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 09-037_**

"Identification of Stochastic Sequential Bargaining Models"

by

Antonio Merlo and Xun Tang

http://ssrn.com/abstract=1492217

# Identification of Stochastic Sequential Bargaining Models[1]

Antonio Merlo

University of Pennsylvania

Xun Tang

University of Pennsylvania

October 15, 2009

ABSTRACT

Stochastic sequential bargaining games (Merlo and Wilson (1995, 1998)) have found wide applications in various fields including political economy and macroeconomics due to their flexibility in explaining delays in reaching agreement. In this paper, we present new results in nonparametric identification of such models under different scenarios of data availability. First, with complete data on players' decisions, the sizes of the surplus to be shared (cakes) and the agreed allocations, both the mapping from states to the total surplus (i.e. the "cake function") and the players' common discount rate are identified, if the unobservable state variable (USV) is independent of observable ones (OSV), and the total surplus is strictly increasing in the USV conditional on the OSV. Second, when the cake size is only observed under agreements and is additively separable in OSV and USV, the contribution by OSV is identified provided the USV distribution satisfies some distributional exclusion restrictions. Third, if data only report when an agreement is reached but never report the cake sizes, we propose a simple algorithm that exploits exogenously given shape restrictions on the cake function and the independence of USV from OSV to recover *all* rationalizable probabilities for reaching an agreement under counterfactual state transitions. Numerical examples show the set of rationalizable counterfactual outcomes so recovered can be informative.

Key words: Nonparametric identification, non-cooperative bargaining, stochastic sequential bargaining, rationalizable counterfactual outcomes

JEL codes: C14, C35, C73, C78

[1] We thank Herman Bierens, Hanming Fang, Philip Haile, Isabelle Perrigne, Katja Seim, Elie Tamer, Petra Todd, Quang Vuong, Ken Wolpin and seminar participants at Penn State and Wharton School at U Penn for comments and discussions. The usual disclaimer applies.

# 1 Introduction

Starting with the seminal contributions of Stahl (1972) and Rubinstein (1982), noncooperative (or strategic) bargaining theory has flourished in the last thirty years. The original model of bilateral bargaining with alternating offers and complete information has been extended in a number of directions allowing for more general extensive forms, information structure and more than two players (see, e.g., Osborne and Rubinstein (1990), Binmore, Osborne and Rubinstein (1992) for surveys). The development of the theoretical literature has gone hand in hand with, and for a large part has been motivated by, the broad range of applications of bargaining models. These include labor, family, legal, housing, political, and international negotiations (see, e.g., Muthoo (1999)). The increased availability of data on the outcomes of such negotiations as well as on the details of the bargaining process has also stimulated a surge in empirical work, where casual empiricism has progressively lead the way to more systematic attempts to take strategic bargaining models to data.[2]

A theoretical framework that has been extensively used in empirical applications is the stochastic bargaining model proposed by Merlo and Wilson (1995, 1998). In this model, the surplus to be allocated (or the "cake") and the bargaining protocol (i.e., the order in which players can make offers and counteroffers), are allowed to evolve over time according to a stochastic process. This feature makes the model flexible (it provides a unified framework for a large class of bargaining games), and rationalizes the occurrence of delays in reaching agreement, which are often observed in actual negotiations, in bargaining environments with complete information. Moreover, for the case where players share a common discount factor and their utility is linear in the amount of surplus they receive (which we refer to as the "canonical model"), the game has a unique subgame perfect equilibrium when there are only two players bargaining, and a unique stationary subgame perfect equilibrium (SSPE) when negotiations are multilateral. The unique equilibrium admitted by the model is stochastic and characterized by the solution of a fixed-point problem which can be easily computed. For all these reasons, the stochastic bargaining framework naturally lends itself to estimation and has been used in a variety of empirical applications that range from the formation of coalition governments in parliamentary democracy (Merlo (1997), Diermeier, Eraslan and Merlo (2003)), to collective bargaining agreements (Diaz-Moreno and Galdon (2000)), to corporate bankruptcy reorganizations (Eraslan (2008)), to the setting of industry standards in product markets (Simcoe (2008)), and to sovereign debt renegotiations (Benjamin and Wright (2008), Bi (2008)).

To date, the existing literature on the structural estimation of noncooperative bargaining models has been entirely parametric. In addition to the body of work cited above based on the stochastic framework, other bargaining models have also been specified and parametrically estimated using a

[2] This path shares many similarities with the development of auction theory in industrial organization.

variety of data sets.[3] However, little is known about whether the structural elements of these models or the bargaining outcomes in a counterfactual environment can be identified without imposing parametric assumptions. This paper contributes to the literature on the estimation of sequential bargaining models by providing positive results in the nonparametric identification of stochastic bargaining models. Our work is not intended to advocate the complete removal of parametric assumptions on the primitives of these models in structural estimations, as in most cases such assumptions are instrumental for attaining point-identification and can be tested. Rather, our main objective is to understand the limit of what can be learned about the model structure and rationalizable counterfactual outcomes when researchers wish to remain agnostic about unknown structural elements of the model.[4]

Empirical contexts of stochastic bargaining games may differ in what the econometricians observe in the data. These differences will in general have important implications on identification of the model structures. Here, we consider three scenarios with increasing data limitations. We refer to these scenarios as "complete data" (where econometricians observe the total surplus to be allocated or "the size of the cake" in each period regardless of whether an agreement is reached), "incomplete data with censored cake sizes" (where econometricians only observe the size of the cake in the period when an agreement is reached), and "incomplete data with unobservable cake sizes" (where econometricians only observe the timing of agreement, but never observe the value of the surplus). To illustrate the three data scenarios and introduce some useful notation, consider, for example, a situation where a group of investors decide to dissolve their partnership and bargain over how to divide a portfolio they jointly own. The size of the cake is the market value of the portfolio which is determined by state variables, such as market or macroeconomic conditions, that evolve over time according to a stochastic process. The investors share the same discount factor which is the market interest rate. Certain state variables that affect the market value of the portfolio are observed by both the investors and the econometricians (OSV), while other state variables are only known to the investors but not observed by the econometricians (USV). In the complete data scenario, the econometricians observe the evolution of the market value of the portfolio at all dates throughout the negotiation. This situation would arise for example if the portfolio is entirely composed of publicly traded stocks. In the second scenario, the econometricians only observe the

[3]For example, Sieg (2000) and Watanabe (2006) estimate a bargaining model with asymmetric information or with uncommon priors, respectively, to study the timing and terms of medical malpractice dispute resolutions. Merlo, Ortalo-Magne and Rust (2009) estimate a bargaining model with incomplete information to study the timing and terms of residential real estate transactions.

