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 07-017

"Inference in a Synchronization Game with Social Interactions"

by

Aureo de Paula

http://ssrn.com/abstract=988997

Inference in a Synchronization Game with Social Interactions *

Áureo de Paula †

First Version: October 2004

This version: May, 2007

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 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 involving 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. 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 choice 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 then concentrate efforts on a subgroup of community members 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

[1] Manski [31] 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 [16].

participation by individuals, a policy-maker may prefer to identify and directly act on such defining variables.

In a timing framework, statistical inference would typically involve 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 [9] and Sikaraya [45] and suggested in Brock and Durlauf [6]), 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 studied in our paper — 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[3].

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 social welfare program participation, stock market participation (Hong, Kubik and Stein [21]), migration (Orrenius [34]) and even crime recidivism (see for instance the empirical investigation by Sirakaya [45], 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 utilized in the finance and investment literature for the analysis of securities derivatives and real options. 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 [43], which discusses the timing of mob formation. Our paper also relates to the threshold models of collective behavior in Granovetter [20], 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

[3] See for instance Van den Berg, Lindeboom and Ridder [46].

connected to the one developed in Mamer [30] for a discrete-time setting and in a different framework[4]. 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 (see Fudenberg and Tirole [14] (Ch.13) for a review of the literature on such games). This literature is nonetheless 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 [22]. In their case the payoff flows evolve discontinuously (according to Poisson processes) 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 [31]). This difficulty explains our search for structure in the specific situation under analysis. Strategic interactions also pose an additional problem that may hinder identification and estimation: that of multiple equilibria. We circumvent this issue in the present article by focusing on a specific equilibrium.

In order to construct the model and perform inference we use the tools of continuous-time optimal stopping problems which appear in the investment and finance literatures. The basic ingredients are explained in Dixit and Pindyck [10]. 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 [45], in which the author investigates duration dependence in timing of crime recidivism. Brock and Durlauf [6] 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 and focus instead on contextual neighborhood variables. In their analysis of group homogeneity and desertion, Costa and Kahn [9] 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 one may contend that it is still not clear whether the effect captured is one of endogenous interactions or

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

correlated effects.

The structure of the paper is as follows. In the next section we present the general model and establish 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 of the world $\omega \in \Omega$ is chosen according to a probability law $\mathbb{P}$. There are I agents. These agents 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 $i \in I$ chooses to abandon the activity 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 $u_i(x^i_{\tau^i}, \tau^i)$. The stopping strategies are represented by $\tau_i : \Omega \to [0,T]$, a (possibly infinite) stopping time with respect to an individual filtration $\mathbb{F}^i = (\mathcal{F}^i_t)_{t\in[0,T]}$[5] 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[6]. Rigorously, a filtration is a sequence of σ-algebras that specifies how events are revealed over time. Intuitively, a filtration records one's knowledge about the state of the world ω as time evolves. The information to individual i at instant t consists then of all events summarized in the collection $\mathcal{F}^i_t$. 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 modeling assumptions one imposes on what each agent knows or not as time evolves.

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) which may depend directly on the participation of the remaining individuals in the group. This direct influence represents the endogenous effects in our model.

[5] 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 $t \in [0,T]$, $\{\omega : \tau(\omega) \leq t\} \in \mathcal{F}_t$. Some authors use the term Markov time for this definition and refer to stopping times as finite Markov times. In this paper we use infinite and finite stopping times respectively for these objects. Intuitively they represent stopping strategies that rely solely on past information.

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

Let θ_t^i be the process representing the fraction of the population (excluding agent i) that has abandoned the activity before time t. In other words, $\theta_t^i = \sum_{s=1,s\neq i}^{I} \mathbb{I}_{\{\tau_s<t\}}/(I-1)$ (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 according to their preferences. Throughout we assume that $\theta_t^i \in \mathcal{F}_t^i$: one knows how many players stopped up to (but excluding) the current instant. Each individual state variable x_t^i is assumed Markovian and is allowed to differ across individuals. The structure for the multi-person decision problem (payoffs, players, strategy spaces and information assumptions) is presented in the following definition.

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* $\tau : \Omega \to [0, T]$.

Having defined the basic structure of the problem, the idea is that each person i faces a decision problem that is mathematically represented by the following (individual) optimal stopping problem (where τ generically denotes a stopping time with respect to $(\mathcal{F}_t^i)_{t\in[0,T]}$):

$$\begin{cases} & & \mathbb{P}(x_t^i \in \Gamma | \mathcal{F}_s^i),\ \text{Markovian.} \\ V_i(x_i) = \sup_{\tau\in T_i} \mathbb{E}_{x_i}[u_i(x_\tau, \tau)] & \text{s.t.} & \theta_t^i = \sum_{s=1,s\neq i}^{I} \mathbb{I}_{\{\tau_s<t\}}/I \\ & & x_0^i = x_i \end{cases} \tag{1}$$

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

In this paper the state variable is assumed to obey a transition law given by the following expression:

$$dx_t^i = \alpha^i(x_t^i, \theta_t^i, t)dt + \sigma^i(x_t^i, \theta_t^i, t)dW_t^i, \qquad x_0^i \sim F_0^i \tag{2}$$

where W_t^i is a Wiener process defined in the particular probability space we are considering and the drift and dispersion coefficients are assumed to be positive Borel-measurable functions. The initial distribution F_0^i is furthermore independent of the Brownian motion W_t^i. We impose no restrictions on the contemporaneous correlation between the Wiener processes[7]. The presence of any contemporaneous covariance accounts for the correlated effects

[7]One could alternatively represent W_t^i as a linear combination of an I-dimensional (independent) Brownian motion $\mathbf{B}_t$ for each i.

in the model.

One important aspect of the assumed law of motion is 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 in a convenient manner (in contrast, for instance, to the work by Hopenhayn and Squintani [22], where sample paths present discontinuities). Since the stochastic utility processes evolve continuously, the probability that a given individual reaches a stopping region in the state space between t and $t+\epsilon$ conditional on not having stopped before t vanishes as $\epsilon \to 0$. As a consequence, where a counterpart's state variable is located becomes immaterial to the decisions taken within the next infinitesimal period by the agent.

In order to assure that, given a profile of stopping times for each player, this stochastic differential equation has a (strong) solution, we impose the following assumptions on the drift and dispersion coefficients:

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[8]. 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. The following section analyzes the existence of equilibria for this game.

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

[8]An optional time with respect to a filtration $(\mathcal{F}_t)_{t\in[0,T]}$ is a positive random variable τ such that $\{\tau < t\} \in \mathcal{F}_t, \forall t$. A stopping time is easily seen to be an optional time. If the filtration is right-continuous, an optional time is a stopping time.

the other agents. 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 $(\tau_s)_{s\in I-\{i\}}$. A Nash Equilibrium[9] for the above game is then:

Definition 2 (Equilibrium) *A* Nash Equilibrium *for the* Synchronization Game *is a stopping time profile* $\tau^* = (\tau_i^*)_{i\in I}$ *such that:*

$$U_i(\tau^*) \geq U_i(\tau_i, \tau_{-i}^*), \forall i, \tau_i \text{ stopping time.}$$

In order 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 assumed to satisfy:*

- *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.*
- *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* $\partial_\theta \sigma(\cdot,\cdot,\cdot) \leq 0$.

[9] Since the strategies depend on information generated by the state variables and these are Markovian and since optimization follows Bellman's principle of optimatility 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. For a discussion of MPE, see Fundenberg and Tirole [14], chapter 13.

The Exponential Discounting Assumption (2) significantly 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, which is 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) will imply that changes in the profile of stopping decisions will affect the objective function only through the drift of the discounted gain function. Finally, the Complementarity Assumption (5) expresses the fundamental idea that higher participation makes the activity more attractive as well as increases the volatility of the returns. This assumption imposes the idea that one agent's action is a strategic complement to the others' actions. We are now ready to state the following result[10]:

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

Proof. See Appendix.

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

3. A Coordination Game

We now specialize the model and extend the analysis developed so far. Consider initially a hypothetical game where agents contemplate the possibility of exit. As before, a state variable x (which is assumed to evolve according to a certain stochastic process) 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 the reward function $\mathbb{E}^x[e^{-\gamma t}(x_t - C)]$.

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

[10]Mamer [30] obtains existence of equilibria in a similar (but more restrictive) game in discrete time through similar techniques.

according to the following law:

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

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. In other words, $\mathcal{F}_t = \sigma(x_s, \mathbb{I}_{\nu<s}, s \leq t)$. The initial condition is drawn from an independent distribution F_0 as in equation (2). Notice that 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 $\gamma > \alpha$.

