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

PIER Working Paper 09-039

"Independent Durations"
Third Version

by

Bo E. Honore and Aureo de Paula

http://ssrn.com/abstract=XXXXXX

Interdependent Durations *

Bo E. Honoré †
Princeton University

Áureo de Paula ‡
University of Pennsylvania

Abstract

This paper studies the identification of a simultaneous equation model involving duration measures. It proposes a game theoretic model in which durations are determined by strategic agents. In the absence of strategic motives, the model delivers a version of the generalized accelerated failure time model. In its most general form, the system resembles a classical simultaneous equation model in which endogenous variables interact with observable and unobservable exogenous components to characterize an economic environment. In this paper, the endogenous variables are the individually chosen equilibrium durations. Even though a unique solution to the game is not always attainable in this context, the structural elements of the economic system are shown to be semiparametrically identified. We also present a brief discussion of estimation ideas and a set of simulation studies on the model.

JEL Codes: C10, C30, C41.

---

*Versions of this paper at different stages were presented to various audiences. We thank these audiences for their many comments. In particular we thank Herman Bierens, Yi Chen, James Heckman, Wilbert van der Klaauw, Rob Porter, Geert Ridder, Elie Tamer, Michela Tincani, Giorgio Topa and Quang Vuong for their insights. We also thank the editor, Enrique Sentana, and three anonymous referees, whose comments helped us significantly improve the article. Bo Honoré gratefully acknowledges financial support from the National Science Foundation, the Gregory C. Chow Econometric Research Program at Princeton University, and the Danish National Research Foundation (through CAM at the University of Copenhagen).

†Department of Economics, Princeton University, Princeton, NJ 08544. E-mail: `honore@princeton.edu`

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

1 Introduction

This paper investigates the identification of a simultaneous equation model involving durations. We present a simple game theoretic setting in which spells are determined by multiple optimizing agents in a strategic way. As a special case, our proposed structure delivers the familiar proportional hazard model as well as the generalized accelerated failure time model. In a more general setting, the system resembles a classical simultaneous equation model in which endogenous variables interact with each other and with observable and unobservable exogenous components to characterize an economic environment. In our case, the endogenous variables are the individually chosen equilibrium durations. In this context, a unique solution to the game is not always attainable. In spite of that, the structural elements of the economic system are shown to be semiparametrically point identified.

The results presented here have connections to the literatures on simultaneous equations and statistical duration models as well as to the recent research on incomplete econometric models that result from structural (game theoretic) economic models (Berry and Tamer (2006)). The paper also adds to the research on time-varying explanatory variables in duration models. In that literature the time–varying explanatory variable is considered to be "external" (see, for instance, Heckman and Taber (1994) or Hausman and Woutersen (2006)). In an earlier paper, Lancaster (1985) considers a duration model where there is simultaneity with another (non-duration) variable for a single agent. In this paper, we focus on simultaneously determined duration outcomes with more than one agent. More recently, Abbring and van den Berg (2003) consider a model where a duration outcome depends on a time-varying explanatory variable, another duration variable, and endogeneity arises because an unobserved heterogeneity term impacts both of the two durations. One can think of the contribution of this paper as providing an alternative framework that allows for endogeneity.

There are many situations in which two or more durations interact with each other. Park and Smith (2006), for instance, cite circumstances in which late rushes in market entry occur as some pioneer firm creates a market for a new service or good. In our model, the decision by the pioneer is understood as having an impact on the attractiveness of the market to other potential entrants. In another related example, Fudenberg and Tirole (1985) examine technology adoption by a set of agents. In their setting, the adoption time by one agent affects the the other agent's adoption time in a number of ways. Under some circumstances, a "diffusion" equilibrium arises, in which players adopt the new technology sequentially. For other parametric configurations, adoption occurs simultaneously and there are many equilibrium times at which this occurs. Our model allows for similar results where sequential timing arises under some realizations of our game and simultaneous timing occurs as multiple equilibria for other realizations. Peer effects in durations also play a natural role in some empirical examples leading to interdependent durations. In Paula (2009), soldiers in the Union Army during the American Civil War tended to desert in groups. Mass desertion could be thought of as lowering the costs of desertion, directly and indirectly, as well as reducing the combat capabilities of a military company. Another example involves the decision by adolescents to first consume alcohol, drugs or cigarettes or to drop out of high school. In this case, the timing chosen by one individual could have an effect on the decisions of others in a given reference group. Other phenomena that could also be analyzed with our model include the decision to retire among couples, the simultaneous bidding on EBay auctions and the pricing behavior of competing firms.

The examples above typically result in a positive probability of concurrent timing. Let T_i and T_j denote the duration variables for two individuals i and j, and suppose that we are interested in the distribution of T_i conditional on T_j, $\mathbb{P}(T_i \leq t | T_j = t_j)$ (and vice versa). From a statistical viewpoint, one might specify a reduced–form model for the conditional distributions as

$$\mathbb{P}(T_i \leq t | T_j = t_j) = \begin{cases} F_i(t)(1 - \pi_i(t_j)) & \text{if } t < t_j \\ F_i(t)(1 - \pi_i(t_j)) + \pi_i(t_j) & \text{otherwise.} \end{cases}$$

where $i \neq j$, $F_i(\cdot)$ is a continuous CDF and $\pi_i(\cdot)$ is between 0 and 1. In other words, conditional on T_j, T_i has a continuous distribution, except that there is a point mass at T_j. One can motivate such a distribution by a model in which three types of events occur. The first two "fatal events" lead to terminations of the spells for individuals 1 and 2, respectively, and the third will lead both spells to terminate. These "shock" models, introduced by Marshall and Olkin (1967), have been used in industrial reliability and biomedical statistical applications (see, for example, Klein, Keiding, and Kamby (1989)). In these models the relationship between the durations is driven by the unobservables, but no direct relationship exists between them. This is similar to the dependence between two dependent variables in a "seemingly unrelated regressions" framework. In economics, it is interesting to consider models in which durations depend on each other in a structural way, allowing for an interpretation of estimated parameters closer to economic theory. This is the aim of our paper. As such, the difference between Marshall and Olkin's model and ours is similar to the difference between seemingly unrelated regressions and structural simultaneous equations models.

To achieve this, we formulate a very simple game theoretic model with complete information where players make decisions about the time at which to switch from one state to another. Our analysis bears some resemblance to previous studies in the empirical games literature, such as Bresnahan and Reiss (1991) and, more recently, Tamer (2003). Bresnahan and Reiss (1991), building on the work in Amemiya (1974) and Heckman (1978), analyze a simultaneous game with a discrete number of possible actions for each agent. A major pitfall in such circumstances is that "when a game has multiple equilibria, there is no longer a unique relation between players' observed strategies and those predicted by the theory" (Bresnahan and Reiss (1991)). When unobserved components have large enough supports, this situation is pervasive for the class of games they analyze. Tamer (2003) characterizes this particular issue as an "incompleteness" in the model and shows that this nuisance does not necessarily preclude point identification of the deep parameters in the model. Our model also possesses multiple equilibria and, like Tamer, we also obtain point identification of the main structural features of the model. This is possible because certain realizations of the stochastic game we analyze deliver unique equilibrium outcomes with sequential timing choices while multiplicity occurs if and only if spells are concurrent. We are then able to obtain point identification using arguments similar to the ones used to obtain identification in mixed proportional hazards models (see, for example, Elbers and Ridder (1982)).

Since the econometrician observes outcomes for two agents, our model is a multiple duration model. The availability of multiple duration observations for a given unit provides leverage both in terms of identification and subsequent estimation (see Honoré (1993), Horowitz and Lee (2004) and Lee (2003)). In the panel duration literature, subsequent spells, such as unemployment durations for workers or time intervals between transactions for assets, are typically observed for a given individual. This allows for the introduction of individual specific effects. In this paper, parallel individual spells are recorded for a given game, and some elements in our analysis can be made game-specific, mimicking the role of individual specific effects in the panel duration literature. [1]

We use a continuous time setting. This is the traditional approach in econometric duration studies and statistical survival analysis. Many game theoretic models of timing are also set in continuous time. The framework can be understood as the limit of a discrete time game. As the frequency of interactions increases, the setting converges to our continuous time framework, which can in turn be seen as an approximation to the discrete time model.

[1] See Hougaard (2000) and Frederiksen, Honoré, and Hu (2007).

The exercise is thus in line with the early theoretical analysis by Simon and Stinchcombe (1989), Bergin and MacLeod (1993) and others and with most of the econometric analysis of duration models (e.g. Elbers and Ridder (1982), Heckman and Singer (1984), Honoré (1990), Hahn (1994), Ridder and Woutersen (2003), Abbring and van den Berg (2003)). See also van den Berg (2001).

The remainder of the paper proceeds as follows. In the next section we present the economic model. Section 3 investigates the identification of the many structural components in the model. The fourth section discusses extensions and alternative models to our main framework. Section 5 briefly discusses estimation strategies and the subsequent section presents simulation exercises to illustrate the consequences of ignoring the endogeneity problem introduced by the interaction or misspecifying the equilibrium selection mechanism. We conclude in the last section.

2 The Economic Model

The economic model consists of a system of two individuals who interact. Information is complete for the individuals. Each individual i chooses how long to take part in a certain activity by selecting a termination time $T_i \in \mathbb{R}_+, i = 1, 2$. Agents start at an activity that provides a utility flow given by the positive random variable $K_i \in \mathbb{R}_+$. At any point in time, an individual can choose to switch to an alternative activity that provides him or her with a flow utility $U(t, \mathbf{x}_i)$ where the vector $\mathbf{x}_i$ denotes a set of covariates.[2] This utility flow is incremented by a factor e^δ when the other agent switches to the alternative activity. We assume that $\delta \geq 0$. Since only the difference in utilities will ultimately matter for the decision, there is no loss in generality in normalizing the utility flow in the initial activity to be a time-invariant random variable.

In order to facilitate the link of our study to the analysis of duration models, we adopt a multiplicative specification for $U(t, \mathbf{x}_i)$ as $Z(t)\varphi(\mathbf{x}_i)$ where $Z : \mathbb{R}_+ \to \mathbb{R}_+$ is a strictly increasing, absolutely continuous function such that $Z(0) = 0$. Assuming an exponential discount rate ρ, individual i's utility for taking part in the initial activity until time t_i given the other agent's timing choice T_j is:

[2]One could in principle allow for ("external") time-varying covariates, but these would have to be fully forecastable by the individuals.

$$\int_0^{t_i} K_i e^{-\rho s} ds + \int_{t_i}^{\infty} Z(s)\varphi(\mathbf{x}_i) e^{\mathbf{1}_{(s \geq T_j)}\delta} e^{-\rho s} ds$$

The first-order condition for maximizing this with respect to t_i is based on:

$$K_i e^{-\rho t_i} - Z(t_i)\varphi(\mathbf{x}_i) e^{\mathbf{1}_{(t_i \geq T_j)}\delta} e^{-\rho t_i} \tag{1}$$

where $\mathbf{1}_A$ is an indicator function for the event A. This may not be equal to zero for any t_i since it is discontinuous at $t_i = T_j$. Given the opponent's strategy, the optimal behavior of an agent in this game consists of monitoring the (undiscounted) marginal utility $K_i - Z(t).\varphi(\mathbf{x}_i).e^{\mathbf{1}_{(t \geq T_j)}.\delta}$ at each moment of time t. As long as this quantity is positive the individual participates in the initial activity, and he or she switches as soon as the marginal utility becomes less than or equal to zero.

As mentioned previously, the relative flow between the inside and outside activities is the ultimate determinant of an individual's behavior. As is the case with the familiar random utility model, our model identifies relative utilities. For example, suppose that the destination state is retirement, with utility flow given by $Z_1(t)\varphi_1(\mathbf{x}_i)$, and that the utility flow in the non-retirement state is $K_i Z_2(t)\varphi_2(\mathbf{x}_i)$ (where K_i represents initial health, t is age, and $\mathbf{x}_i$ is a set of covariates, and we abstract from the interaction term e^{δ}). This would be observationally equivalent to a model where the utility flow in the current state is K_i and utility in the outside activity is $Z(t)\varphi(\mathbf{x}_i)$ with $Z(t) \equiv Z_1(t)/Z_2(t)$ and $\varphi(\mathbf{x}_i) \equiv \varphi_1(\mathbf{x}_i)/\varphi_2(\mathbf{x}_i)$.