[4]In this respect, our work is related to the growing literature on nonparametric identification and tests of empirical auction models, pioneered by Laffont and Vuong (1996), Guerre, Perrigne and Vuong (2000), Athey and Haile (2002), Haile and Tamer (2003), Haile, Hong and Shum (2004), Hendricks, Pinkse and Porter (2003). A recent paper by Chiappori and Donni (2006) also addresses related questions in the context of a static, cooperative (or axiomatic) bargaining framework and derives sufficient conditions on the auxiliary assumptions of the model under which the Nash bargaining solution generates testable restrictions. We do not review the (theoretical or empirical) literature on cooperative bargaining here since it is outside of the scope of this paper.

market value of the portfolio when an agreement is reached but not in any other period during the negotiation. This would be the case if for example the portfolio is composed of non-publicly traded securities, but the sale price is recorded. Finally, in the third scenario with the least data, the econometricians only observe the timing of agreements but never observe the market value of the portfolio. This would be the case if for example the only available information is when a partnership is dissolved but the details of the settlement are kept confidential (e.g., because of a court order).

Under each of the three data scenarios described above, we provide conditions that attain positive results in the identification of model structures or counterfactuals for the canonical model of stochastic bargaining. With complete data, we show that the common discount factor can be uniquely recovered from the distribution of cake sizes and the occurrence of agreements, if the USV and the order of moves are independent of history conditional on the current OSV. Furthermore, if given any state observed the total surplus to be allocated is strictly increasing in the USV, then the cake function (which maps from states into the total surplus) is identified from the distribution of cake sizes, provided (i) the USV is independent of OSV and (ii) the bargaining protocol is independent of the cake sizes given observable states. In the second scenario with censored cake sizes, we show that when the cake function is additively separable in OSV and USV, it is nonparametrically identified if the USV distribution is conditionally independent of a subvector of OSV, or has multiplicative heterogeneity.

In the third scenario with unobservable cake sizes, we first note that earlier, known results for identifying optimal stopping problems apply to our setting. In particular, Berry and Tamer (2006) showed an additively separable cake function is identified if the USV is conditionally independent of all past states given any contemporary OSV, and if the USV distribution is known to econometricians. However, our main contribution under this data scenario is to relax the unrealistic assumption of known USV distribution, and show partial identification of counterfactual outcomes (i.e. probability for reaching an agreement conditional on the OSV) is possible under nonparametric shape restrictions of the cake function and independence of USV from OSV. Our approach is motivated by the fact that in practice researchers often know the cake function satisfies certain shape restrictions derived exogenously from economic theory, or common senses. For example, expected market value of a portfolio of foreign assets must be monotone in exchange rates holding other state variables fixed. We argue such knowledge can be exploited to at least confine rationalizable counterfactual outcomes to an informative subset of the complete outcome space, with the aid of stochastic restrictions such as independence of USV from OSV. To our knowledge, this is the first positive result in identifying counterfactuals in a structural optimal stopping model without assuming knowledge of the USV distribution. We propose a simple but novel algorithm to recover the complete set of rationalizable counterfactual outcomes (RCO), which are defined as outcomes in a counterfactual context that are consistent with players' dynamic rationality, the shape and stochastic restrictions of the model, as well as the actual outcomes observed in the data. We use numerical examples to show the set of RCO recovered can be quite informative and small relative to the complete outcome space.

We also address the identification of two extensions of the canonical model of stochastic bargaining where the players evaluate the cake according to a concave utility function, or the discount factors are heterogeneous across players.[5] We show that if players across all bargaining games in the data are known to follow strategies that lead to the same profiles of expected stationary subgame perfect equilibrium (SSPE) payoffs, then heterogenous discount rates and utility functions can both be identified in the case with complete data under fairly weak restrictions on players' risk attitudes.[6]

The remainder of the paper is organized as follows. Section 2 introduces the canonical model of stochastic sequential bargaining and characterizes players' payoffs in stationary subgame perfect equilibria. Sections 3, 4 and 5 present identification results in each of the three scenarios with different data availability. Section 6 studies identification in extensions of the canonical model with concave utility functions or heterogenous discount rates. Section 7 concludes. Proofs are included in the appendix.

## 2 The Canonical Model of Stochastic Bargaining

Consider an infinite-horizon bargaining game with $K \geq 2$ players (denoted as $i = 1,.,K$) who share the same discount factor $\beta \in (0,1)$. In each period (indexed by $t$), all players observe a vector of states $S_t$ with support $\Omega_S \subseteq \mathbb{R}^{D_S}$ where $D_S$ denotes the dimension of $S$. (Throughout the paper, we use upper case letters for random variables and lower case letters for their realizations. We use $\Omega_R$ to denote the support of a generic random vector $R$, and $R^t$ to denote its history up to, and including, period $t$, i.e. $R^t \equiv \{R_1, R_2, ., R_t\}$.) The set of feasible utility vectors to be allocated in period $t$ with realized state $s_t$ is given by $C(s_t) = \{u \in \mathbb{R}^K : \sum_{i=1}^K u_i \leq c(s_t)\}$, where $c(s_t) : \Omega_S \to \mathbb{R}^1_+$ is the "cake function".[7] In each period $t$, the order of moves among players is

[5] In the terminology of Merlo and Wilson (1995, 1998), these are stochastic bargaining games with non-transferable utility, which typically have multiple equilibria.

[6] In either of the two cases, the profiles of SSPE payoffs for players are not unique in general. The single-SSPE-payoff assumption above is analogous to the "single-equilibrium" assumption used in the estimation of simultaneous games with incomplete information. Such assumptions allow econometricians to link model structures to observable distributions using theoretical characterizations of Bayesina Nash equilibria (BNE) or (SSPE payoffs), while remaining agnostic about which BNE (or SSPE payoff) is realized in the data-generating process.

[7] This environment assumes that the players have time-separable quasi-linear von Neumann-Morgenstern utility functions over the commodity space and that a good with constant marginal utility to each player (e.g., money) can be freely transferred. In the terminology of Merlo and Wilson (1995,1998), this environment is defined as a stochastic bargaining model with transferable utility.

given by a permutation of $\{1, 2, ., K\}$, denoted as $\rho_t$, whose $i$-th coordinate $\rho_{t,(i)}$ is the identity of the player who makes the $i$-th move. Let $\Omega_\rho$ denote the set of all possible permutations of the $K$-vector. Let $\kappa_t \equiv \rho_{t,(1)}$ denote the proposer in period $t$. The transition of states and the order of moves satisfy the following restriction.

