1-\lambda_k(t)dt)\mathbb{E}[J(x+dx,k,t)|x]\}\Big\},\quad k\le n-1\end{aligned}$$

and $J(x,n,t)=\underline{J}(x)$, which is the value function for the optimal stopping problem when the log-linear diffusion has the lowest drift. In the continuation region, the second argument in the right-hand expression is the largest of the two and it can be seen that the value function satisfies:

$$(\gamma+\lambda_k(t))J(x,k,t)\ge \mathcal{A}_kJ(x,k,t)+J_t(x,k,t)+\lambda_k(t)J(x,k+1,t).$$

where $\mathcal{A}_k$ is the infinitesimal generator for a log-linear diffusion with drift coefficient $\alpha-k\Delta\alpha$. The left-hand side indicates the loss from waiting one infinitesimal instant, whereas the right-hand side stands for the benefit of waiting one infinitesimal instant — the expected appreciation in the value function. This expression holds in the continuation region and the typical

$$J(z_k(t),k,t)=z_k(t)-C,\quad \forall t\quad \text{(value matching)}$$
$$J_x(z_k(t),k,t)=1,\quad \forall t\quad \text{(smooth fit)}$$

implicitly define the thresholds z_k.

More rigorously,[30] let $J : \mathbb{R}_{++} \times \{1, \ldots, n\} \times \mathbb{R}_+ \rightarrow \mathbb{R}$ be twice differentiable on its first argument with an absolutely continuous first derivative such that:

1. $J(x, k, t) \geq x - C$;

2. $-\gamma J(x, k, t) + \mathcal{A}_k J(x, k, t) + J_t(x, k, t) + \lambda_k(t)(J(x, k+1, t) - J(x, k, t)) \leq 0$, with equality if $J(x, k, t) > x - C$;

3. $\forall s < \infty, \mathbb{E}[\int_0^\infty e^{-\gamma t} J_x(x_t, k_t, t) x_t^2 dt] < \infty$

Let $\mathcal{S}_k = \{(x, t) : J(x, k, t) \leq x - C\}$ be the stopping region when the regime is k and consider $\tau^* = \inf\{t : x_t \in \mathcal{S}_{k_t}\}$. Then

$$J(x, k, t) = \sup_{\tau \geq t} \mathbb{E}[e^{-\gamma\tau}(x_\tau - C) | x_t = x, k_t = k]$$

and τ^* attains the supremum.

To see this, consider a stopping time τ and let $\tau_m = \tau \wedge m$. Then (2), (3) and Dynkin's formula (Rogers and Williams (1994), pp. 252-4) deliver that

$$J(x, k, t) \geq \mathbb{E}[e^{-\gamma\tau_m} J(x_{\tau_m}, k_{\tau_m}, \tau_m) | x_t = x, k_t = k].$$

Using (1):

$$J(x, k, t) \geq \mathbb{E}[e^{-\gamma\tau_m}(x_{\tau_m} - C) | x_t = x, k_t = k].$$

By Fatou's Lemma, $\liminf_m \mathbb{E}[e^{-\gamma\tau_m}(x_{\tau_m} - C) | x_t = x, k_t = k] \geq \mathbb{E}[e^{-\gamma\tau}(x_\tau - C) | x_t = x, k_t = k]$ and we have that

$$J(x, k, t) \geq \mathbb{E}[e^{-\gamma\tau}(x_\tau - C) | x_t = x, k_t = k].$$

for an arbitrary stopping time τ. Using (2) and (3) plus Dynkin's formula one can then obtain that τ^* attains the supremum. The value matching and smooth pasting conditions are then consequences of J being C^1. As explained earlier, these two conditions implicitly define the thresholds $z_k(t)$.

That $z_k(t) > z_{k+1}(t), \forall t$ can be seen in the following manner. Let $x_t(x, k)$ be the process initialized at the level x and regime k. Since the drifts in successive states are strictly smaller, a comparison result such as the one in Karatzas and Shreve (1991), Proposition V.2.18, or Port and Stone (1991), Theorem V.54, can be established to show that:

$$e^{-\gamma t}(x_t(x, k) - C) > e^{-\gamma t}(x_t(x, k+1) - C), \quad \forall t \quad \mathbb{P}\text{-a.s.}$$

[30]The reasoning is in the spirit of similar arguments in Kobila (1993) and Scheinkman and Zariphopoulou (2001).

This should be enough to imply that the maximum attainable value is decreasing in k:

$$J(x,k,t) > J(x,k+1,t), \quad \forall t.$$

Consequently,

$$J(z_k(t),k,t) > J(z_k(t),k+1,t), \quad \forall t.$$

So, stopping at regime k implies stopping at regime $k+1$, whereas the opposite does not hold. This suffices to argue that

$$z_k(t) > z_{k+1}(t), \quad \forall t \quad k \in \{1,\ldots,n-1\}. \qquad \blacksquare$$

Proof of Proposition 2

<u>Step 1</u>: (Optimal policy characterization) As in Proposition 1, the value function characterizes the thresholds. Notice, though, that at any instant t the probability that another individual's latent utility process hits the stopping region in the next infinitesimal instant, given that it has not occurred so far, is negligible, since this process is a diffusion. As time goes by, though, the likelihood that such an event occurs increases toward one and the value of staying should decrease accordingly. So, we require the function in the limit to agree with the value function in the next regime, which ultimately brings it to the lowest drift regime. Let $J^i(x,k,t) = \sup_{\tau \geq t} \mathbb{E}[e^{\gamma\tau}(x^i_\tau - C^i)|x^i_t = x, k^i_t = k]$ be the value function for individual $i \in I$. Following the steps in Proposition 1, one can see that

1. $J^i(x,k,t) \geq x - C^i$;
2. $-\gamma J^i(x,k,t) + \mathcal{A}^i_k J^i(x,k,t) + J^i_t(x,k,t) \leq 0$, with equality if $J^i(x,k,t) > x - C^i$;
3. $\lim_{t\to\infty} J^i(x,k,t) = \underline{J}^i(x)$.

where $\underline{J}^i(x)$ is the value function for the optimal stopping problem with the lowest drift log-linear diffusion.