An appropriate concept for optimality in the presence of the interaction represented by δ is that of mutual best responses. Consider the optimal T_i of individual i given that individual j has chosen T_j. It is clear from (1) that

$$\begin{aligned} T_1 &= \inf\{t_1 : K_1 - Z(t_1).\varphi(\mathbf{x}_1).e^{\mathbf{1}_{(t_1 \geq T_2)}.\delta} < 0\} \\ T_2 &= \inf\{t_2 : K_2 - Z(t_2).\varphi(\mathbf{x}_2).e^{\mathbf{1}_{(t_2 \geq T_1)}.\delta} < 0\} \end{aligned} \tag{2}$$

In the absence of interaction ($\delta = 0$), the individual switches at $T_i = Z^{-1}(K_i/\varphi(\mathbf{x}_i))$ or

$$\ln Z(T_i) = -\ln \varphi(\mathbf{x}_i) + \underbrace{\epsilon_i}_{\equiv \ln k_i}$$

which is a semi-parametric generalized accelerated failure time (GAFT) model like the one discussed in Ridder (1990). For example, if $Z(t) = \lambda t^{\alpha_i}$, $\varphi(\mathbf{x}_i) = \exp(\mathbf{x}_i'\beta)$ and $K_i \sim \exp(1)$, the cumulative distribution function of T_i is given by

$$\begin{aligned} F_{T_i}(t) &= \mathbb{P}[(K_i e^{-\mathbf{x}_i'\beta}/\lambda)^{1/\alpha_i} \leq t] \\ &= \mathbb{P}(K_i \leq t^{\alpha_i}\lambda e^{\mathbf{x}_i'\beta}) \\ &= 1 - \exp(-t^{\alpha_i}\lambda \exp(\mathbf{x}_i'\beta)) \end{aligned}$$

and the model corresponds to a proportional hazard duration model with a Weibull baseline hazard.

When $\delta > 0$, the solution to (2) depends on the realization of (K_1, K_2). There are five scenarios depicted in Figure 1.

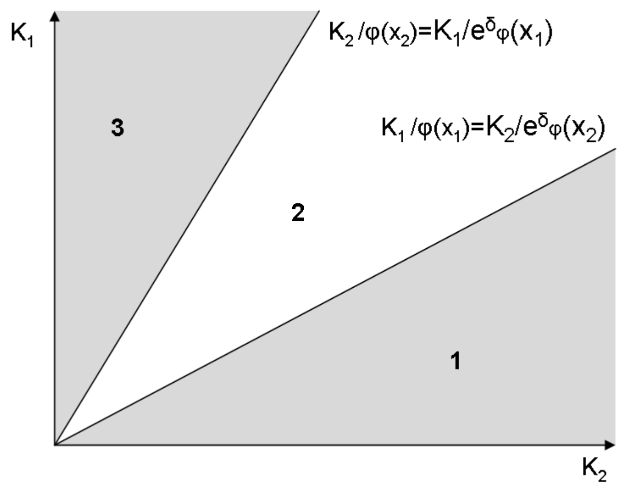

Figure 1: Equilibrium Regions

To understand the alternative scenarios, we first define $\underline{T}_i$ and $\overline{T}_i$, $i = 1, 2$ as the values that set expression (1) to zero when $e^{\mathbf{1}_{(t_i \geq T_j)}\delta} = e^\delta$ and when $e^{\mathbf{1}_{(t_i \geq T_j)}\delta} = 1$, respectively:

$$\begin{aligned} \underline{T}_i &= Z^{-1}(K_i e^{-\delta}/\varphi(\mathbf{x}_i)),\ i = 1, 2 \\ \overline{T}_i &= Z^{-1}(K_i/\varphi(\mathbf{x}_i)),\ i = 1, 2 \end{aligned}$$

Because $\delta > 0$, $\underline{T}_i < \overline{T}_i$, $i = 1, 2$. If $t < \underline{T}_i$ then $Z(t)\varphi(\mathbf{x}_i) - K_i < Z(t)\varphi(\mathbf{x}_i)e^\delta - K_i < 0$, and as a result agent i wouldn't like to switch activities regardless of the other agent's action. Analogously, if $\underline{T}_i < t < \overline{T}_i$, then $Z(t)\varphi(\mathbf{x}_i)e^\delta - K_i > 0$ but $Z(t)\varphi(\mathbf{x}_i) - K_i < 0$, and agent i would switch if the other agent switches, but not if the other player does not. Finally, if $t > \overline{T}_i$, then $Z(t)\varphi(\mathbf{x}_i) - K_i > 0$ and the agent is better off switching at a time less than t.

In region 1 of Figure 1, $T_1 < T_2$ and the equilibrium is unique. This is because the region is such that $K_1/\varphi(\mathbf{x}_1) < K_2 e^{-\delta}/\varphi(\mathbf{x}_2)$ and hence $\overline{T}_1 < \underline{T}_2$. Here, for any t less than $\overline{T}_1$, $Z(t)\varphi(\mathbf{x}_2)e^{\delta} - K_2$ is less than zero and agent 2 has no incentive to switch even if agent 1 has already switched. Also $Z(t)\varphi(\mathbf{x}_2)e^{\delta} - K_2$ is less than zero and agent 1 would not switch either. Once $t > \overline{T}_1$, then $Z(t)\varphi(\mathbf{x}_2)e^{\delta} - K_2$ is strictly greater than 0 and agent one will prefer to have switched earlier, no matter what action the second agent might take. It is therefore optimal for agent 1 to switch at $T_1 = \overline{T}_1$. This in turn induces agent 2 to switch at $T_2 = \underline{T}_2 > T_1$.

In region 2, $T_1 = T_2$ and there are multiple equilibria. This region is given by $K_1/\varphi(\mathbf{x}_1) > K_2 e^{-\delta}/\varphi(\mathbf{x}_2)$ and $K_2/\varphi(\mathbf{x}_2) > K_1 e^{-\delta}/\varphi(\mathbf{x}_1)$. This implies that $\overline{T}_1 > \underline{T}_2$ and $\overline{T}_2 > \underline{T}_1$. To see that individuals will stop simultaneously and there are many equilibria, let

$$\underline{T} = \max\left(\underline{T}_1, \underline{T}_2\right)$$

and

$$\overline{T} = \min\left(\overline{T}_1, \overline{T}_2\right)$$

Because $\overline{T}_1 > \underline{T}_2$ and $\overline{T}_2 > \underline{T}_1$, we have that $\underline{T} \leq \overline{T}$. We now consider three cases depending on t's location relative to $\underline{T}$ and $\overline{T}$. For $t < \underline{T}$, let j be the agent such that $\underline{T} = \underline{T}_j$. Since $t < \underline{T}_j$, individual j would not be willing to switch regardless of the action of the other agent, i. Also since $t < \overline{T}_i$, individual i will not switch either given that individual j does not switch. Hence no agent switches when $t < \underline{T}$. For $\underline{T} \leq t \leq \overline{T}$, $\underline{T}_i \leq t \leq \overline{T}_i$ for each agent. At each point in time in the interval, an agent can therefore do no better than the alternative activity if the other agent has already switched. Hence, any profile such that $\underline{T} \leq T_1 = T_2 \leq \overline{T}$ will be an equilibrium. Finally, for $\overline{T} < t$, $\overline{T}_i < t$ for both individuals and each has an incentive to decrease his or her switching time toward $\overline{T}$ regardless of what the other agent does. Hence, simultaneous switching at any t in the interval $[\underline{T}, \overline{T}]$ is an equilibrium.

Region 3 is similar to region 1. the only difference is that the subscripts have been exchanged. In this region, $T_2 < T_1$ and the equilibrium is unique.

The final two cases are when $K_1/\varphi(\mathbf{x}_1) = K_2 e^{-\delta}/\varphi(\mathbf{x}_2)$ or $K_1/\varphi(\mathbf{x}_1) = K_2 e^{-\delta}/\varphi(\mathbf{x}_2)$. In these cases, the equilibrium is unique and individuals switch simultaneously. Since K_1 and K_2 are continuous random variables, these regions occur with probability zero and we therefore skip a detailed analysis. Regions 1 and 3 also deliver a unique equilibrium. In region 2, a simultaneous switch at any t in $[\underline{T}, \overline{T}]$ would be an equilibrium. This interval will be degenerate if δ is equal to zero. It is also important to note that region 2 can be distinguished from regions 1 and 3 by the econometrician, since this will be used in the identification of the model.

We end this section with a brief discussion on the multiple equilibria encountered in region 2. In our approach, we are agnostic as to which of these equilibria is selected. Some of the solutions in that region may be singled out by different selection criteria nevertheless. The Nash solution concept we use is equivalent to that of an open-loop equilibrium (as discussed, for example, in Fudenberg and Tirole (1991), Section 4.7): one in which individuals condition their strategies on calendar time only and hence commit to this plan of action at the beginning of the game. If individuals can react to events as time unfolds, a closed-loop solution concept that here would be equivalent to subgame perfection would single out the earliest of the Nash equilibria, in which individuals switch at $\underline{T}$. Intuitively, an optimal strategy in region 2 contingent on the game history would prescribe switching simultaneously at any time between $\underline{T}$ and $\overline{T}$. Faced with an opponent carrying such a (closed-loop) strategy, an individual might as well switch as soon as possible to maximize his or her own utility flow. This outcome also corresponds to the Pareto-dominant equilibrium. In this case, the equilibria displayed in our analysis would still be Nash, but not necessarily subgame-perfect. In selecting one of the multiple equilibria that may arise, the early equilibrium is nevertheless a compelling equilibrium and we give it special consideration in the simulation exercises performed later in the paper.

Other selection mechanisms may nonetheless point to later equilibria among the many Nash equilibria available. Players need to know when to act and do so in a coordinated way: to take the initiative a person needs to be confident that he or she will not be acting alone as the switching decision is irreversible. This coordination risk may lead to later switching times. For this reason, we remain agnostic as to which Nash equilibrium is selected.

3 Identification

In this section we ask what aspects of the model can be identified by the data once one recognizes the endogeneity of choices and abstains from an equilibrium selection rule. The proof strategy is similar to that in, for example, Elbers and Ridder (1982) and Heckman and Honoré (1989) applied to the events $T_1 < T_2$ and $T_1 > T_2$. Like those papers, we rely crucially on the continuous nature of the durations, and it is not straightforward to generalize our results to the case where one observes discretized versions of the durations.