*CI-1 (Conditional independence of histories) Conditional on $S_t$, (i) $\rho_t$ is independent of past states and orders of moves $\{S^{t-1}, \rho^{t-1}\}$, and (ii) $S_{t+1}$ is independent of $\rho_t$ and history $\{S^{t-1}, \rho^{t-1}\}$.*

The assumption *CI-1* is a rather weak restriction that can be satisfied in lots of empirical contexts. For example, suppose the order of moves in each period $t$ is determined by a function $\rho(S_t, \zeta_t)$, where $\zeta_t$ consists of noises excluded from $S_t$ and unobservable both to players and to econometricians. Then conditions in *CI-1* are satisfied if (i) given $S_t$, $\zeta_t$ is independent of $\{S^{t-1}, \zeta^{t-1}\}$ and (ii) given $S_t$, $S_{t+1}$ is independent of $\zeta_t$ and $\{S^{t-1}, \zeta^{t-1}\}$. Condition in (i) does rule out the case where players take deterministic, alternating turns to make offers. Under *CI-1*, the transition between information variables is reduced to

$$\tilde{H}_t(S_{t+1}, \rho_{t+1}|S^t, \rho^t) = \tilde{L}_t(\rho_{t+1}|S_{t+1})H_t(S_{t+1}|S_t) \tag{1}$$

Throughout the paper, we maintain that both the first-order Markov transition between states $H_t(.|.)$ and the conditional multinomial distribution $\tilde{L}_t$ are time-homogenous. Thus we will drop the subscript $t$ from $(\tilde{L}, H)$, and use $(R, R')$ to denote random vectors in the current and the next period respectively when there is no confusion.

The game is played as follows. At the beginning of each period, players observe the realized states $s$ and the order of moves $\rho \equiv (\rho_{(1)}, ., \rho_{(K)})$ in that period. The proposer $\kappa \equiv \rho_{(1)}$ then chooses to either pass or propose an allocation in $C(s)$. If he proposes an allocation, player $\rho_{(2)}$ responds by either accepting or rejecting the proposal. Each player then responds in the order prescribed by $\rho$ until either some player rejects the offer or all players accept it. If no proposal is offered and accepted by all players, the game moves to the next period where a new state $s'$ and an order of moves $\rho'$ are realized according to the Markov process $\tilde{H}$. The procedure is then repeated except that the set of feasible proposals is given by $C(s')$ in the new period. This game continues until an allocation is proposed and accepted by all players (if ever). Parameters $(H, c, \tilde{L}, \beta)$ are common knowledge among all players but not known to econometricians. Let $\tilde{S}_t \equiv (S_t, \rho_t)$ denote the information revealed to players in period $t$, and let $\tilde{S}^t$ denote the history of information from the initial period 0 up to period $t$. Given any initial state $\tilde{S}_0 = (s, \rho)$, an outcome $(\tau, \eta_\tau)$ consists of a stopping time $\tau$ and a random $k$-vector $\eta_\tau$ that is measurable with respect to $\tilde{S}^\tau$ such that $\eta_\tau \equiv (\eta_{\tau,i})_{i=1}^K \in C(S_\tau)$ if $\tau < +\infty$ and $\eta_\tau = 0$ if $\tau = +\infty$. (Note the set of feasible allocations is independent of the order of moves.) Given a realization of $(\tilde{s}_0, \tilde{s}_1, \tilde{s}_2, ..)$ with $\tilde{s}_t \equiv (s_t, \rho_t)$, $\tau$ denotes the period in which a proposal is accepted, and $\eta_\tau$ denotes the proposed allocation which is accepted in state $s_\tau$ when the order of moves is $\rho_\tau$. For a game starting with state $s$ and order of moves $\rho$, an outcome $(\tau, \eta_\tau)$ implies a von Neumann-Morgenstern payoff to player $i$, i.e. $E[\beta^\tau \eta_{\tau,i}|\tilde{S}_0 = (s, \rho)]$.

A stationary outcome is such that $\exists$ a measurable subset $\tilde{S}(\mu) \subseteq \Omega_{\tilde{S}} \equiv \Omega_{S,\rho}$ and a measurable function $\mu : \tilde{S}(\mu) \to \mathbb{R}^k$ such that (i) $\tilde{S}_t \notin \tilde{S}(\mu)$ for all $t = 0, 1, ., \tau - 1$; (ii) $\tilde{S}_\tau \in \tilde{S}(\mu)$; and (iii) $\eta_\tau = \mu(\tilde{S}_\tau)$. That is, no allocation is implemented until some state and order of moves $(s, \rho) \in \tilde{S}(\mu)$ is realized, in which case a proposal $\mu(s, \rho) \in C(s)$ is accepted. Using property (iii), we let $v^\mu(s, \rho) \equiv E[\beta^\tau \mu(\tilde{S}_\tau)|\tilde{S}_0 = (s, \rho)]$ denote the von-Neumann-Morgenstern payoff vector given initial state and order of moves $(s, \rho)$. It follows from the definition of stationary outcome that $v^\mu(s, \rho) = \mu(s, \rho)$ for all $(s, \rho) \in \tilde{S}(\mu)$ and $v^\mu(s, \rho) = E[\beta^\tau \mu(\tilde{S}_\tau)|\tilde{S}_0 = (s, \rho)]$ for all $(s, \rho) \notin \tilde{S}(\mu)$. Alternatively we denote a stationary outcome by $(\tilde{S}(\mu), \mu, \tau)$. A history up to a period $t$ is a finite sequence of realized states, orders of moves, and the actions taken at each state in the sequence up to period $t$. A strategy for player $i$ specifies a feasible action at every history at which he must act. A strategy profile is a measurable $k$-tuple of strategies, one for each player. At any history, a strategy profile induces an outcome and hence a payoff for each player. A strategy profile is a subgame perfect equilibrium (SPE) if, at every history, it is a best response to itself. We refer to the outcome and payoff functions induced by a subgame perfect strategy profile as an SPE outcome and SPE payoff respectively. A strategy profile is stationary if the actions prescribed at any history depend only on the current state and current offer. A stationary SPE (SSPE) outcome and payoff are the outcome and payoff generated by a subgame perfect strategy profile which is stationary. Let $v_i : \Omega_{S,\rho} \to \mathbb{R}^1_+$ denote SSPE payoffs for player $i = 1, ., K$, and $w = \sum_{i=1}^K v_i$ denote total SSPE payoffs of all players in the bargaining games. Let $F^K$ denote the set of bounded measurable functions on $\Omega_{S,\rho}$ taking values in $\mathbb{R}^K$. Lemma 1 collects main results characterizing agents' behaviors and outcomes in SSPE of the bargaining game.