Let $\overline{x}$ be the process corresponding to $\nu(\omega) = \infty, \forall \omega \in \Omega$ (i.e. a geometric Brownian motion with drift and diffusion 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 drift and diffusion coefficients $(\alpha - \Delta\alpha)x$ and σx). We can use dynamic programming to show that the optimal stopping times for these two processes are characterized by threshold levels $\overline{z} = z(\alpha, \sigma, C, \gamma)$ and $\underline{z} = z(\alpha - \Delta\alpha, \sigma, C, \gamma)$, where

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

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 [10], p.140-144). The agent will stop the process as soon as it hits the level z. For notational 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:

$$\begin{gathered} \{x \leq z \equiv z(\alpha, \sigma, C, \gamma; \Delta\alpha, t, \text{other parameters})\} \text{ if } t \leq \nu \\ \{x \leq \underline{z}\} \text{ if } t > \nu \end{gathered} \tag{5}$$

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*

$$\{x \leq z_{k-1}\} \text{ if } t \leq \nu_k, k = 1, \ldots, n$$
$$\{x \leq z_n\} \text{ if } t > \nu_n$$

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* $z_n \equiv \underline{z}$.

Proof. See Appendix.

Consider now a game with two agents indexed by $i = 1, 2$. They both contemplate an exit decision that will cost them $C_i, i = 1, 2$. In return, they collect a value $x_i, i = 1, 2$, just as in the previous setup. The difference is that now the latent utility process for one agent is negatively affected once the other agent decides to leave the activity. In analogy with the previous individual setup, we consider a situation in which 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.

In particular, consider all the above parameters indexed by i and the latent utility process, given by:

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

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, $\gamma > \alpha^i$. Notice that the contemporaneous correlation between the Brownian motions is left unconstrained and that $\Delta\alpha^i$ measures the external effect of the other agent's decision on i. As pointed out in the previous section, 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.

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) is represented. Since $\Delta\alpha^i > 0, i = 1, 2$, we should have $z_i(t) > \underline{z}_i, i = 1, 2$. As in the previous analysis, 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. A symmetric situation occurs if we interchange the agents roles.

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 a point where $x^2 \geq \underline{z}^2$. 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 which is not present in standard statistical models that would handle timing situations as the one at hand: the positive probability for simultaneous events even when time is observed continuously. If not properly accounted for, this can bias results towards 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

[11] Van den Berg, Lindeboom and Ridder [46], for instance, point to a negative duration dependence bias in estimates if simultaneity is left unaccounted.

immediately? Such considerations point to the importance of beliefs in these environments and are a relevant consideration in Hopenhayn and Squintani [22], 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 $\epsilon \to 0$ (as is the case with the hazard rate for an Inverse Gaussian random variable). 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. In order to state this result in more generality, assume as before that an exit decision costs an agent $C_i, i \in I$ in return for a payoff $x^i, i \in I$. The latent utility process is given by:

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

where τ^j is the stopping time adopted by the agent j. Notice that the external effect of other agents on i is given by $\Delta\alpha^i > 0$ 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_m^i :	$z(\alpha^i, \sigma^i, C^i, \gamma^i, m, \Delta\alpha^i, t)$ where $i, m \in I$
$\mathcal{S}_m$:	$\{(x^1, x^2, \ldots, x^I) \in \mathbb{R}_+^I : \exists i$ such that $x^i \geq z_m^i\}$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 :	$[a_{kl}^m]_{I\times I}$ where $a_{kl}^m = \mathbb{I}_{x_{\tau_m}^i < z_m^i}$ if $k = l = i$ and $a_{kl}^m = 0$ otherwise and $m \in I$

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

The idea is that the game starts out with no defection and agents hold the highest barrier $z_1^i(t)$ as the initial exit rule. The first exit then occurs at τ_1, which is the hitting time for the stopping region $\mathcal{S}_1$. As the vector process reaches this set, one or more agents will quit. This will shift the stopping thresholds down as well as the stopping region moves to $\mathcal{S}_2$.

In order to track the players who drop out at τ_1 we utilize the matrix A_1. This is a diagonal matrix with ones for those agents who did not drop out at τ_1 and zeros, otherwise. Proceeding analogously through further stopping rounds, defections will occur at the stopping times $\tau_.$ and $\mathbf{1}'A_.\mathbf{1}$ basically 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:*

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

Proof. See Appendix.

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

In this subsection, we argue that the above equilibrium is the only equilibrium which is robust to positive delays in information about the exit of others. As we drive this delay to zero, this parallels the intuitive idea that each agent observes 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.

If 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 Brunnemeier [1], Brunnemeier and Morgan [7] and Morris [33]). In fact, the following quote presents one of the earliest discussions of this problem:

> It is usually the essence of mob formation that the potential members have to know not only where and when to meet but just when to act so that they act in concert. (...) In this case the mob's problem is to act in unison without overt leadership, to find some common signal that makes everyone confident that, if he acts on it, he'll not be acting alone. (Schelling [43])

In order 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:*

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

Proof. See Appendix.

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 which is robust to such perturbations and corresponds to the one displayed in the last subsection.

We draw attention to the fact that other information structures would nonetheless be susceptible to multiple equilibria.[12] One point that is worth noting is that the occurrence of joint exit with positive probability is robust to the existence of multiple equilibria and the issue of equilibrium selection. Given this, we point out that some of our results in this and the subsequent section are robust even to the existence of multiple equilibria. In the next section, we discuss the empirical implications of the model.

4. Empirical Implications

In this section we investigate the empirical implications of the model. We have in mind a sample in which the unit of observation is a game[13] and N such units are recorded. Recall that, by endogenous effects, we mean the effect of other agents' participation on

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

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

the transition law for the individual state variables. Correlated effects refer to the possible contemporaneous correlation between the Brownian motions that drive the individual latent utility processes. 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.

Each agent's latent utility 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 the 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, $\alpha^i = \alpha(w_i)$. In benefit of 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 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

We are now in shape to look at the outcomes in the presence of interactions and correlated effects. The next proposition states that simultaneous departures only occur in the presence of endogenous effects.[14]

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

Proof. See Appendix.

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

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

equilibrium. Notice that traditional econometric models in duration analysis typically do not generate clustering in timing, i.e. the probability of simultaneous exit is zero and such incompatibility may provoke biased estimates and contaminate conclusions.

This result relies basically on the continuity of the sample paths for the stipulated process. If discontinuities are allowed, this may not hold any longer[15]. The problem would nonetheless be diluted if one knew the timing of such shocks. If one observes clustering in other moments, this is 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 ($\Delta\alpha > 0$).*

Proof. See Appendix.

Notice that the direction in which the equilibrium stopping times are affected is not clear. If, on the one hand, the presence of 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.

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

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

and allow the discretization to depend on the sample size.

Imagine that there are no endogenous interactions. In this case, for a small enough

[15] 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.

discretization, doubling the observation interval would roughly double the probability of 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 presence of endogenous effect through variation in the interval of observation.

In our model, when there are no endogenous effects, the function $p(\cdot)$ can be seen to be continuous and such that $p'(\cdot) > 0$. Also notice that $p(0) = 0$ 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 $I_n = 2$ it can also be seen that $\text{var}(\mathbf{y}_{n,\Delta_N}) = p(\Delta_N)(1-p(\Delta_N))$. 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, one is then invited to consider $\overline{\mathbf{y}}_{N,\Delta_N} = N^{-1}\sum_{n=1}^{N} \mathbf{y}_{n,\Delta_N}$. The following result is then established:

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 hypothesis that there are no endogenous effects ($p(0) = 0$),

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

where

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

and

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

whereas if there are endogenous effects ($p(0) > 0$)*,*

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

Proof. See Appendix.

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.[16] 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.

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 presumably observes the equilibrium exit strategies $(\tau_1, \ldots, \tau_I)$ for a certain number of realizations of the game. What parameters of the model can be retrieved given data on the situation under analysis? Could two different parameter vectors generate the same distribution for the data? 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 3 (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.* $\Rightarrow \psi = \hat{\psi}$)*.*

[16] 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 as well as in the empirical game estimation literature. Our empirical application employs data at a daily frequency, which is a 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[17]. 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. 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. Let $g(t; \psi, \mathbf{w})$ denote the probability density function for the first desertion time under the parameters $\psi = (\mathbf{x}, \beta, \sigma, \rho, \gamma, C)$ 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, which is a measure of how far apart two probability distributions are[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,*

$$\partial_{w_{il}} \int \log \Big[\frac{g(t; \psi, \mathbf{w})}{g(t; \hat{\psi}, \mathbf{w})}\Big] g(t; \psi, \mathbf{w}) dt \neq 0 \tag{6}$$

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

Proof. See Appendix.