The subsequent analysis relies on the following assumptions:

Assumption 1 *K_1 and K_2 are jointly distributed according to $G(\cdot,\cdot)$, where $G(\cdot,\cdot)$ is a continuous cumulative distribution function with full support on $\mathbb{R}^2_+$. Furthermore, its corresponding probability density function $g(\cdot,\cdot)$ is bounded away from zero and infinity in a neighborhood of zero.*

Assumption 2 *The function $Z(\cdot)$ is differentiable with positive derivative.*

Assumption 3 *At least one component of $\mathbf{x}_i$, say $\mathbf{x}_{ik}$, is such that supp($\mathbf{x}_{ik}$) contains an open subset of $\mathbb{R}$.*

Assumption 4 *The range of $\varphi(\cdot)$ is $\mathbb{R}_+$ and it is continuously differentiable with non–zero derivative.*

In Assumption 1, we require that $g(0,0)$ be bounded away from zero and infinity. This assumption is related to assumptions typically used in the MPH/GAFT literature with respect to the distribution of the unobserved heterogeneity component. To see this, consider a bivariate mixed proportional hazards model with durations T_i, $i=1,2$ that are independent conditional on observed and unobserved covariates. The integrated hazard is given by $Z(\cdot)\varphi(\mathbf{x}_i)\theta_i, i=1,2$ with $Z(\cdot)$ as the baseline integrated hazard; $\varphi(\mathbf{x}_i)$, a function of observed covariates $\mathbf{x}_i$; and θ_i, a positive unobserved random variable. In other words, for this model, at the optimal stopping time and when $T_i < T_j$:

$$Z(T_i)\varphi(\mathbf{x}_i) = \tilde{K}_i/\theta_i \equiv K_i, \quad i=1,2$$

where $\tilde{K}_i$ follows a unit exponential distribution (independent of $\mathbf{x}$'s and θ's). See, for example, Ridder (1990). Let $f(\cdot,\cdot)$ denote the joint probability density function for (θ_1,θ_2). Then the joint density for (K_1,K_2), $g(\cdot,\cdot)$, is:

$$g(k_1,k_2) = \int_{\mathbb{R}_+}\int_{\mathbb{R}_+} \theta_1\theta_2 e^{-k_1-k_2} f(\theta_1,\theta_2)d\theta_1 d\theta_2.$$

This gives $g(0,0)=\mathbb{E}(\theta_1\theta_2)$, which is positive by assumption. Our requirement that it be finite is then essentially the finite mean assumption in the traditional mixed proportional hazards model identification literature. Economically, it is clear that the model is observationally equivalent to a model in which the same monotone transformation is applied to the utilities in the two activities. Since a power–transformation would preserve the multiplicative structure assumed here, this means that the model should only be identified up to power transformations. Assumption 1 rules out such a transformation, since the transformed K's would not have finite, nonzero density at the origin.

Assumptions 2-4 are stronger than necessary. Most importantly, the appendix shows that for some of the identification results one can allow $\mathbf{x}_i$ to have a discrete distribution. The identification of $\varphi(\cdot)$ uses variation in at least one component of $\mathbf{x}_i$.

The following results establish that assumptions 1-4 are sufficient (though not necessary in many cases) for the identification of the different components in the model. We begin by analyzing $\varphi(\cdot)$.

Theorem 1 (Identification of $\varphi(\cdot)$) *Under Assumptions 1 and 2, the function $\varphi(\cdot)$ is identified up to scale if* $\operatorname{supp}(\mathbf{x}_1, \mathbf{x}_2) = \operatorname{supp}(\mathbf{x}_1) \times \operatorname{supp}(\mathbf{x}_2)$.

Proof. Consider the absolutely continuous component of the conditional distribution of (T_1, T_2), the switching times for the agents, given the covariates $\mathbf{x}_1, \mathbf{x}_2$. When $T_1 < T_2$, using the fact that $T_1 = Z^{-1}(K_1/\varphi(\mathbf{x}_1))$ and $T_2 = Z^{-1}(K_2 e^{-\delta}/\varphi(\mathbf{x}_2))$, we can use the Jacobian method to obtain the probability density function for (T_1, T_2) on the set $\{(t_1, t_2) \in \mathbb{R}_+^2 : t_1 < t_2\}$. It is given by:

$$f_{T_1,T_2|\mathbf{x}_1,\mathbf{x}_2}(t_1, t_2|\mathbf{x}_1, \mathbf{x}_2) = \lambda(t_1)\varphi(\mathbf{x}_1)\lambda(t_2)\varphi(\mathbf{x}_2)e^{\delta} g(Z(t_1)\varphi(\mathbf{x}_1), Z(t_2)\varphi(\mathbf{x}_2)e^{\delta})$$

where

$$Z(t) = \int_0^t \lambda(s)ds.$$

Given two sets of covariates $(\mathbf{x}_1, \mathbf{x}_2)$ and $(\mathbf{x}_1', \mathbf{x}_2')$ we obtain that

$$\begin{aligned}
\lim_{\substack{(t_1,t_2)\to(0,0)\\ t_1<t_2}} \frac{f_{T_1,T_2|\mathbf{x}_1,\mathbf{x}_2}(t_1, t_2|\mathbf{x}_1', \mathbf{x}_2')}{f_{T_1,T_2|\mathbf{x}_1,\mathbf{x}_2}(t_1, t_2|\mathbf{x}_1, \mathbf{x}_2)} &= \lim_{\substack{(t_1,t_2)\to(0,0)\\ t_1<t_2}} \frac{\varphi(\mathbf{x}_1')\varphi(\mathbf{x}_2')g(Z(t_1)\varphi(\mathbf{x}_1'), Z(t_2)\varphi(\mathbf{x}_2')e^{\delta})}{\varphi(\mathbf{x}_1)\varphi(\mathbf{x}_2)g(Z(t_1)\varphi(\mathbf{x}_1), Z(t_2)\varphi(\mathbf{x}_2)e^{\delta})} \\
&= \frac{\varphi(\mathbf{x}_1')\varphi(\mathbf{x}_2')}{\varphi(\mathbf{x}_1)\varphi(\mathbf{x}_2)} \qquad (3)
\end{aligned}$$

where the last equality uses the fact that $\lim_{t\to 0} Z(t) = 0$. Setting $\mathbf{x}_2 = \mathbf{x}_2'$, which can be done because $\operatorname{supp}(\mathbf{x}_1, \mathbf{x}_2) = \operatorname{supp}(\mathbf{x}_1) \times \operatorname{supp}(\mathbf{x}_2)$, identifies $\varphi(\cdot)$ up to scale. ∎

The condition that $\operatorname{supp}(\mathbf{x}_1, \mathbf{x}_2) = \operatorname{supp}(\mathbf{x}_1) \times \operatorname{supp}(\mathbf{x}_2)$ is stronger than necessary for the identification of $\varphi(\cdot)$. In order to identify $\varphi(\mathbf{x}_1)/\varphi(\mathbf{x}_1')$ all we need is to be able to find $\mathbf{x}_2$ such that $(\mathbf{x}_1, \mathbf{x}_2)$ and $(\mathbf{x}_1', \mathbf{x}_2)$ are in the support. Under certain circumstances, such as in interactions between husband and wife, the players in the games sampled may be easily labeled, say $i = 1, 2$. The proof strategy also allows $\varphi(\cdot)$ to depend on i. We also point out that $\mathbf{x}_i$ is not required to contain continuously distributed components. Finally, the identification of $\varphi(\cdot)$ from (3) would still hold even if the players shared the same covariates $\mathbf{x}_1 = \mathbf{x}_2 = \mathbf{x}$ as long as $\varphi(\cdot)$ is the same for both.

Having identified $\varphi(\cdot)$, we can establish the identification of δ.

Theorem 2 (Identification of δ) *δ is identified under Assumptions 1-4 .*

Proof. Consider the probability

$$\mathbb{P}(T_1 < T_2|\mathbf{x}) = \mathbb{P}\left(\frac{K_1}{\varphi(\mathbf{x}_1)} < \frac{K_2 e^{-\delta}}{\varphi(\mathbf{x}_2)}\right) = \mathbb{P}\left(\ln K_1 - \ln K_2 + \delta < \ln\left(\varphi(\mathbf{x}_1)/\varphi(\mathbf{x}_2)\right)\right). \quad (4)$$

Since $\varphi(\cdot)$ is identified up to scale (because of Assumptions 1 and 2), as one varies $\mathbf{x}_1$ and $\mathbf{x}_2$, the probability above traces the cumulative distribution function for the random variable $W = \ln K_1 - \ln K_2 + \delta$ (given Assumptions 3 and 4). Likewise, the probability

$$\mathbb{P}(T_1 > T_2|\mathbf{x}) = \mathbb{P}\left(\frac{K_1 e^{-\delta}}{\varphi(\mathbf{x}_1)} < \frac{K_2}{\varphi(\mathbf{x}_2)}\right) = \mathbb{P}\left(\ln K_1 - \ln K_2 - \delta > \ln\left(\varphi(\mathbf{x}_1)/\varphi(\mathbf{x}_2)\right)\right) \quad (5)$$

traces the survivor function (and consequently the cumulative distribution function) for the random variable $\ln K_1 - \ln K_2 - \delta = W - 2\delta$. Since this is basically the random variable W displaced by 2δ, this difference is identified as the (horizontal) distance between the two cumulative distribution functions that are identified from the data (the events $T_1 > T_2$ and $T_1 < T_2$ conditioned on $\mathbf{x}$). Figure (2) illustrates this idea.

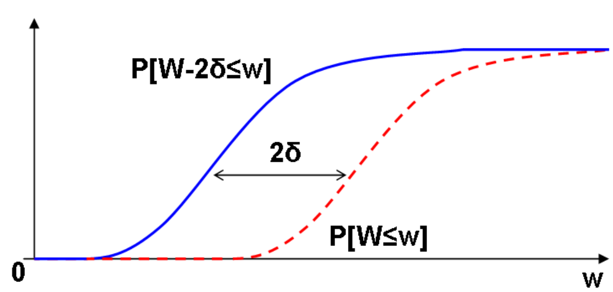

Figure 2: Identification of δ

From this argument, the parameter δ is identified. ■

In the proof of Theorem 2, Assumptions 1 and 2 are invoked to guarantee the identification of $\varphi(\cdot)$. If this function is identified for other reasons, we can dispense with this assumption.

Finally we establish the identification of $Z(\cdot)$ and $G(\cdot,\cdot)$, the join distribution of K_1 and K_2.

Theorem 3 (Identification of $Z(\cdot)$ and $G(\cdot,\cdot)$) *Under Assumptions 1-4, the function $Z(\cdot)$ is identified up to scale, and the distribution $G(\cdot,\cdot)$ is identified up to a scale transformation.*

Proof. We first consider identification of $Z(\cdot)$. On the set $\{(t_1,t_2)\in\mathbb{R}_+^2 : t_1<t_2\}$, consider the function

$$\begin{aligned} h(t_1,t_2,\mathbf{x}_1,\mathbf{x}_2) &= \int_0^{t_1}\int_{t_2}^{\infty} f_{T_1,T_2|\mathbf{x}_1,\mathbf{x}_2}(s_1,s_2|\mathbf{x}_1,\mathbf{x}_2)ds_2ds_1 \\ &= \int_0^{t_1}\int_{t_2}^{\infty} \lambda(s_1)\varphi(\mathbf{x}_1)\lambda(s_2)\varphi(\mathbf{x}_2)e^{\delta}g(Z(s_1)\varphi(\mathbf{x}_1), Z(s_2)\varphi(\mathbf{x}_2)e^{\delta})ds_2ds_1 \end{aligned}$$

Consider the change of variables:

$$\xi_1 = Z(s_1)\varphi(\mathbf{x}_1) \qquad \xi_2 = Z(s_2)e^{\delta}\varphi(\mathbf{x}_2)$$

and rewrite h as

$$h(t_1,t_2,\mathbf{x}_1,\mathbf{x}_2) = \int_0^{Z(t_1)\varphi(\mathbf{x}_1)}\int_{Z(t_2)e^{\delta}\varphi(\mathbf{x}_2)}^{\infty} g(\xi_1,\xi_2)d\xi_1 d\xi_2$$

Then notice that

$$\frac{\partial h/\partial t_1}{\partial h/\partial \mathbf{x}_{1k}} = \frac{\lambda(t_1)\varphi(\mathbf{x}_1)}{Z(t_1)\partial_k\varphi(\mathbf{x}_1)} = \frac{d\ln Z\,(t_1)}{dt_1}\frac{\varphi(\mathbf{x}_1)}{\partial_k\varphi(\mathbf{x}_1)}.$$

integrating and exponentiating yields

$$CZ(s)^{\varphi(\mathbf{x}_1)/\partial_k\varphi(\mathbf{x}_1)}$$

where C is a constant. Given the identification of $\varphi(\cdot)$ up to scale, $Z(\cdot)$ is therefore identified up to scale (the constant C).