**Lemma 1** *Suppose CI-1 holds. Then (a) $f \in F^K$ is a unique SSPE payoff if and only if $A(f) = f$ where for all $(s, \rho) \in \Omega_{S,\rho}$,*

$$\begin{aligned} A_i(f)(s, \rho) &\equiv \max\{c(s) - \beta E[\textstyle\sum_{j \neq i} f_j(S', \rho')|S = s] \ , \ \beta E[f_i(S', \rho')|S = s]\}, \text{ if } \rho_{(1)} = i \\ A_j(f)(s, \rho) &\equiv \beta E[f_j(S', \rho')|S = s], \text{ if } \rho_{(1)} \neq j \end{aligned}$$

*(b) the SSPE total payoff $w$ is independent of $\rho$ given $s$, and solves*

$$w(s) = \max\{c(s), \beta E[w(S')|S = s]\}$$

*for all $s \in \Omega_S$; (c) A unanimous agreement is reached in state $s$ if and only if $c(s) \geq \beta E[w(S')|S = s]$.*

The proof uses results in Theorems 1-3 in Merlo and Wilson (1998) and exploits conditions in *CI-1* to show that the total payoff in SSPE and the occurrence of agreement are independent of the order of moves conditional on realized states. This important property of SSPE is instrumental for some of the positive identification results below. Econometricians are interested in recovering the parameters $(H, c, \tilde{L}, \beta)$ underlying the bargaining game using the distribution of states and decisions of offers/acceptances observed. The data report players' proposals and decisions in a large number of bargaining games. In each period, the state variable $S$ consists of $X \in \Omega_X \subseteq \mathbb{R}^{D_X}$ (which is observed by players and econometricians) and $\epsilon \in \mathbb{R}^1$ (which is only observed by players but not econometricians). For each of the bargaining games, econometricians observe time to agreement after the initial period, the identity of the proposer and observable states $X$ in every period, but not $\epsilon$. In this paper, we discuss identification of the model under different scenarios where cake sizes and agreed proposals may or may not be observable in the data. A typical structural approach for inferring parameters from observables posits that all bargaining games observed in the data share (i) the same transition of states (given by the conditional multinomial distribution $\tilde{L}$ and the Markov process $H$ in (1)); and (ii) the same cake function $c : \Omega_S \to \mathbb{R}^1$. Furthermore, the players in all observed bargaining games follow SSPE strategies.

In practice, data may report cross-sectional variations in the number of players $K$ and their individual characteristics $Z^K$, where $Z^K \equiv (Z_1, ., Z_K)$ with $Z_i \in \mathbb{R}^J$ for $i = 1, ., K$. Such profiles of individual characteristics vary across bargaining games observed in the data, but remain the same throughout each given game. Of course the primitives $(H, c, \tilde{L})$ may also depend on $(K, Z^K)$. These individual characteristics are perfectly observable in data and fixed over time, and our identification arguments throughout the paper are presented as conditional on $(K, Z^K)$. We suppress dependence of structural elements $c$, $\tilde{L}$, $H$ and the distributions of $(X, \rho)$ observed on the vector $(K, Z^K)$ only for the sake of notational simplicity.

## 3 Identification with Complete Data

In this section we consider identification of the cake function and the common discount rate in the canonical stochastic bargaining model when econometricians observe a complete history of *(i)* observable states $X_t$ and sizes of the cake $Y_t = c(X_t, \epsilon_t)$ (but not $\epsilon_t$); *(ii)* whether a unanimous agreement is reached in period $t$ (denoted by a dummy variable $D_t$); and *(iii)* the order of moves and the identity of the proposer (denoted $\kappa_t$) in *each* period throughout the bargaining game. Econometricians also observe the division of the cake when an agreement occurs (denoted $(\eta_{i,\tau})_{i=1}^K \in \mathbb{R}_+^K$ where $\tau$ is the termination period in which an agreement is reached), but may not observe details of the proposal in any period when no agreement is reached. We shall show that with such complete data, all model primitives can be identified under fairly weak, stochastic restrictions on state transitions and shape restrictions on the cake function. For any two generic random vectors $R_1, R_2$, we use $F_{R_2|R_1}$ to denote distribution of $R_2$ conditional on $R_1$. Let $F_{X_0}$ denote the initial distribution of observable states $X_0$ at the start of the bargaining game, and let $\Omega_X$ denote its support. We maintain the following restrictions on the transition between states throughout this section.

*CI-2 (C.I. of unobservable states) (i) Conditional on $X_{t+1}$, $\epsilon_{t+1}$ is independent of $(X_t, \epsilon_t)$ for all $t$; and (ii) conditional on $X_t$, $X_{t+1}$ is independent of $\epsilon_t$ for all $t$.*

The condition *CI-2* requires dynamics between current and next period's states $S$ and $S'$ to be captured by persistence between observable states $X$ and $X'$ only. Let $G_{X'|X}$ denote transitions between $X$ and $X'$, and $F_{\epsilon|X}$ denote the conditional distribution of the unobservable state given $X$. Then *CI-2* implies for all $t$,

$$H(S_{t+1}|S_t) = F(\epsilon_{t+1}|X_{t+1})G(X_{t+1}|X_t)$$

This assumption appears in a wide range of structural dynamic models in industrial organization and labour economics (e.g. Rust (1987)). An important implication of *CI-1,2* is that conditional on $X_t$, $(S_{t+1}, \rho_{t+1})$ are jointly independent of $\epsilon_t$.[8] Throughout the paper, we maintain the regularity condition that for all $t$ and $x \in \Omega_X$, $\Pr(X_{t+1} \in \omega | X_t = x) > 0$ for all $\omega \subseteq \Omega_X$ s.t. $\Pr(X_0 \in \omega) > 0$. Under *CI-1,2*, parameters $\theta \equiv (\beta, c, \tilde{L}_{\rho|S}, F_{\epsilon|X})$ remain to be identified, while both the OSV transition $G_{X'|X}$ and the distribution of initial states $X_0$ can be directly recovered from data. Define a feature $\Gamma(.)$ as a mapping from a vector of parameters $\theta$ to some space of features. For example, $\Gamma(\theta)$ can be a subvector of $(\beta, c, \tilde{L}_{\rho|S}, F_{\epsilon|X})$, or some functional of $c$ or $F_{\epsilon|X}$. Below we give a formal definition of identification with complete data under *CI-1,2*.