As in Proposition 1, we have the *value matching* and *smooth pasting* conditions determining the relevant thresholds:

$$J^i(z^i_m(t),m,t) = z^i_m(t) - C^i, \quad \forall t \quad \text{(value matching)}$$
$$J^i_x(z^i_m(t),m,t) = 1, \quad \forall t \quad \text{(smooth fit)}$$

As before $z^i_m(t) > z^i_{m+1}(t), \quad \forall t \quad m \in \{1,\ldots,n-1\}$.

<u>Step 2</u>: (Stopping times are an increasing sequence) Notice that, by definition, $\tau_0 \leq \tau_1 \leq \cdots \leq \tau_I$ and consequently form an increasing sequence of stopping times.

<u>Step 3</u>: (At each stage at least one agent stops) $\forall k \in I, \exists j : \tau^*_j = \tau_k$. Take a stopping time τ_k. There are two possibilities, represented by two disjoint subsets of Ω, say Ω_1 and Ω_2:

1. Ω_1. The vector process $A_{k-1}x_t$ hits $A_{k-1}\mathcal{S}_{I+1-\mathbf{1}'A_{k-1}\mathbf{1}}$ where $(\exists i \in I : x^i \geq z^i_k$ and $\forall j \neq i, x^j < z^j_{k+1})$. In this case, $\tau^*_i(\omega) = \tau_k(\omega)$ (provided i hasn't stopped yet), $\forall \omega \in \Omega_1$.

2. Ω_2. The above does not happen. In this case, $\exists j : z^j_{k+1} \leq x^j_{\tau_k}$ (provided j hasn't stopped yet). In this case it can be seen that $\tau_{k+1} = \tau_k$. Then, $x^j_{\tau_k} = x^j_{\tau_{k+1}} \geq z^j_{k+1}$ and this implies that $\tau^*_j(\omega) = \tau_{k+1}(\omega) = \tau_k(\omega), \forall \omega \in \Omega_2$.

This means that, at each stopping time τ_k, the drift of x^i drops by $\Delta\alpha^i/(I-1)$.
STEP 4: (τ^*_i is optimal) Apply Proposition 1. ■

Proof of Proposition 3

STEP 1: (The strategy profile is an equilibrium) Set $\nu = \tau^*_j$ in Proposition 2. Consider $\overline{\tau}^i = \inf\{t : x_t > z_i(t)\}$, where $z_i(t)$ is obtained as in Proposition 2. Agent i should use $\overline{\tau}_i$ on $\{\overline{\tau}^i < \tau^*_j\}$ and $\inf\{t > \tau^*_j : x^i_t > \underline{z}_i\}$ on the complementary set.

Now notice that:

$$x^i_{\tau_{\mathcal{S}}} = z_i(t) \Rightarrow \overline{\tau}_i = \tau_{\mathcal{S}}$$

When the vector process hits $\mathcal{S}$ on the subset where $x^i = z_i(t)$, the hitting times for the vector process to reach $\mathcal{S}$ and for the component process to hit $z_i(t)$ coincide. Since $\tau^*_j \geq \tau_{\mathcal{S}}$ by construction, we should also conclude that:

$$\{x^i_{\tau_{\mathcal{S}}} = z_i(t)\} \subset \{\overline{\tau}_i \leq \tau^*_j\}$$

Agent i should then use $\overline{\tau}_i$ (which coincides with $\tau_{\mathcal{S}}$ on this set).

On the other hand,

$$x^i_{\tau_{\mathcal{S}}} \neq z_i(t) \Rightarrow \begin{cases} \overline{\tau}_i > \tau_{\mathcal{S}} \\ (x^j_{\tau_{\mathcal{S}}} > \underline{z}_j \Rightarrow \tau^*_j = \tau_{\mathcal{S}}) \end{cases} \Rightarrow \overline{\tau}_i > \tau^*_j$$

So, we are in the complementary set, in which it is sensible to use $\inf\{t > \tau^*_j : x^i_t > \underline{z}_i\} = \inf\{t > \tau_{\mathcal{S}} : x^i_t > \underline{z}_i\}$.
STEP 2: (The equilibrium is unique) To see that this is the *unique* equilibrium, notice that

1. This is the unique equilibrium in which $x^1_{\tau^*_1 \wedge \tau^*_2} = z_1(t)$ or $x^2_{\tau^*_1 \wedge \tau^*_2} = z_2(t)$. In other words, any equilibrium profile of stopping strategies will have at least one stopper in the first round of exits at his or her threshold;

2. If there is another equilibrium, it should then involve first stoppers quitting at points lower than their initial thresholds. If only one agent drops, this can be shown to be suboptimal according to the reasoning of Proposition 2. If both stop at the same time and since $\mathbb{P}(\{\epsilon_1 = \epsilon_2\}) = 0$, there is an incentive for one of the agents to deviate and wait. ■

Proof of Proposition 4

Let $\mathcal{S} = \{(\mathbf{x}, t) \in \mathbb{R}^I_{++} \times \mathbb{R}_+ : \exists i$ such that $x^i \geq z^i_1 = z(\alpha^i, \sigma^i, C^i, \gamma^i, \Delta\alpha^i, t)\}$ and $\tau_{\mathcal{S}} = \inf\{t > 0 : \mathbf{x}_t \in \mathcal{S}\}$. Since the sample paths are continuous $\mathbb{P}$-almost surely, by Theorem 2.6.5 in Port and Stone (1978) the distribution of $(x_{\tau_{\mathcal{S}}}, \tau_{\mathcal{S}})$ will be concentrated on $\partial\mathcal{S}$. Also, it is true that $\mathbb{P}(\tau_{\mathcal{S}} < \infty) > 0$.