We next turn to identification of $G(\cdot,\cdot)$. Note that h defines the cumulative distribution function of $(K_1,-K_2)$, which can be traced out by varying $Z(t_1)\varphi(\mathbf{x}_1)$ and $Z(t_2)e^{\delta}\varphi(\mathbf{x}_2)$ (making sure that $t_1<t_2$). Since δ is identified and $Z(\cdot)$ and $\varphi(\cdot)$ are identified up to scale, the distribution of $(K_1,-K_2)$ is identified up to a scale transformation. The distribution of (K_1,K_2) is therefore identified up to a scale transformation. ∎

The mechanics of the proof suggests that we can also allow $Z(\cdot)$ to depend on i as is the case with $\varphi(\cdot)$, but the characterization of the equilibrium in section 2 assumes $Z(\cdot)$ to be the same for both individuals. As in the previous result, the identification would still hold were the covariates for the two agents identical for a given draw of the game ($\mathbf{x}_1 = \mathbf{x}_2 = \mathbf{x}$). The requirement that $\mathbf{x}_i$ contain a continuously distributed component is not necessary either. In the appendix we present an alternative proof that dispenses with that assumption.

4 Extensions and Alternative Models

In this section, we discuss results for some variations on the model depicted in section 2.

4.1 Individual–specific δ

As mentioned earlier, in certain problems (such as the interaction between husband and wife) players may be easily labeled. In this case, one can consider different δs for different players: $\delta_i,\ i = 1,2$. The previous result would render identification for $\delta_1+\delta_2$. The following establishes the identification of $\delta_1 - \delta_2$ and hence of $\delta_i,\ i = 1,2$.

Theorem 4 (Identification of $\delta_i,\ i = 1,2$) *$\delta_i,\ i = 1,2$ are identified under Assumptions 1-4.*

Proof. The sum $\delta_1+\delta_2$ is identified according to the arguments in the previous theorem. Let $k > 1$. Then

$$\begin{aligned}\frac{\lim_{s\to 0} f_{T_1,T_2|\mathbf{x}_1,\mathbf{x}_2}(s,ks|\mathbf{x}_1,\mathbf{x}_2)}{\lim_{s\to 0} f_{T_1,T_2|\mathbf{x}_1,\mathbf{x}_2}(ks,s|\mathbf{x}_1,\mathbf{x}_2)} &= \frac{\lim_{s\to 0}\lambda(s)\lambda(ks)\varphi(\mathbf{x}_1)\varphi(\mathbf{x}_2)e^{\delta_2}}{\lim_{s\to 0}\lambda(ks)\lambda(s)\varphi(\mathbf{x}_1)\varphi(\mathbf{x}_2)e^{\delta_1}} \\ &= \frac{\varphi(\mathbf{x}_1)\varphi(\mathbf{x}_2)e^{\delta_2}}{\varphi(\mathbf{x}_1)\varphi(\mathbf{x}_2)e^{\delta_1}} \times \lim_{\substack{s\to 0\\ k>1}} \frac{\lambda(s)\lambda(ks)}{\lambda(ks)\lambda(s)} \\ &= e^{\delta_2-\delta_1}\end{aligned}$$

which identifies $\delta_2 - \delta_1$. This and the previous result identify $\delta_i,\ i = 1,2$. ■

It is also possible to allow δ_1 and δ_2 to depend on x_1 and x_2, respectively.[3] In that case the right–hand side of (3) becomes $\frac{\varphi(\mathbf{x}_1')\varphi(\mathbf{x}_2')e^{\delta(\mathbf{x}_2')}}{\varphi(\mathbf{x}_1)\varphi(\mathbf{x}_2)e^{\delta(\mathbf{x}_2)}}$, which again identifies φ up to scale

[3] We thank a referee for pointing this out.

(by varying $\mathbf{x}_1'$). Varying $\mathbf{x}_1$ in (4) and $\mathbf{x}_2$ in (5) identify the cumulative distribution function of $\ln K_1 - \ln K_2 + \delta_2(\mathbf{x}_2)$ and $\ln K_1 - \ln K_2 - \delta_1(\mathbf{x}_1)$, so $\delta_2(\mathbf{x}_2) + \delta_1(\mathbf{x}_1)$ is identified and $\delta_2(\mathbf{x}_2) - \delta_1(\mathbf{x}_1)$ is identified by the same argument as in Theorem 4. Finally, the proof of Theorem 3 is unchanged.

4.2 Common Shock

Since we do not impose independence between K_1 and K_2, some association in the latent utility flow obtained in the initial activity is allowed. Another source of correlation may be represented by a common shock that drives both individuals to the outside activity concurrently. Even under such extreme circumstances, some aspects of the structure remain identified.

A natural way to introduce this non-strategic shock in the model would follow the motivation in Cox and Oakes (1984). Assume that a common shock that drives both spells to termination at the same time happens at a random time $V > 0$. Denote the probability density function of V by $h(\cdot)$. Individuals switch for two possible reasons: either they deem the decision to be optimal as in the original model; or they are driven out of the initial activity by the common shock. If both individuals are still in the initial activity when the shock arrives, they both switch simultaneously. If one of them switches before the shock arrives, the second one is driven out of the initial activity earlier than he or she would have voluntarily chosen.[4] In keeping with the notation used so far, let T_i be the switching time chosen by individual i and $\tilde{T}_i = \min\{T_i, V\}$, the switching time observed by the econometrician. We then have the following result:

Theorem 5 (Identification of $\varphi(\cdot)$ with Common Shocks) *Suppose Assumptions 1 and 2 hold and* $\operatorname{supp}(\mathbf{x}_1, \mathbf{x}_2) = \operatorname{supp}(\mathbf{x}_1) \times \operatorname{supp}(\mathbf{x}_2)$. *Furthermore assume that the common shock, V, is independent of* $\mathbf{x}_i, K_i,\ i = 1, 2$. *Then the function $\varphi(\cdot)$ is identified up to scale.*

Proof. The proof is similar to that of Theorem 1. Consider the absolutely continuous component of the conditional distribution of $(\tilde{T}_1, \tilde{T}_2)$, the observed switching times for the individuals, given the covariates $\mathbf{x}_1, \mathbf{x}_2$. As in the proof for Theorem 1 and using the definition

[4]The optimal switching times derived in section 2 would still hold. Should the realizations of V happen after that chosen time, the individual would have no incentives to wait. If v arrives earlier than the optimal time, there would be no incentive to anticipate the switch nor would there be anything to be done about it after the shock.

of $\tilde{T}_i = \min\{T_i, V\}$, we can obtain that the probability density function for this pair on the set $\{(\tilde{t}_1, \tilde{t}_2) \in \mathbb{R}_+^2 : \tilde{t}_1 < \tilde{t}_2\}$ is given by:

$$\begin{aligned} f_{\tilde{T}_1,\tilde{T}_2|\mathbf{x}_1,\mathbf{x}_2}(\tilde{t}_1, \tilde{t}_2|\mathbf{x}_1, \mathbf{x}_2) &= \lambda(\tilde{t}_1)\varphi(\mathbf{x}_1)\lambda(\tilde{t}_2)\varphi(\mathbf{x}_2)e^{\delta}g(Z(\tilde{t}_1)\varphi(\mathbf{x}_1), Z(\tilde{t}_2)\varphi(\mathbf{x}_2)e^{\delta})\mathbb{P}(V > \tilde{t}_2) \\ &\quad +\lambda(\tilde{t}_1)\varphi(\mathbf{x}_1)h(\tilde{t}_2)\int_{\tilde{t}_2}^{\infty} g(Z(\tilde{t}_1)\varphi(\mathbf{x}_1), Z(s)\varphi(\mathbf{x}_2)e^{\delta})ds \end{aligned}$$

where

$$Z(t) = \int_0^t \lambda(s)ds, i = 1, 2.$$

Given two sets of covariates $(\mathbf{x}_1, \mathbf{x}_2)$ and $(\mathbf{x}_1', \mathbf{x}_2)$ we can again obtain that

$$\lim_{\substack{(\tilde{t}_1,\tilde{t}_2)\to(0,0)\\ \tilde{t}_1<\tilde{t}_2}} \frac{f_{\tilde{T}_1,\tilde{T}_2|\mathbf{x}_1,\mathbf{x}_2}(\tilde{t}_1, \tilde{t}_2|\mathbf{x}_1', \mathbf{x}_2)}{f_{\tilde{T}_1,\tilde{T}_2|\mathbf{x}_1,\mathbf{x}_2}(\tilde{t}_1, \tilde{t}_2|\mathbf{x}_1, \mathbf{x}_2)} = \frac{\varphi(\mathbf{x}_1')}{\varphi(\mathbf{x}_1)}$$

using the assumption that $\lim_{t\to 0} Z(t) = 0$. So, $\varphi(\cdot)$ is identified up to a scale transformation. ■

The assumption that $\text{supp}(\mathbf{x}_1, \mathbf{x}_2) = \text{supp}(\mathbf{x}_1) \times \text{supp}(\mathbf{x}_2)$ is stronger than necessary. The proof strategy also allows $\varphi(\cdot)$ to depend on i.

Theorem 5 establishes that it is possible to identify the effects of covariates in a model that also allows for common shocks. We next address the question of whether our strategic model is generically distinguishable from the model proposed in Marshall and Olkin (1967). We do this in a setting without covariates. This is equivalent to allowing for covariates in a completely general way and then conditioning on them.

Marshall and Olkin (1967) present a model with three types of shock: one leading to joint spell termination and two leading to individual spell terminations. The corresponding survivor function is given by:

$$S(t_1, t_2) = \exp\left(-H_1(t_1) - H_2(t_2) - H_{12}(\max(t_1, t_2))\right)$$

where H_i, $i = 1, 2$ represent the integrated hazards for the two individual shocks and H_{12} denotes the integrated hazard for the joint shock.[5] We will assume H_1, H_2 and H_{12} are continuously differentiable and strictly increasing with $H_1(0) = H_2(0) = H_{12}(0) = 0$ and $\lim_{t\to\infty} H_1(t) = \lim_{t\to\infty} H_2(t) = \lim_{t\to\infty} H_{12}(t) = \infty$. In other words, the durations until each shock are continuously distributed, strictly positive and finite random variables.

[5]In the original paper, $H_i(t)$, $i = 1, 2$ and $H_{12}(t)$ are linear functions of time.

This leads to the following density on $\mathbb{R}_+^2 - \{(t_1, t_2) \in \mathbb{R}_+^2 : t_1 = t_2\}$:

$$f(t_1, t_2) = \begin{cases} (H_1'(t_1) + H_{12}'(t_1)) H_2'(t_2) \exp(-H_1(t_1) - H_2(t_2) - H_{12}(t_1)) & t_1 > t_2 \\ H_1'(t_1)(H_2'(t_2) + H_{12}'(t_2)) \exp(-H_1(t_1) - H_2(t_2) - H_{12}(t_2)) & t_1 < t_2 \end{cases}$$

For comparison, a version of our model without covariates would define outside utility functions for individuals 1 and 2 as

$$Z_1(t) e^{\delta \mathbf{1}_{(t>t_2)}}$$

and

$$Z_2(t) e^{\delta \mathbf{1}_{(t>t_1)}},$$

respectively. The inside utility flows are given by K_i $(i = 1, 2)$. In order to simplify the comparison to Marshall and Olkin (1967) we assume that the K_i's are independent unit exponential random variables. We will assume that Z_1 and Z_2 are continuously differentiable and strictly increasing with $Z_1(0) = Z_2(0) = 0$ and $\lim_{t\to\infty} Z_1(t) = \lim_{t\to\infty} Z_2(t) = \infty$. In other words, in the absence of the other player, each agent would have a continuously distributed, strictly positive and finite duration.