**Definition 1** *Let $\Theta$ denote a set of unknown parameters $(\beta, c, \tilde{L}_{\rho|S}, F_{\epsilon|X})$ that satisfy certain known restrictions. Two parameters $\theta, \theta'$ are observationally equivalent (denoted $\overset{o.e.}{\sim}$) under $\Theta$ if $\theta, \theta' \in \Theta$ and both generate the same joint distribution of stopping time $\tau$, agreed allocations $\eta_\tau \equiv (\eta_{\tau,i})_{i=1}^K$ and $\{X_t, Y_t, \rho_t\}_{t=0}^{\tau}$. A feature of the true parameter $\theta^*$ (denoted $\Gamma(\theta^*)$) is identified under $\Theta$ if $\Gamma(\theta^*) = \Gamma(\theta)$ for all $\theta \overset{o.e.}{\sim} \theta^*$ in $\Theta$.*

By definition, any feature of the truth $\Gamma(\theta^*)$ that can be expressed in terms of observable distributions is identified. Note observational equivalence and identification with complete data are defined under assumptions *CI-1,2* and any possible additional restrictions on parameters $(\beta, c, \tilde{L}_{\rho|S}, F_{\epsilon|X})$. Our point of departure in discussing identification is that the data-generating process (DGP) is correctly specified with certain parameters under *CI-1,2*. This implicitly requires that the observable distributions necessarily satisfy testable restrictions implied by *CI-1,2* (see the lemma below), so that the identification region for the parameter $\theta^*$ (i.e. the set of $\theta$ such that $\theta \overset{o.e.}{\sim} \theta^*$) is not vacuously empty.

**Lemma 2** *Suppose CI-1,2 hold. Then (a) a joint distribution of $(\tau, \eta_\tau, \{X_t, Y_t, \rho_t\}_{t=0}^{\tau})$ is generated*

[8] This is because $F(S', \rho'|S) = F(\rho'|S', S)H(S'|S) = \tilde{L}(\rho'|S')F(\epsilon'|X')G(X'|X)$ under *CI-1,2*.

*by some $\theta$ under SSPE only if (i) $Y_\tau = \sum_i \eta_{\tau,i}$, (ii) when $\tau \geq 1$,*

$$F_{Y_{t+1},D_{t+1},\rho_{t+1},X_{t+1}|x^t,\rho^t} = F_{Y_{t+1},D_{t+1},\rho_{t+1}|X_{t+1}} G_{X_{t+1}|x_t} \ , \ \forall x^t, \rho^t \tag{2}$$

$$F_{\eta_{t+1},\rho_{t+1}|D_{t+1}=1,x_{t+1},x^t,\rho^t} = F_{\eta_{t+1},\rho_{t+1}|D_{t+1}=1,x_{t+1}} \ , \ \forall x^{t+1}, \rho^t \tag{3}$$

*for all $x^t, \rho^t$ and $1 \leq t \leq \tau - 1$, and (iii) $F_{Y_t,D_t,\rho_t|X_t}$ and $F_{\eta_t,\rho_t|D_t=1,X_t}$ are time-homogenous (i.e. the same for all $t \geq 0$); and (b) any two $\theta \overset{o.e.}{\sim} \theta'$ if and only if they also generate the same static, time-homogenous conditional distributions $F_{\rho,D|X}$, $F_{Y|D=0,X,\rho}$ and $F_{\eta|D=1,X,\rho}$ almost everywhere on $\Omega_X$ in SSPE.*

Part (a) of the lemma summarizes the necessary testable restrictions on observable distributions under *CI-1,2*, while part (b) gives a simpler formulation of the conditions under which $\theta, \theta'$ are observationally equivalent.

*MT (Monotonicity) Both $c(x, \varepsilon)$ and $F_{\epsilon|X=x}(\varepsilon)$ are strictly increasing in $\varepsilon$ for all $x \in \Omega_X$.*

**Lemma 3** *Under CI-1,2 and MT, the common discount factor $\beta$ is identified.*

The intuition of this result is as follows. Under *CI-1,2*, the ex ante total continuation payoff (i.e. $\pi_w(s) \equiv E[w(S')|s]$) must be a function of $x$ only, and does not depend on $\varepsilon$. As a result, $\pi_w(x)$ can alternatively be expressed as the unique solution of a "quasi-structural" fixed-point (QSFP) equation defined by the unknown $\beta$ and the distribution of $(Y, D, X)$ observed. The prefix "quasi-" here is intended to highlight that structural elements $\beta, c, F_{\epsilon|X}$ enter the fixed-point formulation indirectly through distributions of $(Y, D, X)$. The assumption MT is instrumental, as it ensures a one-to-one mapping between $Y$ and $\epsilon$ given $X$. (The discount rate $\beta$ also enters the QSFP equation directly.) With observable distributions fixed from data, $\pi_w$ as a solution to the QSFP equation is shown to be strictly monotone in $\beta$. This implies the probability for reaching an agreement $\Pr(D = 1|x) = 1 - F_{Y|X=x}(\beta \pi_w(x))$ is also strictly monotone in $\beta$ under *MT*, with distribution of $(Y, D, X)$ observed and fixed from data. Therefore, $\beta$ is point-identified.

Additional restrictions are necessary for identifying the cake function $c$, the proposer-choosing mechanism $\tilde{L}_{\rho|S}$, and the USV distribution $F_{\epsilon|X}$.