[17]The Inverse Gaussian is the distribution of the hitting time of a Brownian motion on a given barrier $\log z$. In our case, the initial position is $\log x_i$; the drift coefficient, $\alpha^i - \sigma^2/2$ and the diffusion coefficient, σ. If $\alpha^i - \sigma^2/2$ the barrier is reached in finite time with probability one. Otherwise, with a certain probability the barrier is never reached and, conditional on hitting the threshold, the distribution of the random time is an Inverse Gaussian. Whitmore [48] names this last case a *defective* Inverse Gaussian. Chhikara and Folks [8] provide an extensive characterization of the Inverse Gaussian distribution.

[18]Another potential avenue for identification would be through the results presented in McManus [32]. 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.

In order to check condition (6) one should obtain the density g[19] . One possible route is to use the close association between the theory of stochastic processes and the study of differential equations[20]. We assume that the equilibrium played is the one selected in the previous section. Because the equilibrium strategies can then be expressed through hitting times to certain sets, it is possible to characterize the probability density of interest through associated partial differential equations. As in the previous section, let $z_k(\alpha^i, \sigma, \gamma, C, \Delta\alpha, t)$ be the optimal threshold levels defined in Proposition 2. Here, $G(t, \mathbf{x})$ is the probability that the players will abandon the activity *after* time t when the vector of initial conditions is given by $\mathbf{x}$. The density $g(\cdot)$ can then be obtained as $-dG(\cdot)/dt$. The following result then holds[21, 22]:

Proposition 6 *Let $G(t, \mathbf{x}) = \mathbb{P}[\tau_i^* > t, i \in I | \mathbf{x}_0 = \mathbf{x}]$. Then G is the unique solution to*

$$\partial G/\partial t = [\mathcal{A}((\alpha^i)_{i\in I}, \rho, \sigma) + \mathcal{L}_1((\alpha^i)_{i\in I}, \sigma, \gamma, C, \Delta\alpha, t)]G \text{ in } \mathcal{C}_{t=0}, t > 0$$
$$G(0, \mathbf{x}) = \mathbb{P}(\tau_i^* < \infty, i \in I), \mathbf{x} \in \mathcal{C}_{t=0}$$
$$G(t, \mathbf{x}) = 0, \mathbf{x} \in \partial\mathcal{S}_{t=0} \text{ and } t \geq 0$$

where $\mathcal{S}_{t=0} = \{\mathbf{x} \in \mathbb{R}_+^I : \exists i \text{ such that } x^i \geq z_1(\alpha^i, \sigma, \gamma, C, \Delta\alpha, t = 0)\}, \mathcal{C}_{t=0} = \mathcal{S}_{t=0}^c$,

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

which is the infinitesimal generator for the I-dimensional diffusion representing the latent utility vector process with killing time at $\tau_{\mathcal{S}} : \{\mathbf{x}_t : t \leq \tau_{\mathcal{S}}\}$ and

$$\mathcal{L}_1((\alpha^i)_{i\in I}, \sigma, \gamma, C, \Delta\alpha, t)f = -\sum_{i\in I} \frac{dz_1}{dt}(\alpha^i, \sigma, \gamma, C, \Delta\alpha, t)\frac{\partial f}{\partial x_i}.$$

Proof. See Appendix.

Under certain conditions, namely $\alpha^i - \sigma^2/2 > 0$ for some i, $\mathbb{P}(\tau_i^* < \infty, i \in I) = 1$. If not, it can be obtained from another partial differential equation (a proposition similar to the one just displayed can be stated for this case) or directly estimated from the data.

[19] When $\Delta\alpha = \rho = 0$, the Kullback-Leibler information can be obtained in closed form from the Inverse Gaussian probability density function.

[20] For an introduction to the interplay between diffusion processes and partial differential equations, see Karatzas and Shreve [25].

[21] In certain special cases this solution may be available analytically. This is the case when there are only two players or when there are no interaction effects.

[22] We conjecture that, in the presence of discontinuities, a similar result may be attained relying on partial differential equations for the characterization of equilibrium exit distributions.

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[23], 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 desertion decision as an isolated one, historical studies and anecdotal evidence point to the opposite. In the Confederate South, for instance, and especially toward the end of the war, mass desertion tended to be more prevalent. In this regard, Bearman [4] asserts that "*[d]esertion rates were highest in companies that evidenced a high degree of local homogeneity — company solidarity thus bred rather than reduced desertion rates. There is no support for any of the historical models of desertion that search for individual-level determinants, such as social class, occupation, status, family structure, age, or time of enlistment".* Even if one is not as skeptical as this author about individual and other contextual determinants of desertion, there seems to be some evidence in favor of endogenous effects. Evidence of simultaneous desertion (on both sides) is pervasive in Lonn's [28] volume: "*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.*" (p.152-3) The author goes on to point 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 direct impact on the behavior of others in his company. If no one deserts, the social sanctions attached to exit tend to be high; whereas if there is mass exit, such sanctions tend to be minimized as well as the effectiveness of the military company, to decrease. Furthermore, such decision entailed costs — the probability of being caught and facing the military court.[24] 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

[23] The Civil War lasted four years: from the firing at Fort Sumter on April 14, 1861 until Lee's surrender in Appomattox on April 1865. Its human cost was tremendous. It is estimated that 620,000 soldiers died during the conflict (360,000 Union and 260,000 Confederate).

[24] 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 by the President or (later during the war) the commanding general. According to Costa and Kahn [9], out of estimated 200,000 deserters, 80,000 were caught, of which only 147 were executed. Specially in the early years, punishments were notoriously mild, consisting of dismissal with loss of pay and towards the end, imprisonment for the duration of the war.

a company since its inception and hence there was very little flow of soldiers into or out of the unit.

Whereas standard statistical duration models could be employed on the time until desertion to identify the existence of duration dependence among agents — as indeed is done in Costa and Kahn [9] and Sirakaya [45] and suggested in Brock and Durlauf [6] — it is still unclear whether such effects are obtained from endogenous influences or correlated unobservables. In contrast, our model clearly separates both channels and lays out the circumstances under which each of these sources is identifiable.

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 (absence without leave), desertion and furlough.[25] 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 [28] (see Chapter IX), desertion was markedly higher among foreigners, substitutes and "bounty-jumpers". Substitutes and "bounty-jumpers" appeared as the government started inducing enlistment through enrollment bounties — which created the figure of the "bounty-jumper" who would enlist, collect the reward and desert just to

[25]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 towards the end of the war, especially after the institution of the office of provost marshall general in September, 1862 (see Lonn [28]).

repeat the scheme in another state or county — and the possibility for draftees to hire substitutes to replace them. In the data it is possible to identify foreigners and substitutes. In order to assess the effect of "bounty-jumpers" on desertion we try the bounty amount paid to each recruit as a proxy variable.[26] 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[27] and time to desertion is measured from the earliest muster-in date for recruits in a given company.[28]

One of the implications of our model is that company size will affect the equilibrium exit strategies only in the presence of endogenous effects. Table 1 presents evidence of this phenomenon. The regressions investigate the effect of certain variables on the mean (log of the) time to desertion at the individual level[29] Company size is a significant and robust determinant for the timing of desertion. In addition to the displayed regressions, 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 represent an omitted factor other than the numbers in the group.

TABLE 1 HERE

In order to further investigate the presence of endogenous effects in our data, we compute the statistics in Theorem 2 for various discretization levels. All of them yield results that reject the null hypothesis of no endogenous effects, as displayed on 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.

TABLE 2 HERE

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

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

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

[28]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.

[29]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.