When $T_1 > T_2$

$$\begin{aligned} K_1 &= Z_1(T_1) e^{\delta} = \widetilde{Z}_1(T_1) \\ K_2 &= Z_2(T_2). \end{aligned}$$

This yields the following density for $t_1 > t_2$:

$$e^{\delta} Z_1'(t_1) Z_2'(t_2) \exp\left(-Z_1(t_1) e^{\delta} - Z_2(t_2)\right)$$

and analogously the density is

$$e^{\delta} Z_1'(t_1) Z_2'(t_2) \exp\left(-Z_1(t_1) - Z_2(t_2) e^{\delta}\right)$$

when $t_1 < t_2$. For the two models to coincide when $t_1 > t_2$, we would need that

$$\begin{gathered} (H_1'(t_1) + H_{12}'(t_1)) H_2'(t_2) \exp(-H_1(t_1) - H_2(t_2) - H_{12}(t_1)) = \\ e^{\delta} Z_1'(t_1) Z_2'(t_2) \exp\left(-Z_1(t_1) e^{\delta} - Z_2(t_2)\right). \end{gathered} \tag{6}$$

Taking logs and differentiating with respect to t_2 implies that

$$\frac{H_2''(t_2)}{H_2'(t_2)} - H_2'(t_2) = \frac{Z_2''(t_2)}{Z_2'(t_2)} - Z_2'(t_2) \quad \forall t_2.$$

Integrating, exponentiating and integrating again yields

$$\exp\left(-H_2\left(t_2\right)\right) = c_1 \exp\left(-Z_2\left(t_2\right)\right) + c_2$$

Using $H_2(0) = Z_2(0) = 0$ yields $c_1 + c_2 = 1$. The assumption that $\lim_{t\to\infty} H_2\left(t\right) = \lim_{t\to\infty} Z_2\left(t\right) = \infty$ yields $c_1 = 1$. Hence $H_2(t) = Z_2(t)$. A symmetric argument leads to $H_1(t) = Z_1(t)$.

Replacing these in expression (6) and rearranging, we obtain that:

$$\exp(e^\delta) = (1 + Z_{12}'(t_1)/Z_1'(t_1)) \exp(-Z_1(t_1) - Z_{12}(t_1) - Z_1(t_1)e^\delta)$$

Taking limits as $t_1 \to 0$ we get:

$$\lim_{t_1\to 0} \frac{Z_{12}'(t_1)}{Z_1'(t_1)} = \exp(e^\delta) - 1 \Leftrightarrow \lim_{t_1\to 0} \frac{Z_1'(t_1)}{Z_{12}'(t_1)} = (\exp(e^\delta) - 1)^{-1}$$

Analogously,

$$\lim_{t_2\to 0} \frac{Z_2'(t_2)}{Z_{12}'(t_2)} = (\exp(e^\delta) - 1)^{-1}.$$

Now note that the strategic model implies that

$$T_1 \geq \tilde{Z}_1^{-1}\left(K_1\right)$$

$$T_2 \geq \tilde{Z}_2^{-1}\left(K_2\right)$$

and consequently

$$\begin{aligned} P\left(T_1 \leq s, T_2 \leq s\right) &\leq P\left(Z_1^{-1}\left(K_1 e^{-\delta}\right) \leq s, Z_2^{-1}\left(K_2 e^{-\delta}\right) \leq s\right) \\ &= P\left(K_1 \leq Z_1\left(s\right) e^\delta, K_2 \leq Z_2\left(s\right) e^\delta\right) \\ &= \left(1 - \exp\left(-Z_1\left(s\right) e^\delta\right)\right)\left(1 - \exp\left(-Z_2\left(s\right) e^\delta\right)\right) \end{aligned} \tag{7}$$

for all s. At the same time, Marshall-Olkin's model would yield

$$\begin{aligned} P\left(T_1 \leq s, T_2 \leq s\right) = & \left(1 - \exp\left(-H_{12}\left(s\right)\right)\right) + \left(1 - \exp\left(-H_1\left(s\right)\right)\right)\left(1 - \exp\left(-H_2\left(s\right)\right)\right) - \\ & - \left(1 - \exp\left(-H_{12}\left(s\right)\right)\right)\left(1 - \exp\left(-H_1\left(s\right)\right)\right)\left(1 - \exp\left(-H_2\left(s\right)\right)\right). \end{aligned} \tag{8}$$

Now let

$$\begin{aligned} a(s) &= \exp\left(-Z_1\left(s\right)\right) = \exp\left(-H_1\left(s\right)\right) \\ b(s) &= \exp\left(-Z_2\left(s\right)\right) = \exp\left(-H_2\left(s\right)\right) \\ c(s) &= \exp\left(-H_{12}\left(s\right)\right). \end{aligned}$$

Suppressing the argument, s, (7) and (8) imply that

$$c(ab-b-a)+1 \leq (1-a^{\exp(\delta)})(1-b^{\exp(\delta)}).$$

For $s>0$, the left-handside expression is positive, since it is the joint cumulative distribution at $t_1=t_2=s$ for the Marshall-Olkin model. Then,

$$1 \leq \frac{(1-a^{\exp(\delta)})(1-b^{\exp(\delta)})}{c(ab-b-a)+1}$$

Taking limits as $s \to 0$:

$$1 \leq \lim_{s\to 0} \frac{(1-a(s)^{\exp(\delta)})(1-b(s)^{\exp(\delta)})}{c(s)(a(s)b(s)-b(s)-a(s))+1}$$

$$= \lim_{s\to 0} \frac{-a'e^{\delta}a^{\exp(\delta)-1}(1-b^{\exp(\delta)})-b'e^{\delta}b^{\exp(\delta)-1}(1-a^{\exp(\delta)})}{a'bc+ab'c+abc'-b'c-bc'-a'c-ac'}$$

where the equality uses l'Hôpital's rule and arguments are omitted for notational convenience. Divide numerator and denominator in the last expression by $Z_{12}'(s)$ and notice that

$$\lim_{s\to 0} a(s) = \lim_{s\to 0} b(s) = \lim_{s\to 0} c(s) = 1$$

and

$$\lim_{s\to 0} \frac{a'(s)}{Z_{12}'(s)} = \lim_{s\to 0} \frac{b'(s)}{Z_{12}'(s)} = -(\exp(e^{\delta})-1)^{-1}.$$

The last line follows from $a'(s) = -Z_1'(s)a(s)$ and $b'(s) = -Z_2'(s)b(s)$ plus the fact that $\lim_{s\to 0}\frac{Z_1'(s)}{Z_{12}'(s)} = \lim_{s\to 0}\frac{Z_2'(s)}{Z_{12}'(s)} = (\exp(e^{\delta})-1)^{-1}$. We can similarly obtain that $\lim_{s\to 0}\frac{c'(s)}{Z_{12}'(s)} = -1$. These then imply that for the numerator

$$\lim_{s\to 0}\left[-\frac{a'}{Z_{12}'}e^{\delta}a^{\exp(\delta)-1}(1-b^{\exp(\delta)})-\frac{b'}{Z_{12}'}e^{\delta}b^{\exp(\delta)-1}(1-a^{\exp(\delta)})\right]$$
$$= 2(\exp(e^{\delta})-1)^{-1}e^{\delta}\times 1\times(1-1)=0.$$

The denominator on the other hand yields

$$\begin{aligned} & \lim_{s\to 0}\left[\frac{a'}{Z_{12}'}bc+a\frac{b'}{Z_{12}'}c+ab\frac{c'}{Z_{12}'}-\frac{b'}{Z_{12}'}c-b\frac{c'}{Z_{12}'}-\frac{a'}{Z_{12}'}c-a\frac{c'}{Z_{12}'}\right] \\ = \; & -(\exp(e^{\delta})-1)^{-1}-(\exp(e^{\delta})-1)^{-1}-1+(\exp(e^{\delta})-1)^{-1}+1+(\exp(e^{\delta})-1)^{-1}+1 \\ = \; & 1. \end{aligned}$$

This leads to the contradiction $1 \leq 0$, and the two models cannot be observationally equivalent.

As will be seen shortly, the Marshall-Olkin model is closer to the strategic model with an additive externality than to the model where it is multiplicative. It is therefore also interesting to investigate whether such a model is distinguishable from the Marshall–Olkin model. We specify that the outside utility for agents 1 and 2 equals

$$Z_1(t) + Z_{12}(t)\mathbf{1}_{(t>t_2)}$$

and

$$Z_2(t) + Z_{12}(t)\mathbf{1}_{(t>t_1)},$$

respectively. Here, the externality is allowed to be a non–decreasing, time–dependent function Z_{12}. The inside utility flows are again given by independent unit exponential random variables, $K_i, \;\; i = 1, 2$. When $T_1 > T_2$

$$\begin{aligned} K_1 &= Z_1(T_1) + Z_{12}(T_1) = \widetilde{Z}_1(T_1) \\ K_2 &= Z_2(T_2) \end{aligned}$$

Consequently, the density of (T_1, T_2) when $t_1 > t_2$ is:

$$(Z_1'(t_1) + Z_{12}'(t_1))\, Z_2'(t_2) \exp(-Z_1(t_1) - Z_2(t_2) - Z_{12}(t_1))$$

and analogously we obtain that the density is

$$Z_1'(t_1)\,(Z_2'(t_2) + Z_{12}'(t_2)) \exp(-Z_1(t_1) - Z_2(t_2) - Z_{12}(t_2))$$

when $t_1 < t_2$. If $Z_i(t) = H_i(t)$, $i = 1, 2$ and $Z_{12}(t) = H_{12}(t)$ the two models coincide for $t_1 \neq t_2$. This is why we consider it more natural to compare the Marshall-Olkin model to the additive specification of the strategic model. An argument similar to that for the multiplicative model yields that the two coincide for $t_1 \neq t_2$ only if $Z_i(t) = H_i(t)$, $i = 1, 2$ and $Z_{12}(t) = H_{12}(t)$. Note then that the strategic model implies that

$$T_1 \geq \tilde{Z}_1^{-1}(K_1)$$

$$T_2 \geq \tilde{Z}_2^{-1}(K_2)$$

and consequently

$$\begin{aligned} P(T_1 \le s, T_2 \le s) &\le P\left(\widetilde{Z}_1^{-1}(K_1) \le s, \widetilde{Z}_2^{-1}(K_2) \le s\right) \\ &= P\left(K_1 \le \widetilde{Z}_1(s), K_2 \le \widetilde{Z}_2(s)\right) \\ &= \left(1 - \exp\left(-\widetilde{Z}_1(s)\right)\right)\left(1 - \exp\left(-\widetilde{Z}_2(s)\right)\right) \\ &= (1 - \exp(-Z_1(s) - Z_{12}(s)))(1 - \exp(-Z_2(s) - Z_{12}(s))). \end{aligned} \tag{9}$$

Defining a, b and c as before,and noting that now $c = \exp(-H_{12}(s)) = \exp(-Z_{12}(s))$, (9) and (8) imply that

$$c \ge 1 \Rightarrow Z_{12}(s) \le 0.$$

This can only happen if $Z_{12}(s) = 0$ and there are no simultaneous exits in either model.

4.3 Gradual Interaction[6]

In our original model, the impact of an agent's transition on the utility flow of the other individual ($e^{\mathbf{1}(s \ge T_j)\delta}$) is immediate and permanent. This may be convenient in many situations. Consider for instance two nearby retail establishments contemplating price changes to the goods they sell. If one of the stores changes its prices, we would expect its competitor to follow suit without much delay, if any. Other examples may call for a more gradual effect. Consider, for example, two people deciding to adopt a new operating system, and one benefits from having other users of the operating system with whom to share applications and knowledge about the program. If it takes time for one individual to learn and adjust to a new operating system, the benefits provided by another user may accrue gradually. This variation may be captured by assuming that the relative utility flow for individual i at a time t is given by:

$$Z(t)\varphi(\mathbf{x}_i)e^{\delta(t-T_j)} - K_i$$

where $\delta(t - T_j)$ is an increasing function with $\delta(t - T_j) = 0$ for $t < T_j$. If $\delta(\cdot)$ is a continuous function, the probability of simultaneous transitions is zero (region 2 collapses) but the endogeneity is still present.

There are now two relevant possibilities: $T_1 > T_2$ and $T_1 < T_2$ (as mentioned, $T_1 = T_2$ occurs with zero probability). The first–order conditions for agents 1 and 2 are:

$$Z(T_i)e^{\delta(T_i - T_j)} = K_i/\varphi(\mathbf{x}_i), \quad i \ne j = 1, 2.$$

[6] We thank a referee for suggesting this extension.