*CI-3 (C.I. of order of moves) For all $t$, the order of moves $\rho_t$ is independent of $\epsilon_t$ given $X_t$.*

Condition *CI-3* requires the order of moves realized in each period to be uninformative about unobserved states given $X$. (Among other things, the condition *CI-3* is satisfied if the order of moves is determined by a function $\rho(x, \zeta)$ where $\zeta$ is independent of $\epsilon$ conditional on $X$.) It implies $Y_t = c(S_t)$ is independent of $\rho_t$ conditional on $X_t$. As Lemma 1 shows, occurrence of a unanimous agreement $D_t$ and the cake sizes $Y_t$ only depend on the current state $S_t$ realized but not on the identity of the proposer (i.e. the "separation principle" in Merlo and Wilson (1995)). Hence *CI-3* implies both $Y_t$ and $D_t$ are independent of the order of moves $\rho_t$ conditional on $X_t$. With this assumption, $\tilde{L}_{\rho|S} = L_{\rho|X}$ and it is directly identified from observable distributions along with $\beta, G_{X'|X}$. Thus under *CI-1,2,3*, only $(c, F_{\epsilon|X})$ remain to be identified in the canonical stochastic sequential bargaining model. Our next result shows that, with $\beta$,$L(\rho|X)$ identified, any two pairs $\sigma \equiv (c, F_{\epsilon|X})$ and $\sigma' \equiv (c', F'_{\epsilon|X})$ are observationally equivalent under *CI-1,2,3* if and only if they generate the same joint distribution of cake sizes and observable states.

**Proposition 1** *Suppose CI-1,2,3 and MT hold. Then both* $\beta$,$L_{\rho|X}$ *are identified, and* $\sigma \overset{o.e.}{\sim} \sigma'$ *if and only if* $F_{Y|X}(\sigma) = F_{Y|X}(\sigma')$ *a.e. on* $\Omega_X$*.*

The proposition holds under weaker restrictions $F(S_{t+1}, \rho_{t+1}|S_t) = F(S_{t+1}, \rho_{t+1}|X_t)$ and *CI-3*.[9] The intuition of the result is as follows. The occurrence of an agreement only depends on the evolution of cake sizes, which is independent of the order of moves conditional on $X$ under the assumptions above. The order of moves only determines who receives the "gains to the proposer".[10] Under *CI-1,2,3*, $L_{\rho|X}$ is identified directly along with $\beta$, and any rationalizable distribution of observables must necessarily satisfy $F_{Y,\rho|X} = F_{Y|X}L_{\rho|X}$. Hence any pair of parameters $\sigma$ and $\sigma'$ that generate the same distribution of cakes $F_{Y|X}$ must necessarily be observationally equivalent.

Without Proposition 1, we would not know what transformations of the truth $(c, F_{\epsilon|X})$ can lead to the same distributions of terminal periods $\tau$ and the division of cakes under unanimous agreements $\eta_\tau$, and therefore would not know what normalizations of $F_\epsilon$ is needed to attain full identification of the model. The result in Proposition 1 implies the same normalizations used for identifying the nonadditive random function $Y = c(X, \epsilon)$ are also necessary for identifying the canonical stochastic bargaining model. Based on this finding, the rest of this section taps into results in nonparametric identification and estimation of non-additive random functions by Matzkin (2003), and shows identification of $c, F_{\epsilon|X}$ under the independence of USV and OSV as well as appropriate normalizations.

*SI (Statistical independence)* $\epsilon_t$ *are i.i.d. across bargaining periods and independent from* $X_t$*.*

Let $\sum_{MS}$ denote the set of $\sigma$ satisfying *MT* and *SI*. Let $\tilde{\sigma} \equiv (\tilde{c}, \tilde{F}_\epsilon)$ and $\sigma \equiv (c, F_\epsilon)$ denote any two generic pairs of structural elements in $\Sigma_{MS}$.

---

[9] With the addition of CI-3, the testable restrictions for observable distributions in (2) is strengthened to

$$F_{Y_{t+1},D_{t+1},\rho_{t+1},X_{t+1}|x^t,\rho^t} = F_{Y_{t+1},D_{t+1}|X_{t+1}} L_{\rho_{t+1}|X_{t+1}} G_{X_{t+1}|x_t}$$

for all $x^t, \rho^t$.

[10] Without *CI-3*, the result of the proposition would be "$\theta \overset{o.e.}{\sim} \theta'$ if and only if $F_{Y,\rho|X}(\theta) = F_{Y,\rho|X}(\theta')$ a.e. on $\Omega_X$". To find out what normalizations are innocuous for identification, we would need to introduce further structures on how the distribution of $\rho$ is related to $Y$ given $X$.

**Corollary 1** *(Matzkin (2003)) Suppose CI-1,2,3, MT and SI hold. Then (a)* $\tilde{\sigma} \overset{o.e.}{\sim} \sigma$ *for* $\sigma, \tilde{\sigma} \in \sum_{MS}$ *if and only if* $\tilde{c}(x, \tilde{\varepsilon}) = c(x, g^{-1}(\tilde{\varepsilon}))$ *and* $\tilde{F}_\epsilon(\tilde{\varepsilon}) = F_\epsilon(g^{-1}(\tilde{\varepsilon}))$ *for some increasing function g; and (b)* $(c, F_{\epsilon|X})$ *are identified if* $c(\bar{x}, \varepsilon) = \varepsilon$ *for some* $\bar{x}$ *and all* $\varepsilon$*.*

Given the result in Proposition 1, both parts follow immediately from Matzkin (2003). Part (a) implies at least a normalization of the unobserved state is necessary to attain identification of the cake function, even with $\beta$ fixed and known and $\epsilon$ restricted to be independent from $X$. In part (b), the assumption *MT* guarantees the mapping between the cake size and the USV (latent disturbance) is one-to-one for any $x$. Then assumption *SI* ensures $F_\epsilon$ is just-identified as $F_\epsilon(t) = F_{Y|X=\bar{x}}(t)$ under the normalization $c(\bar{x}, \varepsilon) = \varepsilon$ for some $\bar{x}$ and all $\varepsilon$. It then follows $c$ is identified as $c(x, \varepsilon) = F^{-1}_{Y|X=x}(F_{Y|X=\bar{x}}(\varepsilon))$.[11]

# 4 Incomplete Data with Censored Cake Sizes

In this section, we discuss identification of the canonical stochastic bargaining model when the size of the cakes and the accepted proposals can only be observed under a unanimous agreement. In such a case, theory predicts the distribution of cake sizes observed is censored at the expected present value of total continuation payoffs for all players in SSPE (i.e. $\beta E[w(S')|s]$). As in the case with complete data in Section 3, the common discount factor can be recovered from joint distribution of observable states, stopping times and agreed allocations under *CI-1,2*. Hence, we treat $\beta$ as identified throughout this section. The major theme of this section is that additional restrictions on the cake function (i.e. additive separability in $X$ and $\epsilon$) and the unobservable state distribution (such as multiplicative heterogeneity or distributional exclusion restriction) are sufficient for identification of the cake function despite the loss of information about cake sizes when no agreement is reached.