5.2. Estimation

In this subsection we use a simulated minimum distance estimator to obtain the relevant parameters in the model proposed in Section 3 (see, for instance, Gouriéroux and Monfort [19]). We normalize the discount rate ($\gamma = 5\%$ per year)[30], the exit cost ($C = 1$) and the initial condition ($x_0 = 0.1$). 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 $R = 1$. In order to simulate the phenomenon we have to discretize the sample paths. The discretization referred to below is the simulated paths discretization. Consistency and asymptotic normality are a straightforward application of the results in Pakes and Pollard [36]:

Proposition 7 *Assume*

1. ***(Identification)*** $\inf_{\|\psi-\psi_0\|>\delta} \|G(\psi)\| = \inf_{\|\psi-\psi_0\|>\delta} \|\mathbb{E}(m(\tau(\psi_0))|\tau(\psi_0) < T) - \mathbb{E}(m(\tau(\psi))|\tau(\psi_0) < T)\| > 0, \forall\delta > 0$*;*
2. ψ_0 *is in the interior of a compact parameter set;*
3. $\Gamma = -\mathbb{E}\Big[\frac{d}{d\psi}(NR)^{-1}\sum_{r=1}^{NR} m(\tau_r(\psi_0))\Big]$ *is full-rank;*
4. $\tau_n, \tau_r < T$*, for some* $T < \infty$*;*
5. $m(\cdot)$ *is* C^1 *with bounded derivatives on* $(0, T)$*;*
6. *The discretization error for the simulated paths is* $o(\sqrt{N})$*,*

then the simulation estimators are weakly consistent and asymptotically normal with distribution:

$$\sqrt{N}(\hat{\psi} - \psi_0) \xrightarrow{\mathcal{L}} \mathcal{N}(0, (\Gamma'\Gamma)^{-1}\Gamma' V \Gamma(\Gamma'\Gamma)^{-1})$$

where

$$\Gamma = -\mathbb{E}\Big[\frac{d}{d\psi}(NR)^{-1}\sum_{r=1}^{NR} m(\tau_r(\psi_0))\Big] \quad \text{and} \quad V = (1 + 1/R)var(m(\tau_i))$$

[30] A discount rate of 5% *per year* was chosen. For comparison, commercial paper rates in United States during the was fluctuated between 4% and 8% (NBER Macrohistory Database).

Proof. See Appendix.

Notice that Condition 2 requires that $\Delta\alpha > 0$. We assume that to be reasonable given the test statistics obtained on Table 2. The last condition basically allows us to ignore the discretization error in the simulation as the sample size increase. This will basically depend on the discretization scheme used. Under the Euler scheme this requires that the discretization grids be $o(\sqrt{N})$ (see Glasserman [18]). The identification condition will be satisfied if the sufficient conditions in Theorem 3 are shown to hold for every $\psi \neq \psi_0$.[31]

An important simplification is that we assume the thresholds to be constant: $z_k = z(\alpha - \Delta\alpha(k-1)/(I-1), \sigma, \gamma, C)$.[32] The moments matched were: mean, harmonic mean, average number of desertions in each desertion episode and percentage of soldiers leaving before two years[33]. The choice of moments stems basically from the fact that, in the absence of endogenous or correlated effects, the mean and harmonic mean are ML estimators and sufficient statistics for the Inverse Gaussian distribution. The following table displays the results and normalizations used in the estimation:

TABLE 3 HERE

As indicated, the parameters are precisely estimated. The results indicate a substantial endogenous effects. The following exercise illustrates this point: if there are no endogenous effects, desertion times would be distributed according to a (defective) Inverse Gaussian distribution. 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. 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% — an increase by a factor of more than two.[34]

6. Conclusion

The problem studied here is of great importance in many settings. Social welfare program participation, bank runs, migration, marriage and divorce decisions are only a few of the

[31] Notice though that the conditions in Theorem 3 are sufficient but not necessary.

[32] Similar approximations can be found on the treatment of finite horizon options and seem to work satisfactorily. Examples are Huang, Subrahmanyam and Yu [23] and Ju [24].

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

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

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 [31]), the separation is not clear in other approaches that deal with similar situations (as in Brock and Durlauf [6]).

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 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.

Appendix: Proofs

Sketch of Proof for 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 [25], 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 Protter [39], Theorem V.7)

Proof of Theorem 1

Consider a player $i \in I$. Let the stopping strategies for $I - \{i\}$ be given by the following profile of stopping times $\tau_{-i} = (\tau_s)_{s \in I - \{i\}}$. Given Assumption 3, according to Theorem 4 in Fakeev [12], there exists a solution for the optimal stopping time. Let the individual i's best response function $b_i(\cdot)$ map a stopping time profile τ_{-i} onto one such optimal stopping solution. Given this, consider $b(\cdot)$ defined as the following mapping $\tau = (\tau_s)_{s \in I} \mapsto b(\tau) = (b_i(\tau_{-i}))_{i \in I}$. A Nash Equilibrium is then simply a fixed point for the mapping $b(\cdot)$. In order to establish the existence of such a result we use the Knaster-Tarski Fixed Point Theorem, reproduced below from Aliprantis and Border [2], p.6:

> **Knaster-Tarski Fixed Point Theorem**: Let $(X, \geq)$ be a partially ordered set with the property that every chain in X has a supremum. Let $f : X \to X$ be increasing, and assume that there exists some a in X such that $a \leq f(a)$. Then the set of fixed points of f is nonempty and has a maximal fixed point.

In the following discussion we consider the set of stopping time profiles and identify two stopping times that are $\mathbb{P}$-almost everywhere identical. We proceed by steps:

<u>Step 1</u>: (Partial order) The set of stopping times endowed with the relation $\geq$ defined as: $\tau \geq \upsilon$ if and only if $\mathbb{P}(\tau(\omega) \geq \gamma(\omega)) = 1$ is partially ordered. In other words, $\geq$ is reflexive, transitive and anti-symmetric.

<u>Step 2</u>: (Every chain has a supremum) Given a set of stopping times T, we should be able to find a stopping time $\overline{\tau}$ such that 1. $\overline{\tau} \geq \tau, \forall \tau \in T, \mathbb{P}$-a.s. and 2. if $\upsilon \geq \tau, \mathbb{P}$-a.s., $\tau \in T$ then $\upsilon \geq \overline{\tau}, \mathbb{P}$-a.s.. If T is countable $\sup_{\tau \in T} \tau$ is a stopping time and satisfies conditions 1 and 2 (see Karatzas and Shreve, Lemma 1.2.11). If not, first notice that, since the only

structure that matters for this property is the ordering in $\overline{\mathbb{R}}_+$, we can always assume that the stopping times take values on $[0,1]$ (otherwise, pick an increasing mapping from $\overline{\mathbb{R}}_+$ onto $[0,1]$). Let $\mathcal{C}$ be the collection of all countable subsets $C \subset T$. For each such C, define:

$$l_C = \sup_{\tau \in C} \tau \text{ and } v = \sup_{C \in \mathcal{C}} \mathbb{E}(l_C) < \infty$$

By the previous reasoning, l_C is a stopping time. Then, there is a sequence $\{C_n\}_n \subset \mathcal{C}$ such that $v = \lim_{n\to\infty} \mathbb{E}(l_{C_n})$. Now define $\overline{C} = \cup_{n=1}^{\infty} C_n \in \mathcal{C}$. To show that $l_{\overline{C}}$ satisfies condition 1., first notice that $\overline{C} \in \mathcal{C}, v \geq \mathbb{E}(l_{\overline{C}})$. On the other hand, since $C_n \subset \overline{C}$, $\mathbb{E}(l_{\overline{C}}) \geq \mathbb{E}(l_{C_n}) \to_n v$. These two imply that $v = \mathbb{E}(l_{\overline{C}})$.

For an arbitrary $\tau \in T$, set $\overline{C}_\tau = \{\tau\} \cup \overline{C} \in \mathcal{C}$. Now, $l_{\overline{C}_\tau} \geq l_{\overline{C}}$. This renders $v \geq \mathbb{E}(l_{\overline{C}_\tau}) \geq \mathbb{E}(l_{\overline{C}}) = v \Rightarrow \mathbb{E}(l_{\overline{C}_\tau} - l_{\overline{C}}) = 0 \Rightarrow l_{\overline{C}_\tau} = l_{\overline{C}}, \mathbb{P}$-a.s. This and $l_{\overline{C}_\tau} \geq \tau, \mathbb{P}$-a.s. in turn imply that $l_{\overline{C}} \geq \tau, \mathbb{P}$-a.s.

To see that 2. is satisfied, notice that, if $v \geq \tau, \forall \tau \in T$, in particular, $v \geq \tau, \forall \tau \in \overline{C}$. This implies that $v \geq \sup_{\tau \in \overline{C}} \tau = l_{\overline{C}}$.

<u>Step 3</u>: ($\exists a$ such that $a \leq f(a)$) Pick a as the profile of stopping times that are identically zero.