Consider first the case where $T_1 > T_2$. Here,

$$\begin{aligned} T_2 &= Z^{-1}(K_2/\varphi(\mathbf{x}_2)) \\ T_1 &= Z_*^{-1}(K_1/\varphi(\mathbf{x}_1); T_2) \end{aligned}$$

where $Z_*(s;t) = Z(s)e^{\delta(s-t)}$ and we denote its inverse with respect to the first argument for a given t, $Z_*^{-1}(\cdot;t)$. $T_1 > T_2$ will occur if

$$T_1 = Z_*^{-1}(K_1/\varphi(\mathbf{x}_1); Z^{-1}(K_2/\varphi(\mathbf{x}_2)) > Z^{-1}(K_2/\varphi(\mathbf{x}_2)) = T_2$$

which is equivalent to

$$K_1/\varphi(\mathbf{x}_1) > Z_*(Z^{-1}(K_2/\varphi(\mathbf{x}_2)); Z^{-1}(K_2/\varphi(\mathbf{x}_2))) = K_2/\varphi(\mathbf{x}_2).$$

We obtain analogously that $T_2 > T_1$ when $K_2/\varphi(\mathbf{x}_2) > K_1/\varphi(\mathbf{x}_1)$. This makes sense: the person for whom the inside activity utility flow is higher switches states later. An argument like Theorem 1 can then be used to obtain identification of $\varphi(\cdot)$ up to scale. The following result establishes the identification of $Z(\cdot)$, $G(\cdot,\cdot)$ (both up to scale transformations) and $\delta(\cdot)$.

Theorem 6 (Identification of $Z(\cdot), G(\cdot,\cdot)$ and $\delta(\cdot)$ with Gradual Interaction) *If $\delta(\cdot)$ is increasing and differentiable, then under Assumptions 1-4: the function $Z(\cdot)$ is identified up to scale, the distribution $G(\cdot,\cdot)$ is identified up to a scale transformation and $\delta(\cdot)$ is identified .*

Proof. We first consider identification of $Z(\cdot)$. As in Theorem 3, on the set $\{(t,\Delta t) \in \mathbb{R}_+^2\}$, consider the function

$$h(t,\Delta t,\mathbf{x}_1,\mathbf{x}_2) = \int_0^{Z(t)\varphi(\mathbf{x}_1)} \int_{Z(t+\Delta t)e^{\delta(\Delta t)}\varphi(\mathbf{x}_2)}^{\infty} g(\xi_1,\xi_2)d\xi_1 d\xi_2.$$

As in Theorem 3, this function is the probability that agent 1 switches before t and that agent 2 leaves after $t+\Delta t$. Now, define

$$\overline{h}(t,\mathbf{x}) = \lim_{\substack{\Delta t \to 0 \\ \Delta t > 0}} h(t,\Delta t,\mathbf{x},\mathbf{x}).$$

Then notice that

$$\begin{aligned} \frac{\partial \overline{h}/\partial t}{\partial \overline{h}/\partial_k \mathbf{x}} &= \frac{\lambda(t)\varphi(\mathbf{x})[\int_{Z(t)\varphi(\mathbf{x})}^{\infty} g(Z(t)\varphi(\mathbf{x}),\xi_2)d\xi_2 - \int_0^{Z(t)\varphi(\mathbf{x})} g(\xi_1,Z(t)\varphi(\mathbf{x}))d\xi_1}{Z(t)\partial_k\varphi(\mathbf{x})\left[\int_{Z(t)\varphi(\mathbf{x})}^{\infty} g(Z(t)\varphi(\mathbf{x}),\xi_2)d\xi_2 - \int_0^{Z(t)\varphi(\mathbf{x})} g(\xi_1,Z(t)\varphi(\mathbf{x}))d\xi_1\right]} \\ &= \frac{\lambda(t)\varphi(\mathbf{x})}{Z(t)\partial_k\varphi(\mathbf{x})} \end{aligned}$$

and the proof proceeds as in Theorem 3.

To identify $G(\cdot,\cdot)$, note that

$$\overline{\overline{h}}(t,\mathbf{x}_1,\mathbf{x}_2)=\lim_{\substack{\Delta t\to 0\\ \Delta t>0}} h(t,\Delta t,\mathbf{x}_1,\mathbf{x}_2)$$

defines the cumulative distribution function of $(K_1,-K_2)$, which can be traced out as $Z(t)\varphi(\mathbf{x}_1)$ and $Z(t)\varphi(\mathbf{x}_2)$ are varied. Since $Z(\cdot)$ and $\varphi(\cdot)$ are identified up to scale, the distribution of $(K_1,-K_2)$ is identified up to a scale transformation. Finally, since $(K_1,-K_2)\mapsto(K_1,K_2)$ is a one-to-one mapping, the distribution of (K_1,K_2) is identified up to a scale transformation.

Finally, to identify $\delta(\cdot)$ consider:

$$\frac{\partial h/\partial\Delta t}{\partial h/\partial\mathbf{x}_{2k}}=\frac{\varphi(\mathbf{x}_2)\left(\lambda(t+\Delta t)+Z(t+\Delta t)\delta'(\Delta t)\right)}{\partial_k\varphi(\mathbf{x}_2)Z(t+\Delta t)}$$

or, equivalently:

$$\delta'(\Delta t)=\frac{\partial h/\partial\Delta t}{\partial h/\partial\mathbf{x}_{2k}}\frac{\partial_k\varphi(\mathbf{x}_2)}{\varphi(\mathbf{x}_2)}-\frac{\lambda(t+\Delta t)}{Z(t+\Delta t)}$$

which, given the boundary condition $\delta(0)=0$, identifies $\delta(\cdot)$. ■

5 Estimation Strategies

Consider first the case where $G(\cdot)$ is known. In the absence of interaction effects (δ) and when $G(\cdot)$ is a unit exponential, this would correspond to a classical proportional hazard model. The probability of the event $\{T_1<T_2\}$ this is:

$$\begin{aligned}\mathbb{P}(T_1<T_2|\mathbf{x}_1,\mathbf{x}_2) &= \mathbb{P}(K_1\varphi(\mathbf{x}_2)e^{\delta}/\varphi(\mathbf{x}_1)<K_2|\mathbf{x}_1,\mathbf{x}_2) \\ &= \int_0^{+\infty}\int_{\xi_1\varphi(\mathbf{x}_2)e^{\delta}/\varphi(\mathbf{x}_1)}^{+\infty} g(\xi_1,\xi_2)d\xi_2 d\xi_1\end{aligned} \tag{10}$$

and a similar expression would hold for $\{T_2<T_1\}$. Assume that $Z(\cdot)$, $\varphi(\cdot)$ and $g(\cdot,\cdot)$ are modelled up to the (finite-dimensional) parameters α, β and θ respectively ($Z(\cdot)\equiv Z(\cdot;\alpha)$, $\varphi(\cdot)\equiv\varphi(\cdot;\beta)$ and $g(\cdot,\cdot)\equiv g(\cdot,\cdot;\theta)$). Given data on the realization of the game analyzed in section 3 of this paper and pooling the observations with $T_1=T_2$, we then obtain the

likelihood function

$$\begin{aligned}\mathcal{L}(\alpha,\beta,\theta,\delta) \quad \equiv \quad & \Pi_{t_1<t_2}\Big\{\partial_t Z(t_1;\alpha)\varphi(\mathbf{x}_1;\beta)\partial_t Z(t_2;\alpha)\varphi(\mathbf{x}_2;\beta)e^{\delta} \\ & \qquad \times g(Z(t_1;\alpha)\varphi(\mathbf{x}_1;\beta), Z(t_2;\alpha)\varphi(\mathbf{x}_2;\beta)e^{\delta};\theta)\Big\} \\ & \times\Pi_{t_1>t_2}\Big\{\partial_t Z(t_1;\alpha)\varphi(\mathbf{x}_1;\beta)e^{\delta}\partial_t Z(t_2;\alpha)\varphi(\mathbf{x}_2;\beta) \\ & \qquad \times g(Z(t_1;\alpha)\varphi(\mathbf{x}_1;\beta)e^{\delta}, Z(t_2;\alpha)\varphi(\mathbf{x}_2;\beta);\theta)\Big\} \\ & \times\Pi_{t_1=t_2}\Big\{1-\int_0^{+\infty}\int_{\xi_1\varphi(\mathbf{x}_2;\beta)e^{\delta}/\varphi(\mathbf{x}_1;\beta)}^{+\infty} g(\xi_1,\xi_2;\theta)d\xi_2 d\xi_1 \\ & \qquad -\int_0^{+\infty}\int_{\xi_2\varphi(\mathbf{x}_1;\beta)e^{\delta}/\varphi(\mathbf{x}_2;\beta)}^{+\infty} g(\xi_1,\xi_2;\theta)d\xi_1 d\xi_2\Big\}\end{aligned}$$

where $\Pi_{t_1<t_2}, \Pi_{t_1>t_2}$ and $\Pi_{t_1=t_2}$ denote the product over the observations for which $t_1 < t_2$, $t_1 > t_2$ and $t_1 = t_2$. We use the fact that, for sequential switching ($t_1 < t_2$ or $t_1 > t_2$), there is a unique equilibrium so we know the contribution to the likelihood. For the event in which termination times coincide, we cannot map the duration to a unique (K_1, K_2) and we therefore ignore the exact duration and the contribution to the likelihood function is $\mathbb{P}(T_1 = T_2|\mathbf{x}_1, \mathbf{x}_2)$. Under standard assumptions, this likelihood function provides us with an estimator for the parameters of interest in this model. We conjecture that a sieves approach, for instance, may be adapted to obtain a more general estimation procedure.[7]

The probability in (10) can also be used to obtain an estimator for $\varphi(\cdot;\beta)$ and δ without the assumption that $Z(\cdot)$ is the same across games as long as it is the same for players within the same game. Assume initially that $G(\cdot,\cdot)$ is the bivariate CDF for two independent unit exponential random variables: $G(k_1,k_2) = (1-e^{-k_1})(1-e^{-k_2})\mathbf{1}_{(k_1,k_2)\in\mathbb{R}_+^2}$. Then

[7] In general, we expect a non–parametric estimator to converge at a slower rate than $\sqrt{N}$ as is the case for unrestricted non–parametric estimators in the duration literature (see- for instance- the discussion in Heckman and Taber (1994)).

$$
\begin{aligned}
\mathbb{P}(T_i < T_j|\mathbf{x}_1,\mathbf{x}_2) &= \mathbb{P}(Z^{-1}(K_i/\varphi(\mathbf{x}_i)) < Z^{-1}(K_j e^{-\delta}/\varphi(\mathbf{x}_j))|\mathbf{x}_1,\mathbf{x}_2) \\
&= \mathbb{P}(K_i\varphi(\mathbf{x}_j)e^{\delta}/\varphi(\mathbf{x}_i) < K_j|\mathbf{x}_1,\mathbf{x}_2) \\
&= \int_0^\infty e^{-k_i}\int_{k_i\varphi(\mathbf{x}_j)e^{\delta}/\varphi(\mathbf{x}_i)}^\infty e^{-k_j}dk_j dk_i \\
&= \int_0^\infty e^{-k_i-k_j\varphi(\mathbf{x}_j)e^{\delta}/\varphi(\mathbf{x}_i)}dk_1 \\
&= \frac{1}{1+\varphi(\mathbf{x}_j)e^{\delta}/\varphi(\mathbf{x}_i)} = \frac{e^{\log\varphi(\mathbf{x}_i)-\log\varphi(\mathbf{x}_j)-\delta}}{1+e^{\log\varphi(\mathbf{x}_i)-\log\varphi(\mathbf{x}_j)-\delta}}
\end{aligned}
$$

Taking $\varphi(\mathbf{x};\beta) = \exp(\mathbf{x}'\beta)$, for example, this becomes $\Lambda\left((\mathbf{x}_i - \mathbf{x}_j)'\beta - \delta\right)$. where $\Lambda(\cdot)$ is the CDF for the logistic distribution.