## 4.1 Identification under multiplicative heterogeneity

We start with a class of models where unobservable state variables (USV) are known to belong to a "location-scale" family. This subsumes zero-mean normally distributed USV with variance depending on observable states.

*AS (Additive Separability) The cake function is given by* $c(x, \varepsilon) = \tilde{c}(x) + \varepsilon$ *for all* $s \in \Omega_S$*.*

[11] Matzkin (2003) also noted that a slight extension of the identification arguments above suggests $c$ can be identified when (i) $X \equiv (X_0, X_1)$ with $\epsilon$ independent of $X_1$ conditional on $X_0$; and (ii) $c(x, \varepsilon)$ is normalized by letting $c(x_0, \bar{x}_1, \varepsilon) = \varepsilon$ for all $x_0, \varepsilon$.

*MH (Multiplicative Heterogeneity) (i) For all $t$, the disturbance $\epsilon_t = \sigma(X_t)\tilde{\epsilon}_t$, where $\tilde{\epsilon}_t$ is i.i.d. across bargaining games and all periods, independent of the process of observable states $\{X_t\}_{t=0}^{+\infty}$, has median 0, and positive densities over $\mathbb{R}^1$; (ii) The scale function $\sigma(X)$ is continuous, strictly positive and bounded on $\Omega_X$.*

*MH* implies the independence of $\epsilon_t$ from history of states conditional on $X_t$, therefore the discounted ex ante total SSPE continuation payoffs $\pi_w(S) \equiv \beta E[w(S')|S]$ must be a function of $X$ only and does not involve $\epsilon$. With $\beta$ identified, $\pi_w(X) \equiv E[w(S')|X]$ can be fully recovered from observed distributions $F_{Y,D|X}$ and $G_{X'|X}$. (Alternatively, $\beta\pi_w(X)$ can be simply identified as the lower end of support of the cake sizes observed under unanimous agreements conditional on $X$.) Under *AS*, the distribution of the gains from being a proposer under state $x$ (defined as $Y^* \equiv Y - \beta\pi_w(X)$) is identical to that of $\max\{\gamma(X) + \epsilon, 0\}$, where $\gamma(x) \equiv \tilde{c}(x) - \beta\pi_w(x)$ should be intuitively interpreted as the conditional median of the proposer gains.

We then apply results in Chen, Dahl and Khan (2005) for censored regression models to identify the function of conditional median cake sizes $\tilde{c}$. The unbounded support condition in *MH* is stronger than necessary, as the identification arguments only need that *for all* $x \in \Omega_X$, support of $\tilde{\epsilon}$ is large enough to ensure the gains to the proposer is greater than zero (and therefore unanimous agreements occur) with positive probability. The zero median is a location normalization necessary for identifying $\gamma$. Most importantly, the support condition and the location-scale specification ensure the conditional quantiles of "normalized" cake sizes are linear in unknown parameters, i.e. $q_\alpha(Y^*|x) = \gamma(x) + \sigma(x)c_\alpha$ for all $x$ (where $c_\alpha$ is the $\alpha$-th quantile of $\tilde{\epsilon}$) for some $\alpha$ close enough to 1. This linearity is crucial for attaining the identification of $\tilde{c}$.

*RG (Regularity) (i) $\tilde{c}(x)$ is continuous and bounded on $\Omega_X$; (ii) for all continuous and bounded functions $b(x)$ on $\Omega_X$, $\int \max\{\tilde{c}(x)+\varepsilon, b(x)\}dF_{\epsilon|X}(\varepsilon|x)$ is bounded and continuous in $x$; (iii) $G(x'|x)$ satisfies the Feller Property (i.e. mapping continuous, bounded functions to continuous, bounded functions).*

*SG (Support of gains)* $\Pr(X \in \Omega_X^+) > 0$ *where* $\Omega_X^+ \equiv \{x : \gamma(x) > 0\}$.

**Proposition 2** *(i) Under CI-1, AS, MH, RG and SG, $\beta$ is identified, and both $\tilde{c}(X)$ and $\sigma(X)$ are identified on $\Omega_X$ from the distribution of cake sizes censored under unanimous agreements. (ii) The condition SG can be tested using the conditional distribution $F_{D|X}$.*

Since *AS* and *MH* imply *CI-2* and *MT*, the same arguments in Lemma 3 above suggest conditions *CI-1* and *MH* guarantee the common discount factor $\beta$ can be uniquely recovered from observable distributions. With $\beta$ (and therefore $\beta\pi_w(x)$) identified, the rest of the identification proof follows from Chen, Dahl and Khan (2005). Condition SG guarantees $\gamma$ can be identified directly as $Med(Y^*|x)$ over $\Omega_X^+$, which happens with a positive probability. Condition RG are regularity conditions that guarantee $\gamma(x)$ is bounded and continuous on $\Omega_X$. This ensures the conditional quantiles $q_\alpha(Y^*|x)$ must be strictly positive for some $\alpha$ greater than 1/2 and close enough to 1 even for all $x \notin \Omega_X^+$. By construction, for such values of $\alpha$, $q_\alpha(Y^*|x)$ must also be positive *for all* $x \in \Omega_X$. Under SG, we can recover these high quantiles of unobservable states (i.e. $c_\alpha$) through a linear system that relates the conditional quantiles of cake sizes to $\gamma(x)$, $\sigma(x)$ and $c_\alpha$ for $x \in \Omega_X^+$. Knowledge of these high quantiles $c_\alpha$ are used to identify $\gamma(x)$ and $\sigma(x)$ for $x \notin \Omega_X^+$ using the equation $q_\alpha(Y^*|x) = \gamma(x) + \sigma(x)c_\alpha$ for $\alpha$ close enough to 1. The major identifying restriction SG is testable, as the event "$\gamma(x) > 0$" is equivalent to "$\Pr\{\epsilon > -\gamma(x)|x\} > \frac{1}{2}$" with $\epsilon$ having strictly positive densities around 0, and $Median(\epsilon|x) = 0$ for all $x \in \Omega_X$. Therefore, this condition is equivalent to $\Pr\{\Pr(D = 1|X) > \frac{1}{2}\} > 0$ and can be tested using observable distributions.