<u>Step 4</u>: ($b(\cdot)$ is increasing) This is the case if each individual best response function $b_i(\cdot)$ is increasing. By the version of Itô's Lemma for twice differentiable functions (see Revuz and Yor [40], p.224, remark 3), and the fact that $u_i(x,t) = e^{-\gamma_i t} g_i(x)$ is twice differentiable (since $g_i(\cdot)$ is twice differentiable), $e^{-\gamma_i t} g_i(x)$ obeys the following stochastic differential equation (given a profile of stopping times τ_{-i}):

$$d[e^{-\gamma_i s} g_i(x_s^i)] = \underbrace{e^{-\gamma_i t}[g_i'(x_t^i)\alpha^i(x_t^i,\theta_t,t) + \frac{1}{2}\sigma^{i2}(x_t^i,\theta_t,t)g_i''(x_t^i) - \gamma_i g_i(x_t^i)]}_{\equiv \mu^i(x^i,\theta_t^i,t)} dt + \\ + \underbrace{e^{-\gamma_i t} g_i'(x_t^i)\sigma^i(x_t^i,\theta_t,t)}_{\equiv \beta^i(x_t^i,\theta_t^i,t)} dW_t^i$$

where the $\mu(\cdot,\cdot,\cdot)$ and $\beta(\cdot,\cdot,\cdot)$ denote the drift and dispersion coefficients of $e^{-\gamma_i t} g_i(x_t^i)$. If $g_i(\cdot)$ is increasing and convex and if $\alpha_i(\cdot,\cdot,\cdot)$ and $\sigma_i(\cdot,\cdot,\cdot)$ are decreasing in θ, the above drift is decreasing in θ.

Now consider a profile of stopping times τ_{-i} and υ_{-i} such that τ_{-i} dominates $\upsilon_{-i}, \mathbb{P}$- a.s. Moving from one profile to another will impact θ and this will have effects on both the drift and the dispersion coefficients of $e^{-\gamma_i t} g_i(x_t^i)$.

The effect on the dispersion coefficient does not affect the objective function of an individual agent. This obtains from the fact that $g'(\cdot)$ is bounded and the Bound on Volatility Assumption. These assumptions deliver that, for each $t < \infty$:

$$\mathbb{E}[\int_0^t (e^{-\gamma s} g'(x_s)\sigma(x_s, \theta_s, s))^2 ds] < K\mathbb{E}[\int_0^t (e^{-\gamma s}\sigma(x_s, \theta_s, s))^2 ds] < \infty$$

for some $K \in \mathbb{R}$. This in turn implies that $z_t = \int_0^t (e^{-\gamma s} g'(x_s)\sigma(x_s, \theta_s, s))dW_s^i$ is a martingale (see Karatzas and Shreve [25], p.139) and by the Optional Sampling Theorem, $\mathbb{E}[\int_0^\tau (e^{-\gamma s} g'(x_s)\sigma(x_s, \theta_s, s))dW_s^i] = 0, \forall \tau$ where τ is an $(\mathcal{F}_t)$-stopping time (see Karatzas and Shreve [25], p.19).

Given τ_{-i} and υ_{-i}, we know that $\theta_t^{i,\tau} \leq \theta_t^{i,\upsilon}, \mathbb{P}$-almost surely, $\forall t$ (where $\theta_t^{i,\tau}$ and $\theta_t^{i,\upsilon}$ aggregate the stopping decisions for the profiles τ_{-i} and υ_{-i}) we will have $\mu(x, \theta_t^{i,\upsilon}, t) \leq \mu(x, \theta_t^{i,\tau}, t), \mathbb{P}$-almost surely, $\forall x, t$. Letting $y_t^{i,\tau}$ be the process given by

$$dy_t^{i,\tau} = \mu^i(x_t^i, \theta_t^{i,\tau}, t)dt + \beta(x_t^i, \theta_t^{i,\tau}, t)dW_t^i$$

and $y_t^{i,\upsilon}$ be the process given by

$$dy_t^{i,\upsilon} = \mu^i(x_t^i, \theta_t^{i,\upsilon}, t)dt + \beta(x_t^i, \theta_t^{i,\tau}, t)dW_t^i$$

using a slight variation of Proposition 5.2.18 in Karatzas and Shreve [25], we get:

$$\mathbb{P}[y_t^{i,\tau} \geq y_t^{i,\upsilon}, \forall 0 \leq t < \infty] = 1$$

Again, a slight variation of the proof of this proposition can be repeated using this fact and focusing on $y_t^{i,\tau} - y_s^{i,\tau} - (y_t^{i,\upsilon} - y_s^{i,\upsilon}), t \geq s$ instead of simply $y_t^{i,\tau} - y_t^{i,\upsilon}$. This is enough to achieve the following result:

$$\mathbb{P}[(y_t^{i,\tau} - y_s^{i,\tau}) - (y_t^{i,\upsilon} - y_s^{i,\upsilon}) \geq 0, \forall 0 \leq s \leq t < \infty] = 1$$

This suffices to show that it is not profitable for agent i to stop earlier when the profile is τ_{-i} than when the profile is υ_{-i}. Suppose not. Then, let $A = \{b_i(\tau_{-i}) < b_i(\upsilon_{-i})\}$. According to Lemma 1.2.16 in Karatzas and Shreve [25], $A \in \mathcal{F}_{b_i(\tau_{-i})} \cap \mathcal{F}_{b_i(\upsilon_{-i})}$. By the above result we can then see that $\mathbb{E}\{\mathbb{I}_A[y^{i,\tau}_{b_i(\upsilon_{-i})} - y^{i,\tau}_{b_i(\tau_{-i})}]\} \geq \mathbb{E}\{\mathbb{I}_A[y^{i,\upsilon}_{b_i(\upsilon_{-i})} - y^{i,\upsilon}_{b_i(\tau_{-i})}]\}$. The RHS expression in this inequality is positive because $A \in \mathcal{F}_{b_i(\tau_{-i})} \cap \mathcal{F}_{b_i(\upsilon_{-i})} = \mathcal{F}_{b_i(\tau_{-i}) \wedge b_i(\upsilon_{-i})}$ which implies that the agent would do better by picking $b_i(\tau_{-i}) \wedge b_i(\upsilon_{-i})$ if the RHS were negative. But this would contradict the fact that $b_i(\upsilon_{-i})$ is a best response. So, if $A \neq \emptyset$, delaying the response by choosing $b_i(\upsilon_{-i}) \vee b_i(\tau_{-i})$ would improve the agent's payoff given that the remaining agents are playing τ_{-i}. ■

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 $k \in \{0, 1, \dots, n\}$ 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 $0 = \nu_0 \leq \nu_1 \leq \cdots \leq \nu_k$, $\lambda_k(t;\omega) = 0$ if $t < \nu_{k-1}(\omega)$. The value function for this problem is then given by:

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

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

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

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:

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

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 — 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[35], let $J : \mathbb{R}_{++} \times \{1, \dots, n\} \times \mathbb{R}_+ \to \mathbb{R}$ 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 [41], p.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

[35] The reasoning is in the spirit of similar arguments in Kobila [27] and Scheinkman and Zariphopoulou [42].

initialized at the level x and regime k. Since the drift in successive states are strictly smaller, a comparison result such as the one in Karatzas and Shreve [25], Proposition V.2.18, or Protter [39], Theorem V.54, can be established to show that:

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

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

<u>Step 2</u>: (Stopping times are an increasing sequence) Notice that, by definition, $\tau_0 \leq \tau_1 \leq \dots \leq \tau_I$ 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 $\Delta\alpha^i/(I-1)$.

<u>Step 4</u>: (τ^*_i is optimal) Apply Proposition 1.