(Sufficiency) If there are endogenous effects, $z^i_1(t) > z^i_2(t), \forall t \quad i \in I$. There will be simultaneous exit whenever $z^i_1 \geq x^i_{\tau_{\mathcal{S}}} \geq z^i_2$, for some $i \in I$. This has positive probability as long as $z^i_1(\tau_{\mathcal{S}}) > z^i_2(\tau_{\mathcal{S}}), i \in I$. To see this, first notice that the latent utilities process can be represented as the following diffusion process with killing time at $\tau_{\mathcal{S}}$:

$$dx^i_t = \alpha^i x^i_t dt + \sum_{j \in I} \tilde{\sigma}_{ij} dB^j_t, \quad i = 1, \ldots, I$$

where $\mathbf{B}_t$ is an I-dimensional Brownian motion (with independent components) and $\tilde{\sigma}_{I \times I} = [\tilde{\sigma}_{ij}]$. Let $\partial\mathcal{S}_H = \{(\mathbf{x}, t) \in \partial\mathcal{S} : z^i_1(t) \geq x^i \geq z^i_2\}$. By Corollary II.2.11.2 in Gihman and Skorohod (1972) (p. 308), one gets that $\mathbb{P}[(\mathbf{x}_{\tau_{\mathcal{S}}}, \tau_{\mathcal{S}}) \in \partial\mathcal{S}_H] = u(\mathbf{x}, t)$ is an $\mathcal{A}$-harmonic function in $\mathcal{C} = \mathcal{S}^c$. In other words,

$$\begin{aligned} &\mathcal{A}u(\mathbf{x}) + u_t(\mathbf{x}, t) = 0 \text{ in } \mathcal{C} \\ &u(\mathbf{x}, t) = 1 \text{ if } (\mathbf{x}, t) \in \partial\mathcal{S}_H \\ &u(\mathbf{x}, t) = 0 \text{ if } (\mathbf{x}, t) \in \partial\mathcal{S} \backslash \partial\mathcal{S}_H \end{aligned}$$

where

$$\mathcal{A}f = \sum_{i \in I} \alpha^i x_i \frac{\partial f}{\partial x_i} + \frac{1}{2} \sum_{\substack{i,j \in I \\ i \neq j}} (\tilde{\sigma}\tilde{\sigma}')_{ij} x_i x_j \frac{\partial^2 f}{\partial x_i \partial x_j}.$$

is the infinitesimal generator associated with the above diffusion. By the Minimum Principle for elliptic operators (see Proposition 4.1.3 in Port and Stone (1978)), if u attains a minimum (which in this case would be zero) on $\mathcal{C}$, it is constant on $\mathcal{C}$. This would in turn imply that $\forall (\mathbf{x}, t) \in \mathcal{C}, u(\mathbf{x}) = \mathbb{P}[(\mathbf{x}_{\tau_{\mathcal{S}}}, \tau_{\mathcal{S}}) \in \partial\mathcal{S}_H | \mathbf{x}_0 = \mathbf{x}] = 0$. But by Proposition 2.3.6 in Port and Stone (1978), one can deduce that $u(\mathbf{x}, t) = \mathbb{P}[(\mathbf{x}_{\tau_{\mathcal{S}}}, \tau_{\mathcal{S}}) \in \partial\mathcal{S}_H | \mathbf{x}_0 = \mathbf{x}] \neq 0$.

(Necessity) If there are no endogenous effects, one agent's drift is never affected by the exit of other agents. Each agent's decision is given by $\tau^*_i = \inf\{t \in \mathbb{R}_+ : x^i_t > z^i = z(\alpha^i, \sigma^i, C^i, \gamma^i)\}$. There will be clustering only if $\tau^*_i = \tau^*_j, i \neq j$. The state-variable vector can be represented as above until the killing time $\tau_{\mathcal{S}}$. Then, there will be clustering only if $\mathbf{x}_t$ hits $\mathcal{S}$ at the point $(z^i)_{i \in I}$. But in $I \geq 2$ dimensions any one-point set A is polar with respect to a Brownian motion, i.e.,

$\mathbb{P}[\tau_A < \infty] = 0$ where τ_A is the hitting time for A (Proposition 2.2.5 in Port and Stone (1978)). So, $\mathbb{P}[\tau_i^* = \tau_j^*, i \neq j] = 0$. ■

Proof of Proposition 5

If there are no endogenous effects, the equilibrium strategies are characterized by the thresholds $z(\alpha, \sigma, \gamma, C)$. The marginal distribution for these are then (possibly defective) Inverse Gaussian distributions, so that $\mathbb{P}(\tau \leq t|\tau < \infty; x_0, z, \alpha, \sigma, I)$ is given by:

$$\Phi\left(\frac{\log(\frac{z}{x_0}) - |\alpha - \frac{\sigma^2}{2}|t}{\sigma\sqrt{t}}\right) - e^{\frac{2|\alpha - \frac{\sigma^2}{2}|(\log(\frac{z}{x_0}))}{\sigma^2}}\Phi\left(\frac{-\log(\frac{z}{x_0}) - |\alpha - \frac{\sigma^2}{2}|t}{\sigma\sqrt{t}}\right)$$

and

$$\mathbb{P}(\tau < \infty) = \begin{cases} 1 & \text{if } \alpha - \sigma^2/2 > 0 \\ \exp\left(\frac{-2\log(z/x_0)|\alpha - \sigma^2/2|}{\sigma^2}\right) & \text{otherwise} \end{cases}$$

(see Whitmore (2006)). Notice that the expression does not depend on I and this completes the proof. ■

Proof of Theorem 2

We start out by proving the conditions for Lyapunov's Central Limit Theorem for an arbitrary combination of $\mathbf{y}_{n,\Delta_{N,1}}$ and $\mathbf{y}_{n,\Delta_{N,2}}$.