If we then define the variable Y by

$$
Y = \begin{cases} 1 & \text{if} \quad T_1 < T_2 \\ 2 & \text{if} \quad T_1 = T_2 \\ 3 & \text{if} \quad T_1 > T_2 \end{cases},
$$

then

$$
\begin{aligned}
\mathbb{P}(Y \le 1|\mathbf{x}_1,\mathbf{x}_2) &= \Lambda((\mathbf{x}_1 - \mathbf{x}_2)\beta - \delta), \\
\mathbb{P}(Y \le 2|\mathbf{x}_1,\mathbf{x}_2) &= \Lambda((\mathbf{x}_1 - \mathbf{x}_2)\beta + \delta).
\end{aligned}
$$

This corresponds to an ordered logit on Y with explanatory variables $\mathbf{x}_1 - \mathbf{x}_2$ and cutoff points at $-\delta$ and δ. If we take $G(\cdot,\cdot)$ to be the bivariate log-normal CDF, an ordered probit is obtained.

When $G(\cdot,\cdot)$ is unknown, but the same across games

$$
\begin{aligned}
\mathbb{P}(Y \le 1|\mathbf{x}_1,\mathbf{x}_2) &= H((\mathbf{x}_1 - \mathbf{x}_2)\beta - \delta) \\
\mathbb{P}(Y \le 2|\mathbf{x}_1,\mathbf{x}_2) &= H((\mathbf{x}_1 - \mathbf{x}_2)\beta + \delta)
\end{aligned} \tag{11}
$$

where $H(w) = \mathbb{P}(\ln K_1 - \ln K_2 \le w)$. Various authors have proposed alternative estimation procedures for the estimation of this semiparametric ordered choice model (for instance, Chen and Khan (2003), Coppejans (2007), Klein and Sherman (2002), Lee (1992), and Lewbel (2003)). If G is game-specific, then (11) can be estimated by a version of Manski's maximum score estimator (Manski (1975)).[8]

[8] This would require a quantile restriction on $K_1 - K_2$ conditional on $(\mathbf{x}_1,\mathbf{x}_2)$

Finally, we note that if $G(\cdot)$, and hence $H(\cdot)$, is known, δ is identified even if $\mathbf{x}_1 = \mathbf{x}_2$, since

$$\delta = -H^{-1}(\mathbb{P}(T_1 < T_2|\mathbf{x})).$$

6 The Effect of Misspecifications

In this section we briefly examine the effect of misspecifications in the economic model or equilibrium selection process on the estimation of the parameters of interest. Throughout K_1 and K_2 are assumed to be independent unit exponentials.

6.1 Ignoring Endogeneity

This subsection investigates the consequences of treating an opponent's decision as exogenous in a parametric version of our model. The first data–generating process is defined by $Z(t) = t^{\alpha}$, $\varphi(\mathbf{x}_i) = \exp(\beta_0 + \beta_1 \mathbf{x}_i)$, $(\alpha, \beta_0, \beta_1, \delta) = (1.0, -3.0, 0.3, 0.3)$ and

$$\begin{pmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \end{pmatrix} \sim N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \right).$$

This implies that without the interaction, T_1 and T_2 would be independent durations from a Weibull proportional hazards model. When the model gives rise to multiple equilibria (and hence simultaneous exit), a specific duration is drawn from a uniform distribution over the possible duration times.[9] Tables 1 and 2 present the results based on 1000 replications of data sets of size 1000. Table 1 is based on a correctly specified likelihood that groups all ties occurring in realizations of region 2 in the previous discussion of the model. Table 2 presents results from a maximum likelihood estimation for agent 1 taking agent 2's action as exogenous.

[9]We experimented with different selection rules and these made no appreciable difference to the results we present here.

TABLE 1: Incorporating Endogeneity

	True Value	Bias	RMSE	Median Bias	Median Abs.Err.
α	1.000	0.001	0.019	0.000	0.013
β_0	-3.000	0.000	0.067	-0.001	0.045
β_1	0.300	0.000	0.018	0.000	0.012
δ	0.300	-0.001	0.023	-0.001	0.016

TABLE 2: Weibull. Dependent variable T_1

	True Value	Bias	RMSE	Median Bias	Median Abs.Err.
α	1.000	-0.079	0.084	-0.080	0.080
β_0	-3.000	0.076	0.116	0.078	0.087
β_1	0.300	-0.005	0.027	-0.005	0.019
δ	0.300	0.523	0.530	0.524	0.524

As expected, the maximum-likelihood estimator that incorporates endogeneity performs well, whereas the Weibull estimator that assumes that the other agent's action is exogenous performs poorly. Specifically, the effect of the opponent's decision is grossly over-estimated. Treating the other agent's action as exogenous also biases estimates toward negative duration dependence. Both of these are expected. In the first case, δ is biased because the estimation does not take into account the multiplier effect caused by the feedback between T_1 and T_2. The assumption of exogeneity also leads to a downward bias on duration dependence as duration lengths reinforce themselves: a shock leading to a longer duration by one agent will tend to lengthen the opponent's duration and hence further reduce the hazard for the original agent. Likewise, some bias is found in the estimation of β_1: changing $\mathbf{x}_i$ leads to a change in T_i, which affects T_j and feeds back into T_i. Ignoring this channel also introduces bias.

The results in Tables 1 and 2 assume symmetry between the two agents in the model. The next design changes this by changing the joint distribution of $(\mathbf{x}_1, \mathbf{x}_2)$ to

$$\begin{pmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \end{pmatrix} \sim N\left(\begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}\right)$$

This makes the first agent likely to move first. When multiple equilibria were possible, an equilibrium was selected as in the previous exercise. The overestimation bias on δ is of a similar magnitude as before. The effect on the estimation of α is different for each individual given the asymmetry in the distribution of the $\mathbf{x}$'s.

TABLE 3: Incorporating Endogeneity

	True Value	Bias	RMSE	median bias	median abs.err.
α	1.000	0.000	0.019	0.000	0.012
β_0	-3.000	0.000	0.067	0.000	0.045
β_1	0.300	0.000	0.017	0.000	0.011
δ	0.300	0.000	0.024	0.000	0.017

TABLE 4: Weibull. Dependent variable T_1

	True Value	Bias	RMSE	median bias	median abs.err.
α	1.000	-0.065	0.071	-0.066	0.066
β_0	-3.000	0.049	0.107	0.052	0.075
β_1	0.300	-0.002	0.026	-0.002	0.018
δ	0.300	0.523	0.530	0.524	0.524

TABLE 5: Weibull. Dependent variable T_2

	True Value	Bias	RMSE	median bias	median abs.err.
α	1.000	-0.095	0.099	-0.095	0.095
β_0	-3.000	0.083	0.121	0.083	0.087
β_1	0.300	-0.007	0.027	-0.008	0.018
δ	0.300	0.530	0.537	0.531	0.531

6.2 Equilibrium Selection

In this subsection, we examine the effect of estimating the model by full maximum likelihood after imposing a potentially incorrect equilibrium selection assumptions in the estimation of an otherwise correctly specified, parametric version of the model.

The data-generating processes for all the results below are based on $Z(t) = t^{\alpha}$, $\varphi(\mathbf{x}_i) = \exp(\beta_0 + \beta_1 \mathbf{x}_{1i} + \beta_2 \mathbf{x}_2)$ and $(\alpha, \beta_0, \beta_1, \beta_2, \delta) = (1.35, -4.00, 1.00, 0.50, 1.00)$, where $\mathbf{x}_{i1}$, $i = 1, 2$ represents an individual specific covariate and $\mathbf{x}_2$, a common covariate. These three variables are independent standard normal random variables. A total of 1000 replications with sample sizes of 2000 observations (games) were generated.

Tables 6 through 10 differ in the way equilibrium is selected when there are multiple equilibria. Aside from the column indicating the value of each of the parameters, each of the tables presents median bias and median absolute error for three alternative estimators: the maximum likelihood estimator from Section 5 that pools equilibria without selecting the equilibrium; a maximum likelihood estimator that assumes the earliest equilibrium ($\underline{T}$) is played when there are multiple equilibria; and a maximum likelihood estimator that takes the latest equilibrium ($\overline{T}$) as the selected equilibrium in case of multiple equilibria.

In Table 6, the latest equilibrium ($\overline{T}$) is selected. As expected, the estimator corresponding to the results in the last two columns performs the best, since it assumes the correct selection rule generating the data. Pooling equilibria in the estimation seems to do an appreciably better job than the estimator that incorrectly assumes the equilibrium selection criterion as the earliest possible equilibrium: although the estimates for β_1 and δ present similar median bias and absolute error, the other parameters appear to present much less bias in the estimator that pools the equilibria. The estimator for the constant term β_0 seems to be particularly biased downward when $\underline{T}$ is assumed to be selected. This makes sense: by assuming an earlier selection scheme the constant is below the true parameter, lowering the hazard and thus increasing the durations to match the data.

TABLE 6: $\overline{T}$ Selected

		Pools Ties		*Assumes $\underline{T}$*		*Assumes $\overline{T}$*	
	True Value	Median Bias	Median Absolute	Median Bias	Median Absolute	Median Bias	Median Absolute
α	1.350	0.018	0.053	−0.025	0.046	0.011	0.041
Constant	−4.000	−0.036	0.160	−0.168	0.189	−0.028	0.129
δ	1.000	−0.003	0.060	−0.001	0.059	0.001	0.054
β_1	1.000	0.014	0.059	−0.015	0.052	0.005	0.046
β_2	0.500	0.006	0.043	−0.033	0.043	0.006	0.038

Table 7 displays a design where the earliest equilibrium ($\underline{T}$) is picked. Here the middle estimator, which correctly assumes the selection scheme generating the data, is as expected the best of the three. The improvement of the pooling estimator over the one that wrongfully assumes the selection mechanism seems even more compelling than in the previous case. The effect of mistaken equilibrium selection on the constant term is again fairly large: in order to accommodate an equilibrium selection rule that chooses later equilibria than the ones actually played, the hazard are overestimated, which lowers the duration.

TABLE 7: $\underline{T}$ Selected

		Pools Ties		*Assumes $\underline{T}$*		*Assumes $\overline{T}$*	
	True Value	Median Bias	Median Absolute	Median Bias	Median Absolute	Median Bias	Median Absolute
α	1.350	0.007	0.049	0.008	0.040	−0.014	0.042
Constant	−4.000	−0.017	0.158	−0.012	0.125	0.321	0.321
δ	1.000	0.005	0.062	0.005	0.062	−0.137	0.137
β_1	1.000	0.006	0.058	0.007	0.046	−0.013	0.046
β_2	0.500	0.003	0.042	0.002	0.038	0.006	0.039

In Table 8, an equilibrium is randomly selected according to a uniform distribution on $[\underline{T}, \overline{T}]$, as was the case in the previous subsection. The performance of the pooling estimator is noticeably better in comparison to the two other estimators except for the estimation on α, the Weibull parameter.

TABLE 8: $U[\underline{T}, \overline{T}]$ Selected

		Pools Ties		*Assumes $\underline{T}$*		*Assumes $\overline{T}$*	
	True Value	Median Bias	Median Absolute	Median Bias	Median Absolute	Median Bias	Median Absolute
α	1.350	0.010	0.048	-0.001	0.041	0.006	0.040
Constant	-4.000	-0.025	0.152	-0.125	0.154	0.116	0.150
δ	1.000	0.005	0.062	0.008	0.060	-0.065	0.071
β_1	1.000	0.011	0.060	0.003	0.046	0.007	0.045
β_2	0.500	-0.002	0.044	-0.020	0.041	0.002	0.038

Table 9 shows the case in which the earliest equilibrium is selected when the common variable $\mathbf{x}_2$ is greater than zero, whereas the latest equilibrium is picked when $\mathbf{x}_2$ is less then zero – this amplifies the effect of this variable on the hazard beyond the impact already present in the multiplicative $\varphi(\cdot)$ term. In this case, the pooling estimator fares better across all the parameters.