This establishes that the equilibrium can be represented through the hitting times. ■

Proof of Proposition 3

<u>Step 1</u>: (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 $\mathcal{S} = \{(\mathbf{x}, t) \in \mathbb{R}^I_{++} \times \mathbb{R}_+ : \exists i \text{ such that } x^i \geq z^i_1 = z(\alpha^i, \sigma^i, C^i, \gamma^i, \Delta\alpha^i, t)\}$ 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 [38] 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, $z_1^i(t) > z_2^i(t), \forall t \quad i \in I$. There will be simultaneous exit whenever $z_1^i \geq x_{\tau_{\mathcal{S}}}^i \geq z_2^i$, for some $i \in I$. This has positive probability as long as $z_1^i(\tau_{\mathcal{S}}) > z_2^i(\tau_{\mathcal{S}}), i \in I$. In order to see this, first notice that the latent utilities process can be represented as the following diffusion process with killing time at $\tau_{\mathcal{S}}$:

$$dx_t^i = \alpha^i x_t^i dt + \sum_{j \in I} \tilde{\sigma}_{ij} dB_t^j, \quad i = 1, \dots, I$$

where $\mathbf{B}_t$ is an I-dimensional Brownian motion (with independent components) and $\tilde{\sigma}_{I \times I} = [\tilde{\sigma}_{ij}]$. Let $\partial \mathcal{S}_H = \{(\mathbf{x}, t) \in \partial \mathcal{S} : z_1^i(t) \geq x^i \geq z_2^i\}$. By Corollary II.2.11.2 in Gihman and Skorohod [17] (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 [38] or Section 6.4 in Evans [11]), 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 [38], 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 $\tau_i^* = \inf\{t \in \mathbb{R}_+ : x_t^i > z^i = z(\alpha^i, \sigma^i, C^i, \gamma^i)\}$. 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 [38]). 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 [48]). 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 (see Pagan and Ullah [37], p.358) 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 $(\alpha_1, \alpha_2) \in \mathbb{R}^2$,

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

Let

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

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

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

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

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

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

For C, observe that

$$
\begin{aligned}
C & \leq \sum_{i=1,2,3} [\mathbb{E}[|\alpha_i \mathbf{y}_{n,\Delta_{N,i}}|]^{2+\delta} + (\alpha_i p(\Delta_{N,i}))^{2+\delta}] \leq \\
& \leq \sum_{i=1,2,3} [\alpha_i^{2+\delta} + (\alpha_i p(\Delta_{N,i}))^{2+\delta}] \longrightarrow \\
& \longrightarrow (1 + p(0))^{2+\delta} \sum_{i=1,2,3} \alpha_i^{2+\delta}.
\end{aligned}
$$

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

$$\zeta_N(\Delta_{N,1}, \Delta_{N,2}, \Delta_{N,3}) \to \zeta_N(0, 0, 0) > 0.$$

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

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

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

$$
\begin{aligned}
\Delta^{-1} \zeta_N(\Delta, 0, 0) &= \Delta^{-1} \alpha_1^2 \mathrm{var}(\mathbf{y}_{n,\Delta}) = \\
&= \Delta^{-1} \alpha_1^2 \mathrm{var} \left(\binom{I_n}{2}^{-1} \sum_{\{i,j\} \in \pi_n} \mathbb{I}_{\{\tau_\Delta^i = \tau_\Delta^j\}} \right) = \\
&= \Delta^{-1} \alpha_1^2 \left[\binom{I_n}{2}^{-1} \mathrm{var} \left(\mathbb{I}_{\{\tau_\Delta^i = \tau_\Delta^j\}} \right) + \right. \\
&+ \left. 2 \binom{I_n}{2}^{-2} \sum_{\{i,j\},\{k,l\} \in \pi_n} \mathrm{cov} \left(\mathbb{I}_{\{\tau_\Delta^i = \tau_\Delta^j\}}, \mathbb{I}_{\{\tau_\Delta^k = \tau_\Delta^l\}} \right) \right] \geq \\
&\geq \Delta^{-1} \alpha_1^2 \binom{I_n}{2}^{-1} \left[p(\Delta)(1 - p(\Delta)) - 2p(\Delta)^2 \right] \overset{\Delta \to 0}{\longrightarrow} \\
&\overset{\Delta \to 0}{\longrightarrow} \alpha_1^2 \binom{I_n}{2}^{-1} p'_+(0) > 0
\end{aligned}
$$

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

$$N^\varepsilon \zeta_N(\Delta_{N,1}, \Delta_{N,2}, \Delta_{N,3}) = N^\varepsilon \times \left(\partial\zeta_N(\widehat{\Delta}_{N,1}, \widehat{\Delta}_{N,2}, \widehat{\Delta}_{N,3})' \begin{bmatrix} \Delta_{N,1} \\ \Delta_{N,2} \\ \Delta_{N,3} \end{bmatrix}\right) \xrightarrow{N}$$

$$\xrightarrow{N} \sum_{i=1,2,3} k_i \partial_i^+ \zeta_N(0,0,0) > 0$$

where $k_i, i = 1,2,3$ 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

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

Since (α_1, α_2) is arbitrary, the Cramér-Wold device dictates that (see van der Vaart [47])

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

STEP 2: (Delta Method) By the uniform delta method (see van der Vaart [47], Theorem 3.8) 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

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

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

$$\begin{aligned}
\frac{p(\Delta_{N,k})}{p(\Delta_{N,1})} - \frac{a_k}{a_1} &= \frac{\Delta_{N,k}p'(0) + \frac{1}{2}\Delta_{N,k}^2 p''(0) + O(\Delta_{N,k}^3)}{\Delta_{N,1}p'(0) + \frac{1}{2}\Delta_{N,1}^2 p''(0) + O(\Delta_{N,1}^3)} - \frac{a_k}{a_1} = \\
&= \frac{a_k + ca_k^2 N^{-\epsilon} + O(N^{-2\epsilon})}{a_1 + ca_1^2 N^{-\epsilon} + O(N^{-2\epsilon})} - \frac{a_k}{a_1} = \frac{ca_k a_1(a_k - a_1)N^{-\epsilon} + O(N^{-2\epsilon})}{a_1[a_1 + ca_1^2 N^{-\epsilon} + O(N^{-2\epsilon})]} = \\
&= \frac{ca_k a_1(a_k - a_1)N^{-\epsilon}}{a_1^2[1 + ca_1 N^{-\epsilon} + O(N^{-2\epsilon})]} + O(N^{-2\epsilon}) \overset{\text{(by Taylor's Expansion)}}{=} \\
&= \frac{ca_k a_1(a_k - a_1)N^{-\epsilon}}{a_1^2} \times [1 - ca_1 N^{-\epsilon} + O(N^{-2\epsilon})] + O(N^{-2\epsilon}) = \\
&= \frac{ca_k a_1(a_k - a_1)N^{-\epsilon}}{a_1^2} + O(N^{-2\epsilon})
\end{aligned}$$

so that

$$\frac{p(\Delta_{N,k})}{p(\Delta_{N,1})} = \frac{a_k}{a_1} + \frac{ca_k a_1(a_k - a_1)N^{-\epsilon}}{a_1^2} + O(N^{-2\epsilon}).$$

This delivers

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

Noticing that

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

gives us that

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

as long as $\epsilon > 1/3$.

If $p(0) > 0$, it is possible to see that the statistic diverges using similar arguments. ■

For the next theorem we will make use of the following result (Theorem 1 in Araújo and Mas-Colell [3]), which we cite as a lemma.

Lemma 1 (Theorem 1 in Araújo and Mas-Colell [3]) *Let Ψ be a topological space, $E \subset \mathbb{R}^n, 1 \leq n \leq \infty$ and ν denote a Borel probability measure on $\mathbb{R}^n$. Assume the following:*

1. *$(\Psi \times \Psi)\backslash\Delta$ is a Lindelöf space (i.e. any open cover has a countable subcover), where $\Delta = \{(x, y) \in \Psi \times \Psi : x = y\}$.*
2. *$F : \Psi \times E \to \mathbb{R}$ is a continuous function.*
3. *$\forall i, \psi \in \Psi$ and $a \in E, \partial_{a_i} F(\psi, a)$ exists and depends continuously on ψ and a.*
4. *ν is a product probability measure, each factor being absolutely continuous with respect to the Lebesgue measure.*
5. *(Identification Condition)[36] If $F(\psi, a) = F(\hat{\psi}, a), \psi \neq \hat{\psi}$, then $\partial_{a_i}(F(\psi, a) - F(\hat{\psi}, a)) \neq 0$ for some i.*

Then, for ν-a.e. $a \in E$, the function $F(\cdot, a) : \Psi \to \mathbb{R}$ has at most one maximizer.

Proof of Theorem 3

<u>Step 1</u>: 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 $\hat{\psi}$ is maximizes the expected log-likelihood if and only if $g(t; \hat{\psi}, \mathbf{w}) = g(t; \psi, \mathbf{w})$ (see Schervish [44], Proposition 2.92). In particular, $\hat{\psi} = \psi$ is one such maximizer.

<u>Step 2</u>: 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. Now, for each fixed value in the support of $(w_i^d)_{i \in I}$, notice that:

1. Ψ is trivially Lindelöf since it is compact[37]

[36] This condition is named the Sondermann condition in Araújo and Mas-Collel [3]. We change its denomination to better suit our application.