<u>Step 1</u>: (Lyapunov's CLT) First, notice that, $\forall \delta > 0$ and $(\alpha_1, \alpha_2, \alpha_3) \in \mathbb{R}^3$,

$$\begin{aligned} \mathbb{E}[|\sum_{i=1,2,3} \alpha_i \mathbf{y}_{n,\Delta_{N,i}}|^{2+\delta}] &\leq 3\max_{i=1,2,3}(|\alpha_i|)\mathbb{E}[|\max_{i=1,2,3}(\mathbf{y}_{n,\Delta_{N,i}})|^{2+\delta}] \\ &\leq 3\max_{i=1,2,3}(|\alpha_i|)\mathbb{E}[\max_{i=1,2,3}(\mathbf{y}_{n,\Delta_{N,i}})] < \infty \end{aligned}$$

Let

$$\zeta(\Delta_{N,1}, \Delta_{N,2}, \Delta_{N,3}) = \begin{bmatrix} \alpha_1 & \alpha_2 & \alpha_3 \end{bmatrix} \text{var} \begin{bmatrix} \mathbf{y}_{n,\Delta_{N,1}} \\ \mathbf{y}_{n,\Delta_{N,2}} \\ \mathbf{y}_{n,\Delta_{N,3}} \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{bmatrix}$$

and observe that, provided $(\alpha_1, \alpha_2, \alpha_3) \neq (0, 0, 0)$, $\zeta(0, 0, 0) = 0 \Leftrightarrow p(0) = 0$ (since $p(0) = 0 \Leftrightarrow v(0) = 0$). Consider then

$$L_{n,N} = \frac{\sum_{i=1,2,3} \alpha_i (\mathbf{y}_{n,\Delta_{N,i}} - p(\Delta_{N,i}))}{\sqrt{N\zeta(\Delta_{N,1}, \Delta_{N,2}, \Delta_{N,3})}}.$$

For Lyapunov's Condition to be satisfied one needs to be able to state that

$$\lim_N \sum_{n=1}^{N} \mathbb{E}|L_{n,N}|^{2+\delta} = 0$$

for some $\delta > 0$. That this is the case can be seen because

$$\textstyle\sum_{n=1}^{N} \mathbb{E}|L_{n,N}|^{2+\delta} = \overbrace{N^{1-(1+\delta/2)(1-\kappa)}}^{A} \overbrace{(N^{\kappa}\zeta(\Delta_{N,1}, \Delta_{N,2}, \Delta_{N,3}))^{-(1+\delta/2)}}^{B} \times$$
$$\underbrace{\times\, \mathbb{E}[|\sum_{i=1,2,3} \alpha_i(\mathbf{y}_{n,\Delta_{N,i}} - p(\Delta_{N,i}))|^{2+\delta}]}_{C}$$

We set $\kappa = \varepsilon$ if $p(0) = 0$ (Assumption 1 holds) and $\kappa = 0$, otherwise.

For C, observe that

$$C \;\leq \textstyle\sum_{i=1,2,3}[\mathbb{E}[|\alpha_i \mathbf{y}_{n,\Delta_{N,i}}|]^{2+\delta} + (\alpha_i p(\Delta_{N,i}))^{2+\delta}] \leq \sum_{i=1,2,3}[\alpha_i^{2+\delta} + (\alpha_i p(\Delta_{N,i}))^{2+\delta}] \xrightarrow{N}$$

$$\longrightarrow (1 + p(0)^{2+\delta}) \textstyle\sum_{i=1,2,3} \alpha_i^{2+\delta}.$$

With respect to B, imposing $\kappa = 0$ and assuming that $p(0) > 0$, one has

$$\zeta(\Delta_{N,1}, \Delta_{N,2}, \Delta_{N,3}) \xrightarrow{N} \zeta(0, 0, 0) > 0.$$

In case $p(0) = 0$, notice that $\zeta(\cdot, \cdot, \cdot)$ is a function of $p(\cdot)$. This being differentiable, by the Mean Value Theorem one has

$$N^{\kappa}\zeta(\Delta_{N,1}, \Delta_{N,2}, \Delta_{N,3}) = N^{\kappa} \times \left(\zeta(0, 0, 0) + \partial\zeta(\widehat{\Delta}_{N,1}, \widehat{\Delta}_{N,2}, \widehat{\Delta}_{N,3})' \begin{bmatrix} \Delta_{N,1} \\ \Delta_{N,2} \\ \Delta_{N,3} \end{bmatrix}\right)$$

where $0 \leq \widehat{\Delta}_{N,k} \leq \Delta_{N,k}, k = 1, 2, 3$ and $\partial\zeta(\cdot, \cdot, \cdot)$ is the gradient vector for $\zeta(\cdot, \cdot, \cdot)$. We draw attention to the fact that $\lim_N \partial_k\zeta(\widehat{\Delta}_{N,1}, \widehat{\Delta}_{N,2}, \widehat{\Delta}_{N,3}) \times \Delta_{k,N} > 0, k = 1, 2, 3$. This is because

$$\Delta^{-1}\zeta(\Delta, 0, 0) = \Delta^{-1}\alpha_1^2 \text{var}(\mathbf{y}_{n,\Delta}) = \Delta^{-1}\alpha_1^2 \text{var}\left(\binom{I_n}{2}^{-1} \sum_{\{i,j\} \in \pi_n} \mathbb{I}_{\{\tau_\Delta^i = \tau_\Delta^j\}}\right) =$$

$$= \Delta^{-1}\alpha_1^2 \left[\binom{I_n}{2}^{-1} \text{var}\left(\mathbb{I}_{\{\tau_\Delta^i = \tau_\Delta^j\}}\right) + 2\binom{I_n}{2}^{-2} \sum_{\{i,j\},\{k,l\} \in \pi_n} \text{cov}\left(\mathbb{I}_{\{\tau_\Delta^i = \tau_\Delta^j\}}, \mathbb{I}_{\{\tau_\Delta^k = \tau_\Delta^l\}}\right)\right] \geq$$

$$\geq \Delta^{-1}\alpha_1^2 \binom{I_n}{2}^{-1} \left[p(\Delta)(1 - p(\Delta)) - 2p(\Delta)^2\right] \xrightarrow{\Delta \to 0} \alpha_1^2 \binom{I_n}{2}^{-1} p'_+(0) > 0$$