TABLE 9: $\underline{T} \cdot \mathbf{1}(\mathbf{x}_2 > 0) + \overline{T} \cdot \mathbf{1}(\mathbf{x}_2 \leq 0)$ Selected

		Pools Ties		*Assumes $\underline{T}$*		*Assumes $\overline{T}$*	
	True Value	Median Bias	Median Absolute	Median Bias	Median Absolute	Median Bias	Median Absolute
α	1.350	0.009	0.051	-0.015	0.043	-0.007	0.042
Constant	-4.000	-0.032	0.154	-0.095	0.146	0.161	0.177
δ	1.000	0.002	0.057	0.005	0.058	-0.069	0.075
β_1	1.000	0.008	0.059	0.085	0.086	0.065	0.070
β_2	0.500	0.007	0.042	-0.016	0.040	0.006	0.037

Finally, Table 10 displays results for a selection mechanism that picks $\underline{T}$ when this quantity is greater than 10 and selects $\overline{T}$ when $\underline{T}$ is less than 10. Again the pooling estimator seems to be the superior one when comparing median bias and median absolute error for the parameters of interest.

TABLE 10: $\underline{T} \cdot \mathbf{1}(\underline{T} > 10) + \overline{T} \cdot \mathbf{1}(\underline{T} \leq 0)$ **Selected**

		Pools Ties		*Assumes* $\underline{T}$		*Assumes* $\overline{T}$	
	True Value	Median Bias	Median Absolute	Median Bias	Median Absolute	Median Bias	Median Absolute
α	1.350	0.014	0.048	0.057	0.059	0.051	0.056
Constant	−4.000	−0.030	0.143	−0.253	0.254	0.020	0.129
δ	1.000	0.009	0.067	−0.006	0.061	−0.091	0.095
β_1	1.000	0.012	0.061	−0.039	0.056	−0.024	0.048
β_2	0.500	0.001	0.042	−0.023	0.041	0.002	0.038

In sum, either ignoring the strategic interaction in the model by assuming exogeneity or misspecifying the equilibrium selection mechanism may lead to erroneous inference.

7 Conclusion

In this article we have provided a new motivation for simultaneous duration models that relies on strategic interactions between agents. The paper thus relates to the previous literature on empirical games. We presented an analysis of the possible Nash equilibria in the game and noticed that it displays multiple equilibria, but in a way that still permits point identification of structural objects.

The maintained assumption in the paper is that agents can exactly control their duration. Heckman and Borjas (1980), Honoré (1993) and Frijters (2002) consider statistical models in which the hazard for one duration depends on the outcome of a previous duration and Rosholm and Svarer (2001) consider a model in which the hazard for one duration depends on the simultaneous hazard for a different duration. It would be interesting to investigate whether a strategic economic model in which agents can control their hazard subject to costs will generate incomplete econometric models and what the effect of this would be on the identifiability of the key parameters of the model.

References

Abbring, J., and G. van den Berg (2003): "The Nonparametric Identification of Treatment Effects in Duration Models," *Econometrica*, 75, 933–64.

Amemiya, T. (1974): "Multivariate Regression and Simultaneous Equation Models when the Dependent Variables Are Truncated Normal," *Econometrica*, 42(6), 999–1012.

Bergin, J., and B. MacLeod (1993): "Continuous Time Repeated Games," *International Economic Review*, 34.

Berry, S., and E. Tamer (2006): "Identification in Models of Oligopoly Entry," in *Advances in Economics and Econometrics*, ed. by R. Blundell, W. Newey, and T. Persson, vol. 2. Cambridge University Press, Cambridge.

Bresnahan, T., and P. Reiss (1991): "Empirical Models of Discrete Games," *Journal of Econometrics*, 48.

Chen, S., and S. Khan (2003): "Rates of Convergence for Estimating Regression Coefficients in Heteroskedastic Discrete Response Models," *Journal of Econometrics*, 117.

Coppejans, M. (2007): "On Efficient Estimation of the Ordered Response Model," *Journal of Econometrics*, 137.

Cox, D., and D. Oakes (1984): *The Analysis of Survival Data*. Chapman and Hall.

Elbers, C., and G. Ridder (1982): "True and Spurious Duration Dependence: The Identifiability of the Proportional Hazard Model," *Review of Economic Studies*, 49.

Frederiksen, A., B. E. Honoré, and L. Hu (2007): "Discrete Time Duration Models with Group–level Heterogeneity," *Journal of Econometrics*, 141, 1014–1043.

Frijters, P. (2002): "The Non-Parametric Identification of Lagged Duration Dependence," *Economics Letters*, 75(3), 289–292.

Fudenberg, D., and J. Tirole (1985): "Preemption and Rent Equalization in the Adoption of New Technology," *Review of Economic Studies*, 52(3).

——— (1991): *Game Theory*. The MIT Press.

Hahn, J. (1994): "The Efficiency Bound for the Mixed Proportional Hazard Model," *Review of Economic Studies*, 61(4).

Hausman, J., and T. Woutersen (2006): "Estimating a Semi-Parametric Duration Model with Heterogeneity and Time-Varying Regressors," MIT Working Paper.

Heckman, J. (1978): "Dummy Endogenous Variables in a Simultaneous Equation System," *Econometrica*, 46.

Heckman, J., and B. Singer (1984): "A Method for Minimizing the Impact of Distributional Assumptions in Econometric Models for Duration Data," *Econometrica*, 52(2).

Heckman, J., and C. Taber (1994): "Econometric Mixture Models and More General Models for Unobservables in Duration Analysis," *Statistical Methods in Medical Research*, 3(3), 277–99.

Heckman, J. J., and G. J. Borjas (1980): "Does Unemployment Cause Future Unemployment? Definitions, Questions and Answers from a Continuous Time Model of Heterogeneity and State Dependence," *Economica*, 47, 247–283.

Heckman, J. J., and B. E. Honoré (1989): "The Identifiability of the Competing Risks Model," *Biometrika*, 76, 325–330.

Honoré, B. (1990): "Simple Estimation of a Duration Model with Unobserved Heterogeneity," *Econometrica*, 58(2).

Honoré, B. E. (1993): "Identification Results for Duration Models with Multiple Spells," *Review of Economic Studies*, 60(1), 241–46.

Horowitz, J. L., and S. Lee (2004): "Semiparametric Estimation of a Panel Data Proportional Hazards Model with Fixed Effects," *Journal of Econometrics*, 119(1), 155–198.

Hougaard, P. (2000): *Analysis of Multivariate Survival Data*. Springer–Verlag, New York.

Klein, J., N. Keiding, and C. Kamby (1989): "Semiparametric Marshall-Olkin Models Applied to the Occurrence of Metastases at Multiple Sites after Breast Cancer," *Biometrics*, 45.

Klein, R., and R. Sherman (2002): "Shift Restrictions and Semiparametric Estimation in Ordered Response Models," *Econometrica*, 70.

Lancaster, T. (1985): "Simultaneous equations models in applied search theory," *Journal of Econometrics*, 28(1), 113 – 126.

Lee, M. (1992): "Median Regression for Ordered Discrete Response," *Journal of Econometrics*, 51.

Lee, S. S. (2003): "Estimating Panel Data Duration Models with Censored Data," Cemmap working papers, Centre for Microdata Methods and Practice, Institute for Fiscal Studies.

Lewbel, A. (2003): "Ordered Response Threshold Estimation," Boston College Working Paper.

Manski, C. F. (1975): "The Maximum Score Estimation of the Stochastic Utility Model of Choice," *Journal of Econometrics*, 3, 205–228.

Marshall, A., and I. Olkin (1967): "A Multivariate Exponential Distribution," *Journal of the American Statistical Association*, 62.

Park, A., and L. Smith (2006): "Caller Number Five: Timing Games that Morph from One Form to Another," University of Toronto Working Paper.

Paula, A. (2009): "Inference in a Synchronization Game with Social Interactions," *Journal of Econometrics*, 148(1), 56–71.

Ridder, G. (1990): "The Non-Parametric Identification of Generalized Accelerated Failure-Time Models," *Review of Economic Studies*, 57.

Ridder, G., and T. Woutersen (2003): "The Singularity of the Information Matrix of the Mixed Proportional Hazard Model," *Econometrica*, 71(5).

Rosholm, M., and M. Svarer (2001): "Structurally Dependent Competing Risks," *Economics Letters*, 73(2), 169–173.

Simon, L., and M. Stinchcombe (1989): "Extensive Form Games in Continuous Time: Pure Strategies," *Econometrica*, 57.

Tamer, E. (2003): "Incomplete Simultaneous Discrete Response Model with Multiple Equilibria," *Review of Economic Studies*, 70.

van den Berg, G. J. (2001): "Duration Models: Specification, Identification and Multiple Durations," in *Handbook of Econometrics*, ed. by J. Heckman, and E. Leamer, vol. 5 of *Handbook of Econometrics*, chap. 55, pp. 3381–3460. Elsevier.

Appendix

We present a proof for identification of $Z(\cdot)$ that dispenses with the assumption that $\mathbf{x}_i$ contains a continuously distributed covariate as in Theorem 3. Specifically assume that $\mathbf{x}_i$ takes two values, a and b. By Theorem 1, $\varphi(\cdot)$ is identified up to scale. Normalize $\varphi(a) = 1$ and $\varphi(b) < 1$. The proof parallels that in Elbers and Ridder (1982). Consider the function:

$$B(s) = \int_0^s \int_{Z(t)e^{\delta}\varphi(\mathbf{x}_2)}^{\infty} g(\xi_1, \xi_2) d\xi_1 d\xi_2, \quad \text{for all } t \geq 0.$$

which is implicitly also a function of δ, $g(\cdot)$, $Z(\cdot)$ and $\varphi(\mathbf{x}_2)$. When evaluated at $Z(t)\varphi(\mathbf{x}_1)$ this function provides the probability that agent 1 leaves before t and agent 2 leaves after t. This function is increasing and, consequently, invertible (holding fixed the other implicit arguments).

Assume that $Z(\cdot)$ is not identified. Then, there is a pair $(\tilde{Z}, \tilde{B})$ such that

$$\begin{aligned} B(Z(t)) &= \tilde{B}(\tilde{Z}(t)), \quad \text{for all } t \geq 0 & (12) \\ B(Z(t)\varphi(b)) &= \tilde{B}(\tilde{Z}(t)\varphi(b)), \quad \text{for all } t \geq 0. & (13) \end{aligned}$$

From equation (12),

$$\varphi(b)\,\tilde{Z}(t) = \varphi(b)\,\tilde{B}^{-1}(B(Z(t))), \quad \text{for all } t \geq 0$$

and from equation (13),

$$\tilde{Z}(t)\varphi(b) = \tilde{B}^{-1}\left(B(Z(t)\varphi(b))\right), \quad \text{for all } t \geq 0$$

and, consequently,

$$\tilde{B}^{-1}(B(Z(t)\varphi(b))) = \varphi(b)\,\tilde{B}^{-1}(B(Z(t))), \quad \text{for all } t \geq 0. \tag{14}$$

Defining $f = \tilde{B}^{-1} \circ B$ we have from equation (14) that

$$f(\varphi(b)\,s) = \varphi(b)\,f(s), \quad \text{for all } s \geq 0 \tag{15}$$

and consequently that $f(0) = 0$. Proceeding as in Elbers and Ridder (1982), this implies that

$$f(\varphi(b)^n\,s) = \varphi(b)^n\,f(s), \quad \text{for all } s \geq 0 \text{ and all } n.$$

after repeated application of (15). Differentiating with respect to s and rearranging:

$$f'(s) = f'(\varphi(b)^n\,s), \quad \text{for all } s \geq 0 \text{ and all } n.$$

Since $\varphi(b) < 1$, taking the limit as $n \to \infty$,

$$f'(s) = f'(0) \equiv c$$

which, along with $f(0) = 0$, implies that

$$\tilde{B}^{-1} \circ B(s) = cs, \quad \text{for all } s$$

establishing that $\tilde{B}(cs) = B(s)$, for all s. Using equation (12) we obtain that $\tilde{B}(cZ(t)) = \tilde{B}(\tilde{Z}(t)) \Rightarrow cZ(t) = \tilde{Z}(t)$ for all t.