[37] 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 [2], Theorem 3.1). This implies that it is Lindelöf (see Aliprantis and Border [2], p.45).

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

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

and

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

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 *et al.* [13] 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 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 Identification Condition holds by assumption.

By Lemma 1, 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. ■

Proof of Proposition 6

Notice that (for $t \in [0, \tau_{\mathcal{S}}]$) the vector process with the latent utilities can be represented as the following diffusion process with killing at time $\tau_{\mathcal{S}}$:

$$dx_t^i = \alpha^i x_t^i dt + \sigma x_t^i dW_t^i, \qquad i \in I$$

Let $\mathcal{S} = \mathcal{S}((\alpha^i)_{i\in I}, \sigma, \gamma, C, \Delta\alpha, t) = \{\mathbf{x} \in \mathbb{R}_+^I : \exists i \text{ such that } x^i \geq z_1^i(t) \equiv z_1(\alpha^i, \sigma, \gamma, C, \Delta\alpha, t)\}$ and denote by $\mathcal{A}((\alpha^i)_{i\in I}, \rho, \sigma)$ the infinitesimal generator associated with the above diffusion (where the argument reminds the reader of the dependence of the operator on the parameters). In other words, $\mathcal{A}((\alpha^i)_{i\in I}, \rho, \sigma)$ is the following differential operator:

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

for f in the appropriate domain (see Karatzas and Shreve [25], p.281).

Denote by $\tau_{\mathcal{S}} \equiv \inf\{t : \mathbf{x}_t \in \mathcal{S}\} = \inf\{t : \exists i \text{ such that } x^i \geq z_1^i(t) \equiv z_1(\alpha^i, \sigma, \gamma, C, \Delta\alpha, t)\} = \inf\{t : \exists i \text{ such that } \hat{x}^i \equiv x^i - (z^i(t) - z^i(0)) \geq z^i(0)\}$. Let $G(t, \mathbf{x})$ be the probability that the diffusion will reach $\mathcal{S}$ after t. In other words, $G(t, \mathbf{x}) = \mathbb{P}[\tau_{\mathcal{S}} > t|\mathbf{x}_0 = \mathbf{x}]$ and represents the survival function for the exit time distribution of the first deserter. Following Gardiner [15], Subsection 5.4.2, this probability can be conveniently written as the solution to the following (parabolic) partial differential equation (Kolmogorov backward equation):

$$G_t = [\mathcal{A}((\alpha^i)_{i\in I}, \rho, \sigma) + \mathcal{L}((\alpha^i)_{i\in I}, \rho, \sigma, \Delta\alpha, t)]G$$

in $\mathcal{C}_{t=0}((\alpha^i)_{i\in I}, \sigma, \gamma, C, \Delta\alpha), \quad t > 0$ with the following conditions:

$$\begin{aligned} &G(0, \mathbf{x}) = \mathbb{P}[\tau_{\mathcal{S}} < \infty|\mathbf{x}_0 = \mathbf{x}], \quad \mathbf{x} \in \mathcal{C}_{t=0}((\alpha^i)_{i\in I}, \sigma, \gamma, C, \Delta\alpha) \\ &G(t, \mathbf{x}) = 0, \quad \mathbf{x} \in \mathcal{S}_{t=0}((\alpha^i)_{i\in I}, \sigma, \gamma, C, \Delta\alpha) \text{ and } t \geq 0 \end{aligned}$$

where the boundary condition follows since $\partial\mathcal{S}_{t=0}((\alpha^i)_{i\in I}, \sigma, \gamma, C, \Delta\alpha) \subset \mathcal{S}_{t=0}((\alpha^i)_{i\in I}, \sigma, \gamma, C, \Delta\alpha)$ and because 0 is an absorbing boundary for $x^i, i \in I$.

Uniqueness is obtained in Theorem 4, Section 7.1.2 in Evans [11]. ■

Proof of Proposition 7

The proof follows from the conditions in Theorems 3.1 and 3.3 in Pakes and Pollard [36]. It is analogous to Example 4.1 in that paper, in which the authors check for these conditions in the stopping model by Pakes [35].

With exception of condition (ii) in Theorem 3.3, the conditions follow pretty much along the same lines as in that example, so we omit them here. Condition (ii) requires that $G(\cdot)$ be differentiable at ψ_0 with derivative matrix Γ, full rank. We prove here that $G(\cdot)$ is differentiable at ψ_0. That the matrix is of full rank is assumed in the statement of the proposition.

For $\tau < T$, we would like to establish that $G(\psi) = \mathbb{E}^x[m(\tau(\psi))|\tau(\psi) < T]$ is differentiable at

ψ_0. In order to do this we use Proposition 1 in Broadie and Glasserman [5]. The proposition consists in imposing condition (A1-A4) for the Lebesgue Dominated Convergence Theorem so that the expectation of a derivative is equal to the derivative of the expectation.

First, we check for the conditions that focus on the differentiability of each realization of the random variable $m(\tau)$. Letting $B_{t,\text{discr}}(\omega)$ be a discretized draw for the (continuous time) I-dimensional Brownian motion governing the behavior of the state variable processes. For this realization, the stopping time for individual i is given implicitly by $\tau(\omega;\psi)$:

$$\sup_{\tau(\omega;\psi)\geq s}\{\mu(\theta_s(\omega,\psi),\alpha,\Delta\alpha,s)-\frac{\sigma^2}{2}s+\sigma B_{s,\text{discr}}(\omega)\}=\log z(\theta_{\tau(\omega;\psi)}(\omega,\psi),\psi).$$

The implicit function theorem guarantees that $\tau(\omega;\psi)$ is differentiable w.r.t. ψ with probability one. This takes cares of A1.

The assumption that $m(\cdot)$ is differentiable and has bounded derivatives on $(0,T)$ is used to satisfy A2 and A3.

The fact that the parameter space is compact guarantees that the derivative of $\tau(\omega;\psi)$ is bounded by an integrable random variable and thus condition A4 is satisfied.

This delivers existence of the derivative as the discretization window goes to zero. ■

References

[1] Abreu, D. and M. Brunnermeier (2002) "Synchronization Risk and Delayed Arbitrage", *Journal of Financial Economics*, **66**.

[2] Aliprantis, C. and K. Border (1999) *Infinite Dimensional Analysis: A Hitchhiker's Guide*. Springer-Verlag.

[3] Araújo, A. and A. Mas-Colell (1978) "Notes on the Smoothing of Aggregate Demand", *Journal of Mathematical Economics* **5**(2).

[4] Bearman, P. (1991) "Desertion as Localism: Army Unit Solidarity and Group Norm in the U.S. Civil War", *Social Forces*, **70**(2).

[5] Broadie, M. and P. Glasserman (1996) "Estimating Security Price Derivatives Using Simulation", *Management Science*, **42**(2).

[6] Brock, W. and S. Durlauf (2002) "Interactions-Based Models" **in** J. Heckman and E. Leamer (eds.), *Handbook of Econometrics*, v.5, Elsevier Science.

[7] Brunnermeier, M. and J. Morgan (2004) "Clock Games: Theory and Experiment", mimeo., Deparment of Economics, Princeton University.

[8] Chhikara, R. and J.L. Folks (1989) *The Inverse Gaussian Distribution*, Marcel Dekker, Inc.

[9] Costa, D. and M. Kahn (2003) "Cowards and Heroes: Group Loyalty during the American Civil War", *Quarterly Journal of Economics*, **118**(2).

[10] Dixit, A. and R. Pindyck (1994) *Investment Under Uncertainty*, Princeton University Press.

[11] Evans, L. (1998) *Partial Differential Equations*, American Mathematical Society.

[12] Fakeev, A.G. (1970) "On Optimal Stopping of Random Processes with Continuous Time", *Theory Prob. Applications*, **15**(2).

[13] 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**.

[14] Fudenberg, D. and J. Tirole (1991) *Game Theory*, The MIT Press.

[15] GARDINER, C.W. (2002) *Handbook of Stochastic Methods for Physics, Chemistry and the Natural Sciences*. Second Edition. Springer-Verlag.

[16] GLAESER, E. AND J. SCHEINKMAN (2002) "Non-Market Interactions" **in** M. Dewatripont, L.P. Hansen and S. Turnovsky (eds.), *Advances in Economics and Econometrics: Theory and Applications, Eight World Congress*, Cambridge University Press.