where the inequality follows because $\text{cov}\left(\mathbb{I}_{\{\tau_\Delta^i = \tau_\Delta^j\}}, \mathbb{I}_{\{\tau_\Delta^k = \tau_\Delta^l\}}\right) = \mathbb{E}\left(\mathbb{I}_{\{\tau_\Delta^i = \tau_\Delta^j\}}\mathbb{I}_{\{\tau_\Delta^k = \tau_\Delta^l\}}\right) - p(\Delta)^2 \geq -p(\Delta)^2$. The statement follows by analogy for the second and third arguments in $\zeta(\cdot, \cdot, \cdot)$. In this case,

$$N^{\varepsilon}\zeta(\Delta_{N,1}, \Delta_{N,2}, \Delta_{N,3}) \;=\; N^{\varepsilon}\left(\partial\zeta(\widehat{\Delta}_{N,1}, \widehat{\Delta}_{N,2}, \widehat{\Delta}_{N,3})' \begin{bmatrix} \Delta_{N,1} \\ \Delta_{N,2} \\ \Delta_{N,3} \end{bmatrix}\right) \xrightarrow{N} \sum_{i=1,2,3} a_i \partial_i^+ \zeta(0, 0, 0) > 0$$

where $a_i, i = 1, 2, 3$ are positive constants by Assumption 2. This suffices to show that B converges to a finite value.

If $p(0) > 0$, $A = N^{-\delta/2} \longrightarrow 0$. When $p(0) = 0$, we can drive A to zero by choosing $\delta > 0$ so that

$$\delta > 2((1-\varepsilon)^{-1} - 1) > 0.$$

Hence, Lyapunov's Condition

$$\lim_N \sum_{n=1}^{N} \mathbb{E}|L_{n,N}|^{2+\delta} = 0$$

holds. The Central Limit Theorem then asserts that

$$\sqrt{N}\frac{\sum_{i=1,2,3} \alpha_i(\overline{\mathbf{y}}_{\Delta_{N,i}} - p(\Delta_{N,i}))}{\sqrt{\zeta(\Delta_{N,1}, \Delta_{N,2}, \Delta_{N,3})}} \xrightarrow{\mathcal{L}} \mathcal{N}(0,1).$$

Since (α_1, α_2) is arbitrary, the Cramér-Wold device implies

$$\sqrt{N}\mathrm{var}\begin{bmatrix} \mathbf{y}_{n,\Delta_{N,3}} \\ \mathbf{y}_{n,\Delta_{N,2}} \\ \mathbf{y}_{n,\Delta_{N,1}} \end{bmatrix}^{-\frac{1}{2}} \left(\begin{bmatrix} \overline{\mathbf{y}}_{\Delta_{N,3}} \\ \overline{\mathbf{y}}_{\Delta_{N,2}} \\ \overline{\mathbf{y}}_{\Delta_{N,1}} \end{bmatrix} - \begin{bmatrix} p(\Delta_{N,3}) \\ p(\Delta_{N,2}) \\ p(\Delta_{N,1}) \end{bmatrix} \right) \xrightarrow{\mathcal{L}} \mathcal{N}(0, I_3)$$

<u>STEP 2</u>: (Delta Method) By the uniform delta method, one obtains that

$$\sqrt{N}\sigma_N^{-\frac{1}{2}}\left[\frac{\overline{\mathbf{y}}_{\Delta_{N,2}}}{\overline{\mathbf{y}}_{\Delta_{N,1}}} - \frac{p(\Delta_{N,2})}{p(\Delta_{N,1})} - \xi\left(\frac{\overline{\mathbf{y}}_{\Delta_{N,3}}}{\overline{\mathbf{y}}_{\Delta_{N,1}}} - \frac{p(\Delta_{N,3})}{p(\Delta_{N,1})}\right)\right] \xrightarrow{\mathcal{L}} \mathcal{N}(0,1)$$

where

$$\sigma_N = \begin{bmatrix}\frac{\xi p(\Delta_{N,3}) - p(\Delta_{N,2})}{p(\Delta_{N,1})^2} & \frac{1}{p(\Delta_{N,1})} & -\frac{\xi}{p(\Delta_{N,1})}\end{bmatrix} \mathrm{var}\begin{bmatrix} \mathbf{y}_{n,\Delta_{N,1}} \\ \mathbf{y}_{n,\Delta_{N,2}} \\ \mathbf{y}_{n,\Delta_{N,3}} \end{bmatrix} \begin{bmatrix} \frac{\xi p(\Delta_{N,3}) - p(\Delta_{N,2})}{p(\Delta_{N,1})^2} \\ \frac{1}{p(\Delta_{N,1})} \\ -\frac{\xi}{p(\Delta_{N,1})} \end{bmatrix}$$

and

$$\xi = \frac{a_2(a_2 - a_1)}{a_3(a_3 - a_1)}.$$

Notice that

$$\frac{p(\Delta_{N,k})}{p(\Delta_{N,1})} = \frac{p(0) + \Delta_{N,k}p'(0) + \frac{1}{2}\Delta_{N,k}^2 p''(0) + O(\Delta_{N,k}^3)}{p(0) + \Delta_{N,1}p'(0) + \frac{1}{2}\Delta_{N,1}^2 p''(0) + O(\Delta_{N,1}^3)}$$

and, if $p(0)=0$, one obtains that

$$\begin{aligned}\frac{p(\Delta_{N,k})}{p(\Delta_{N,1})}-\frac{a_k}{a_1} &= \frac{\Delta_{N,k}p'(0)+\frac{1}{2}\Delta_{N,k}^2p''(0)+O(\Delta_{N,k}^3)}{\Delta_{N,1}p'(0)+\frac{1}{2}\Delta_{N,1}^2p''(0)+O(\Delta_{N,1}^3)}-\frac{a_k}{a_1}= \\ &= \frac{a_k+p''(0)a_k^2N^{-\epsilon}+O(N^{-2\epsilon})}{a_1+p''(0)a_1^2N^{-\epsilon}+O(N^{-2\epsilon})}-\frac{a_k}{a_1}=\frac{p''(0)a_ka_1(a_k-a_1)N^{-\epsilon}+O(N^{-2\epsilon})}{a_1[a_1+p''(0)a_1^2N^{-\epsilon}+O(N^{-2\epsilon})]}= \\ &= \frac{p''(0)a_ka_1(a_k-a_1)N^{-\epsilon}}{a_1^2[1+p''(0)a_1N^{-\epsilon}+O(N^{-2\epsilon})]}+O(N^{-2\epsilon})\overset{\text{(by Taylor's Expansion)}}{=} \\ &= \frac{p''(0)a_ka_1(a_k-a_1)N^{-\epsilon}}{a_1^2}\times[1-p''(0)a_1N^{-\epsilon}+O(N^{-2\epsilon})]+O(N^{-2\epsilon})= \\ &= \frac{p''(0)a_ka_1(a_k-a_1)N^{-\epsilon}}{a_1^2}+O(N^{-2\epsilon}).\end{aligned}$$

This delivers

$$\frac{\overline{\mathbf{y}}_{\Delta_{N,2}}}{\overline{\mathbf{y}}_{\Delta_{N,1}}}-\frac{p(\Delta_{N,2})}{p(\Delta_{N,1})}-\xi\left(\frac{\overline{\mathbf{y}}_{\Delta_{N,3}}}{\overline{\mathbf{y}}_{\Delta_{N,1}}}-\frac{p(\Delta_{N,3})}{p(\Delta_{N,1})}\right)=\frac{\overline{\mathbf{y}}_{\Delta_{N,2}}}{\overline{\mathbf{y}}_{\Delta_{N,1}}}-\frac{a_2}{a_1}-\xi\left(\frac{\overline{\mathbf{y}}_{\Delta_{N,3}}}{\overline{\mathbf{y}}_{\Delta_{N,1}}}-\frac{a_3}{a_1}\right)+O(N^{-2\epsilon}).$$

Noticing that $\sqrt{N}\sigma_N^{-\frac{1}{2}}=O(N^{\frac{1+\epsilon}{2}})$ gives us that

$$\sqrt{N}\sigma_N^{-\frac{1}{2}}\left[\frac{\overline{\mathbf{y}}_{\Delta_{N,2}}}{\overline{\mathbf{y}}_{\Delta_{N,1}}}-\frac{a_2}{a_1}-\xi\left(\frac{\overline{\mathbf{y}}_{\Delta_{N,3}}}{\overline{\mathbf{y}}_{\Delta_{N,1}}}-\frac{a_3}{a_1}\right)\right]\xrightarrow{\mathcal{L}}\mathcal{N}(0,1)$$

as long as $\epsilon>1/3$. If $p(0)>0$, it is possible to see that the statistic diverges using similar arguments. ■

Proof of Theorem 3

STEP 1: Consider the expected log-likelihood function conditioned on $\mathbf{w}$:

$$\mathcal{KL}(\psi,\hat{\psi},\mathbf{w})=\int\log[g(t;\hat{\psi},\mathbf{w})]g(t;\psi,\mathbf{w})dt$$

From the properties of the Kullback-Leibler information criterion (or relative entropy) for two probability distributions, it is obtained that $\hat{\psi}$ maximizes the expected log-likelihood if and only if $g(t;\hat{\psi},\mathbf{w})=g(t;\psi,\mathbf{w})$ (see Schervish (1995), Proposition 2.92). In particular, $\hat{\psi}=\psi$ is one such maximizer.

STEP 2: Take $\Psi=\{\psi,\hat{\psi}\}$. Also, let $w_i=(w_i^c,w_i^d)$ where w_i^c denotes the continuous random covariates and w_i^d, those with a discrete component. We now verify the conditions in Theorem 1 in Araújo and Mas-Colell (1978). For each fixed value in the support of $(w_i^d)_{i\in I}$:

1. Ψ is trivially Lindelöf since it is compact.[31]

[31] We could allow for Ψ to be a subset of $\mathbb{R}^K$. $\mathbb{R}^{2K}$ is Lindelöf since it is separable and metrizable and thus, second countable (see Aliprantis and Border (1999), Theorem 3.1). This implies that it is Lindelöf (see Aliprantis and Border (1999), p. 45).

2. The Kullback-Leibler information criterion is continuous on the parameters and $\mathbf{w}$. To obtain this result, notice that

$$g(;\psi) = -\frac{dG}{dt}(;\psi)$$

and

$$G(t;\psi) = \mathbb{P}(\tau > t;\psi) = \mathbb{E}[\mathbb{I}_{\{\sup_{s\leq t} x^i_s(\psi) - z(t;\psi) > 0 \text{ for some } i\}}] = \mathbb{E}[\phi((x_s)^t_{s=0})|x_0 = x]$$

The derivative of this last expression with respect to the parameters is well defined and can be obtained through Malliavin calculus (see Proposition 3.1 in Fournié, Lasry, Lebuchoux, Lions, and Touzi (1999) for the drift, for example). The assumption that $\alpha(\cdot)$ is a continuous function on the covariates achieves the result.

3. The derivative exists and is continuous since we assume that $\alpha(\cdot)$ is of class C^1 with respect to the continuous random variables.

4. Pick any product measure ν equivalent to the measure represented by the CDF $F_{\mathbf{w}^c}$. Since the latter is assumed to be continuous, its measure is absolutely continuous with respect to the Lebesgue measure and ν is also absolutely continuous with respect to the Lebesgue measure.

5. The Sondermann Condition in Araújo and Mas-Colell (1978) holds by assumption.

By Theorem 1 in Araújo and Mas-Colell (1978), there is at most one maximizer for the expected log-likelihood function $F_{\mathbf{w}^c}$-a.e. for each element in the support of $\mathbf{w}^d$ and we know that ψ maximizes it. The statement is easily extended $F_{\mathbf{w}}$-a.e. since the support of w^d is countable and the union of countable events with null measure — there being more than one maximizer — has zero measure. ■

References

Abreu, D., and M. Brunnermeier (1997): "Synchronization Risk and Delayed Arbitrage," *Journal of Financial Economics*, 66.