[17] GIHMAN, I.I. AND A.V. SKOROHOD (1972) *Stochastic Differential Equations*. Springer-Verlag.

[18] GLASSERMAN, P. (2004) *Monte Carlo Methods in Financial Engineering*. Springer.

[19] GOURIÉROUX, C. AND A. MONFORT (1996) *Simulation-Based Econometric Methods*, Oxford University Press.

[20] GRANOVETTER, M. (1978) "Threshold Models of Collective Behavior", *American Journal of Sociology*, **83**(6).

[21] HONG, H., J. KUBIK AND J. STEIN (2004) "Social Interactions and Stock-Market Participation", *Journal of Finance*, forthcoming.

[22] HOPENHAYN, H. AND F. SQUINTANI (2004) "Preemption Games with Private Information", working paper.

[23] HUANG, J., M. SUBRAHMANYAM AND G. YU (1996) "Pricing and Hedging American Options: A Recursive Integration Method", *The Review of Financial Studies*, **9**(1).

[24] JU, N. (1998) "Pricing an American Option by Approximating Its Early Exercise Boundary as a Multipiece Exponential Function", *The Review of Financial Studies*, **11**(3).

[25] KARATZAS, I. AND S. SHREVE (1991) *Brownian Motion and Stochastic Calculus*. Springer-Verlag.

[26] KARLIN, S. AND H. TAYLOR (1981) *A Second Course in Stochastic Processes*, Academic Press.

[27] KOBILA, T.O. (1993) "A Class of Solvable Stochastic Investment Problems Involving Singular Controls", *Stochastics and Stochastics Reports*, 43.

[28] LONN, E. (1928) *Desertion During the Civil War*, The Century Co.

[29] Lancaster, T. (1972) "A Stochastic Model for the Duration of a Strike" *Journal of the Royal Statistical Society, Series A (General)*, **135**(2)

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

[31] Manski, C. (1993) "Identification of Endogenous Social Effects: The Reflection Problem", *Review of Economic Studies*, **60**(3).

[32] McManus, D. (1992) "How Common Is Identification in Parametric Models?", *Journal of Econometrics*, **53**.

[33] Morris, S. (1995) "Co-operation and Timing", CARESS Working Paper No.95-05.

[34] 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.

[35] Pakes, A. (1986) "Patents as Options: Some Estimates of the Value of Holding European Patent Stocks", *Econometrica*, **54**.

[36] Pakes, A. and D. Pollard (1989) "Simulation and the Asymptotics of Optimization Estimators", *Econometrica*, **57**(5).

[37] Pagan, A. and A. Ullah (1999) *Nonparametric Econometrics*, Cambridge University Press.

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

[39] Protter, P. (1991) *Stochastic Integration and Differential Equations. A New Approach.*, Springer- Verlag.

[40] Revuz, D. and M. Yor (1999) *Continuous Martingales and Brownian Motion*, Springer-Verlag.

[41] Rogers, L.C.G. and D. Williams (1994) *Diffusions, Markov Processes and Martingales, V.1*, Cambridge U. Press, Cambridge.

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

[43] Schelling, T. (1960) *The Strategy of Conflict*, Harvard University Press.

[44] Schervish, M. (1995) *Theory of Statistics*, Springer.

[45] Sirakaya, S. (2006) "Recidivism and Social Interactions", forthcoming in *Journal of the American Statistical Association*.

[46] 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**.

[47] van der Vaart, A.W. (1998) *Asymptotic Statistics*, Cambridge University Press.

[48] Whitmore, G.A. (1979) "An Inverse Gaussian Model for Labour Turnover", *Journal of the Royal Statistical Society, Series A (General)*, **142**(4)

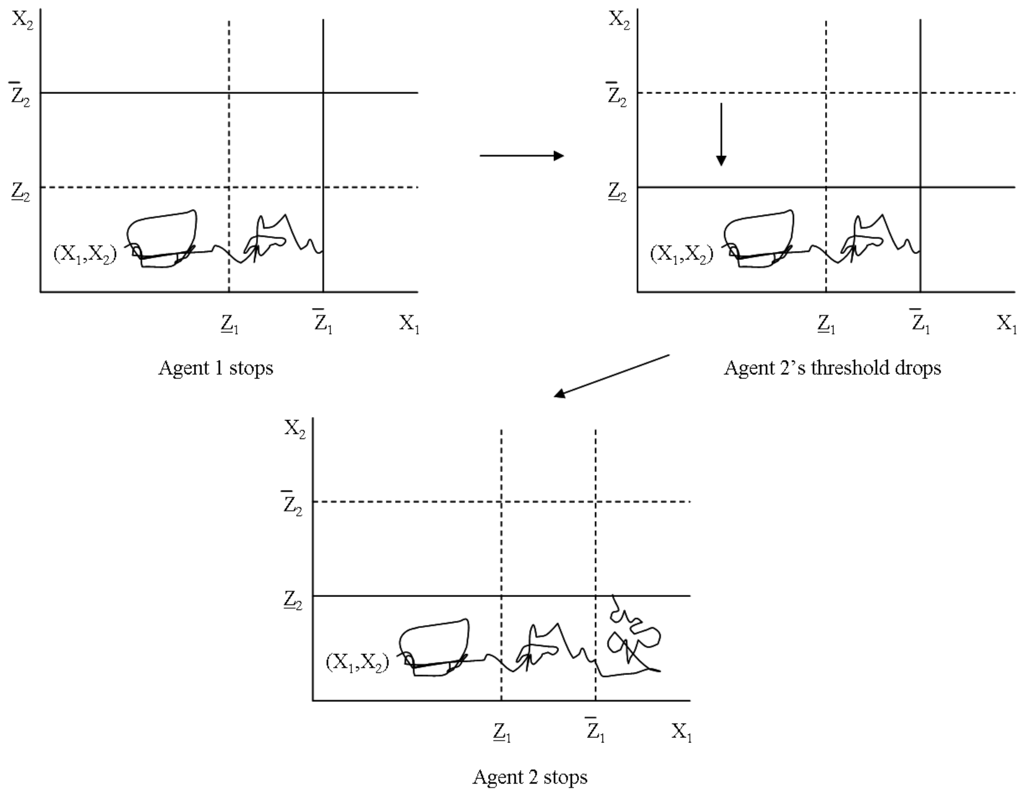

Figure 1: Sequential Stopping

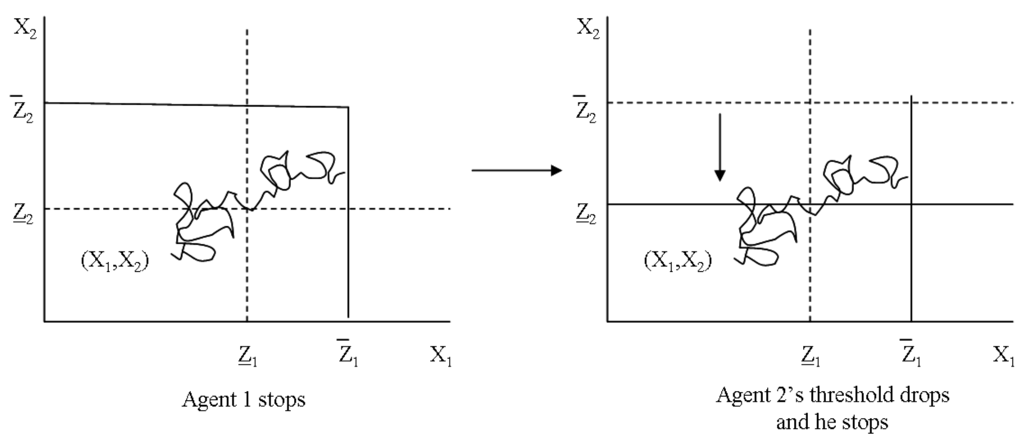

Figure 2: Simultaneous Stopping

Table 1: Individual Regressions

Dependent variable: log(Days Until Desertion)						
	Coef. (Std. Error)	t	P> |t|	Coef. (Std. Error)	t	P> |t|
Company Size	0.0020 (0.0006)	3.43	0.00	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)	2.70	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 = 1. Exit cost = 0.1. Discount rate = 5% *per year*.

Penn Institute for Economic Research

ABOUT PIER

EVENTS

LECTURE SERIES

RESEARCH FELLOWS

VISITORS

RESEARCH PROGRAMS

WORKING PAPER SERIES

3718 Locust Walk, University of Pennsylvania, Philadelphia, PA 19104-6297
Questions or comments may be directed to: pier@econ.upenn.edu