Aliprantis, C., and K. Border (1999): *Infinite Dimensional Analysis: A Hitchhiker's Guide*. Springer–Verlag, New York.

Araújo, A., and A. Mas-Colell (1978): "Notes on the Smoothing of Aggregate Demand," *Journal of Mathematical Economics*, 5.

Brock, W., and S. Durlauf (2001): "Interactions-Based Models," in *Handbook of Econometrics*, ed. by J. Heckman, and E. Leamer, vol. 5 of *Handbook of Econometrics*, chap. 55, pp. 3297–3380. Elsevier.

Brunnermeier, M., and J. Morgan (2004): "Clock Games: Theory and Experiment," Working Paper, Princeton University.

Chhikara, R., and J. Folks (1989): *The Inverse Gaussian Distribution*. Marcel Dekker.

Costa, D., and M. Kahn (2003): "Cowards and Heroes: Group Loyalty during the American Civil War," *Quarterly Journal of Economics*, 118.

Dixit, A., and R. Pindyck (1994): *Investment Under Uncertainty*. Princeton U. Press.

Eckstein, Z., and K. Wolpin (1989): "The Specification and Estimation of Dynamic Stochastic Discrete Choice Models: A Survey," *Journal of Human Resources*, 15.

Fournié, E., J.-M. Lasry, J. Lebuchoux, P.-L. Lions, and N. Touzi (1999): "Applications of Malliavin Calculus to Monte Carlo Methods in Finance," *Finance and Stochastics*, 3.

Gihman, I., and A. Skorohod (1972): *Stochastic Differential Equations*. Springer–Verlag.

Glaeser, E., and J. A. Scheinkman (2002): "Non-Market Interactions," in *Advances in Economics and Econometrics: Theory and Applications, Eight World Congress*, ed. by M. Dewatripont, L. Hansen, and S. Turnovsky. Cambridge University Press.

Glasserman, P. (2004): *Monte Carlo Methods in Financial Engineering*. Springer.

Granovetter, M. (1978): "Threshold Models of Collective Behavior," *American Journal of Sociology*, 83.

He, H., W. Keierstead, and J. Rebholz (1998): "Double Lookbacks," *Math. Finance*, 8.

Hong, H., J. Kubik, and J. Stein (2004): "Social Interactions and Stock-Market Participation," *Journal of Finance*, 59.

Hopenhayn, H., and F. Squintani (2004): "Preemption Games with Private Information," UCLA Working Paper.

Huang, J., M. Subrahmanyam, and G. Yu (1996): "Pricing and Hedging American Options: A Recursive Integration Method," *Review of Financial Studies*, 9.

Iyengar, S. (1985): "Hitting Lines with Two-Dimensional Brownian Motion," *SIAM Journal of Applied Mathematics*, 45.

Ju, N. (1998): "Pricing an American Option by Approximating Its Early Exercise Boundary as a Multipiece Exponential Function," *Review of Financial Studies*, 11.

Karatzas, I., and S. Shreve (1991): *Brownian Motion and Stochastic Calculus*. Springer.

Kobila, T. (1993): "A Class of Solvable Stochastic Investment Problems Involving Singular Controls," *Stochastics and Stochastics Reports*, 43.

Lonn, E. (1928): *Desertion During the Civil War*. The Century Co.

Mamer, J. (1987): "Monotone Stopping Games," *Journal of Applied Probability*, 24.

Manski, C. (1993): "Identification of Endogenous Social Effects: The Reflection Problem," *Review of Economic Studies*, 60.

McManus, C. (1992): "How Common Is Identification in Parametric Models?," *Journal of Econometrics*, 53.

Merton, R. (1990): *Continuous-Time Finance*. Blackwell Publishers.

Morris, S. (1995): "Co-operation and Timing," CARESS Working Paper #95-05.

Orrenius, P. (1999): "The Role of Family Networks, Coyote Prices and the Rural Economy in Migration from Western Mexico: 1965-1994," Discussion Paper, Federal Reserve Bank of Dallas.

Pakes, A., and D. Pollard (1989): "Simulation and the Asymptotics of Optimization Estimators," *Econometrica*, 57.

Port, S., and C. Stone (1978): *Brownian Motion and Classical Potential Theory*. Academic Press.

——— (1991): *Stochastic Integration and Differential Equations. A New Approach*. Springer–Verlag.

Rebholz, J. (1994): "Planar Diffusions with Applications to Mathematical Finance," Ph.D. Dissertation, University of California, Berkeley.

Rogers, L., and D. Williams (1994): *Diffusions, Markov Processes and Martingales*, vol. 1. Cambridge University Press.

Scheinkman, J. A., and T. Zariphopoulou (2001): "Optimal Environmental Management in the Presence of Irreversibilities," *Journal of Economic Theory*, 96.

Schelling, T. (1960): *The Strategy of Conflict*. Harvard University Press.

Schervish, M. (1995): *Theory of Statistics*. Springer.

Sirakaya, S. (2001): "Recidivism and Social Interactions," *Journal of the American Statistical Association*.

van den Berg, G., M. Lindeboom, and G. Ridder (1994): "Attrition in Longitudinal Panel Data and the Empirical Analysis of Dynamic Labour Market Behaviour," *Journal of Applied Econometrics*, 9(4).

Whitmore, G. (2006): "An Inverse Gaussian Model for Labour Turnover," *Journal of the Royal Statistical Society, Series A (General)*, 142.

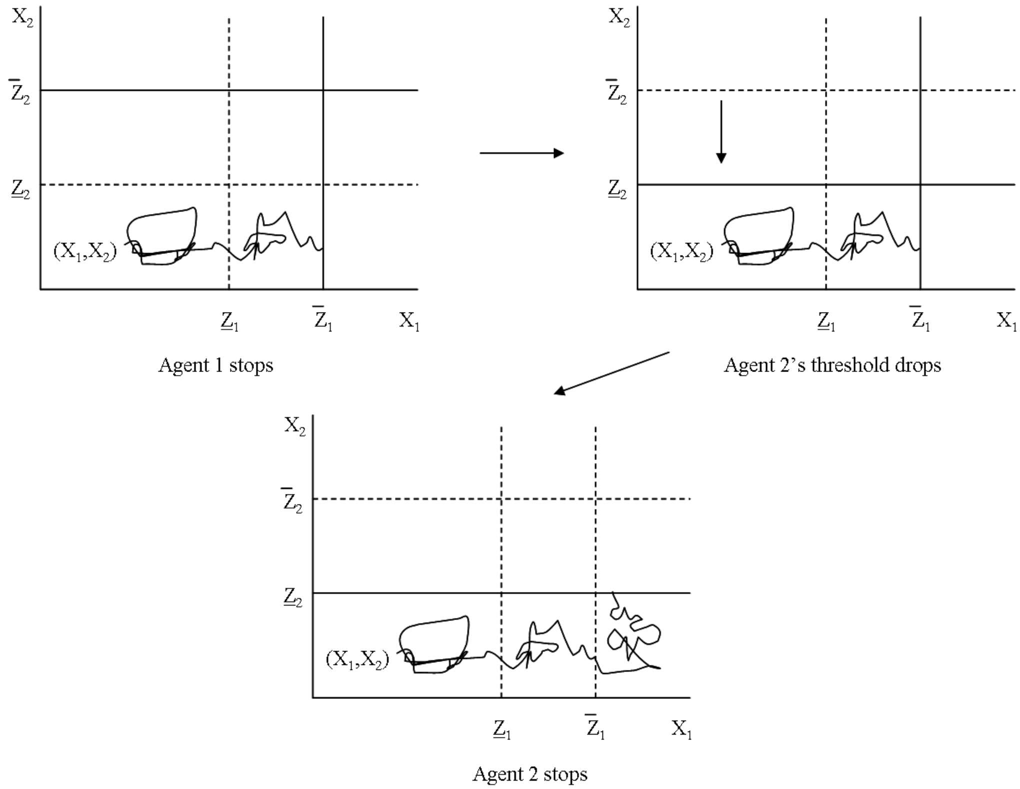

Figure 1: Sequential Stopping

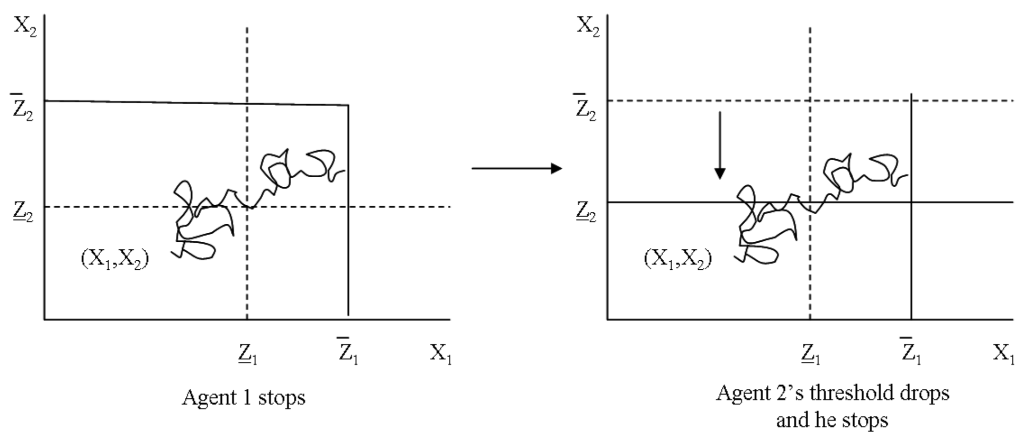

Figure 2: Simultaneous Stopping

Table 1: Individual Regressions

Dependent variable: log(Days Until Desertion)								
	Coef.	Std. Error	t	P> |t|	Coef.	Std. Error	t	P> |t|
Company Size	0.002	0.0006	3.43	0	0.0019	0.0006	3.08	0.00
Bounty Paid					0.0015	0.0006	2.36	0.02
Foreigner					-0.2144	0.0504	-4.25	0.00
Substitute					0.1874	0.0996	1.88	0.06
Age					-0.0095	0.0033	-2.86	0.00
Height					0.0256	0.0095	2.7	0.01
State Controls:	Yes				Yes			
Year Controls:	Yes				Yes			
Number of obs =	3337				3237			
R-squared =	0.2983				0.3076			

† Standard errors are robust standard errors.

Table 2: Test (Non-Battle Desertions)

		$\left(\frac{\overline{y}_{\Delta_2}}{\overline{y}_{\Delta_1}} - \frac{\Delta_2}{\Delta_1}\right) - \xi\left(\frac{\overline{y}_{\Delta_3}}{\overline{y}_{\Delta_1}} - \frac{\Delta_3}{\Delta_1}\right)$			
Δ_2/Δ_1	Δ_3/Δ_1	$\overline{y}_{\Delta_1}$	$\overline{y}_{\Delta_2}$	$\overline{y}_{\Delta_3}$	Test Statistic
2	3	0.002289	0.002693	0.002859	-6.25
2	5	0.002289	0.002693	0.003358	-11.31
2	10	0.002289	0.002693	0.004210	-13.18
3	4	0.002289	0.002859	0.003216	-13.05
3	5	0.002289	0.002859	0.003358	-16.99
3	10	0.002289	0.002859	0.004210	-20.95

† $\Delta_1 = 1$ day. All 303 companies were used.

Table 3: Model Estimation

$\hat{\alpha}$	$\hat{\Delta\alpha}$	$\hat{\sigma}$	$\hat{\rho}$
0.0438	0.0050	5.8610	0.1008
(0.0113)	(0.0000)	(0.0219)	(0.0040)
(per year)	(per year)	(per year)	

† Initial position = 0.1. Exit cost = 1. Discount rate = 5% *per year*.