## 4.2 Identification under exclusion restriction

While contributing to identification of $\tilde{c}$, multiplicative heterogeneity of $\epsilon$ also restricts the set of observable distributions that are consistent with the canonical model of stochastic bargaining. In particular, the MH assumption has stringent restrictions on conditional quantiles of the gains to the proposer given $X$. For *any* pair of observable states $\{x^k\}_{k=1}^2$ and *any* four percentiles $\{\alpha_j\}_{j=1}^4$ such that $q_{\alpha_j}(y^*|x^k) > 0$ for $j = 1, 2, 3, 4$ and $k = 1, 2$, the ratio between conditional interquantile range of the cakes must be independent of realized states. That is,

$$\begin{aligned}\chi(x^1; (\alpha_j)_{j=1}^4) &\equiv \frac{q_{\alpha_1}(Y^*|x^1) - q_{\alpha_2}(Y^*|x^1)}{q_{\alpha_3}(Y^*|x^1) - q_{\alpha_4}(Y^*|x^1)} = \frac{c_{\alpha_1} - c_{\alpha_2}}{c_{\alpha_3} - c_{\alpha_4}} \\ &= \chi(x^2; (\alpha_j)_{j=1}^4)\end{aligned}$$

In this subsection, we consider alternative identifying assumptions that would avoid such a stringent restrictions on observable distributions.

*ER (Exclusion restriction)* $X = (X_0, X_1)$ *and* $\epsilon$ *is independent of* $X_1$ *given* $X_0$ *in each period.*

*RS (Rich support)* $\epsilon$ *has positive densities over* $\mathbb{R}^1$ *conditional on all* $x_0$*, and* $Median(\epsilon|x) = 0$ *for all* $x \in \Omega_X$*.*

*SG2 (Support of gains)* $\Pr(X \in \Omega_X^+|x_0) > 0$ *for all* $x_0$*.*

**Proposition 3** *Under CI-1,2, AS, RG, ER, RS and SG2, the discount rate* $\beta$ *is identified and the cake function* $\tilde{c}(X)$ *is identified on the support* $\Omega_X$*.*

*AS, CI-1,2* guarantee the distribution of gains to the proposer is generated through a censored regression. (Now that *MH* is dropped, *CI-2* needs to be stated explicitly.) Conditions *CI-1,2* deliver the identification of $\beta$ as before. In contrast with the case with *MH*, under *ER* the conditional quantiles of gains to the proposer $Y^* \equiv Y - \beta\pi_w$ are additively separable in $\gamma$ and the quantiles of the USV conditional on observable states. That is, $q_\alpha(Y^*|x) = \gamma(x) + c_\alpha(x)$ for all $x \in \Omega_X$, where $c_\alpha(x)$ is an infinite dimensional nuisance parameter. Conditions *RG* and *RS* ensure that $\gamma$ is bounded on $\Omega_X$ while $\Omega_\epsilon$ is not. Hence $q_\alpha(Y^*|x) > 0$ holds even for $x \notin \Omega_X^+$ provided $\alpha$ is greater than 1/2 and close enough to 1. Conditions *ER* and *SG2* allow us to fix any $x_0$ and exploit "enough" variations in $x_1$ alone to reach some state $\tilde{x} \equiv (x_0, \tilde{x}_1) \in \Omega_X^+$. With a slight abuse of notation, note *ER* implies $c_\alpha(\tilde{x}) = c_\alpha(x_0)$ for any $\alpha \geq 1/2$, and these are recovered as $q_\alpha(Y^*|\tilde{x}) - \gamma(\tilde{x})$ with both components observed directly for $\tilde{x} \in \Omega_X^+$. With knowledge of the nuisance parameter $c_\alpha(x_0)$ for any $\alpha \geq 1/2$, $\gamma(x)$ can be identified for $x \notin \Omega_X^+$ as $q_{\bar{\alpha}}(Y^*|x) - c_{\bar{\alpha}}(x_0)$ for some $\bar{\alpha}$ greater than 1/2 and close enough to 1. Then $\tilde{c}$ is recovered on $\Omega_X$ as the sum of $\gamma$ and $\beta\pi_w(x)$. The exclusion restriction has some simple, testable restrictions on observable distributions:

$$q_\alpha(Y^*|x) - q_\alpha(Y^*|\tilde{x}) = q_{\alpha'}(Y^*|x) - q_{\alpha'}(Y^*|\tilde{x}) = \gamma(x) - \gamma(\tilde{x})$$

for all $\alpha, \alpha', x, \tilde{x}$ such that $\min\{q_\alpha(x), q_\alpha(\tilde{x}), q_{\alpha'}(x), q_{\alpha'}(\tilde{x})\} > 0$ and $x_0 = \tilde{x}_0$. We leave the construction of nonparametric estimators and discussion of their statistical properties for future work.

We conclude this section by noting that with $\beta$ identified first, $\tilde{c}$ is recovered from the censored distribution of cake sizes alone under the assumptions above. The distribution of allocations of the cake under agreements are not involved in the identification of $\tilde{c}$ above. This is because of the additional identifying power from the independence of $\tilde{c}$ from the order of moves given the OSV, and the new structures imposed through the assumptions *AS* and *MH* or *ER*.

## 5 Incomplete Data with Unobservable Cake Sizes

In some other empirical contexts, econometricians can observe realized states in all periods and when a unanimous agreement is reached, but never observe the cake sizes, the order of moves, or the agreed allocations. For example, such a scenario arises when all parties involved in the game choose to keep details of negotiations and agreements confidential, and outsiders (econometricians) only get to observe factors that are known to affect the cake sizes. Econometricians seek to learn enough about the underlying structures (i.e. cake functions, distributions of the USV, etc) to predict probabilities for an agreement under counterfactual contexts (such as when the transition between states are perturbed). Suppose conditions *CI-1,2* hold. Let $\Theta$ denote the set of generic restrictions on unknown parameters $\theta \equiv (\beta, c, F_{\epsilon|X})$.[12]

[12] Throughout this section, we maintain that the common discount factor $\beta$ is known to econometricians. Within the class of canonical models where players' utilities are transferrable, this restriction is often justifiable as the discount rate can usually be recovered exogenously. For example, in some empirical applications, the cake size is measured in monetary terms and the discount rate can be estimated as the interest rate that lasts throughout the bargaining process.

**Definition 2** *Two parameters $\theta$ and $\theta'$ are observationally equivalent (denoted $\overset{o.e.}{\sim}$) under $\Theta$ if $\theta, \theta' \in \Theta$, $F_{D|X}(\theta) = F_{D|X}(\theta')$ a.e. on $\Omega_X$. A feature of the truth $\theta^*$ (denoted $\Gamma(\theta^*)$) is identified under $\Theta$ if $\Gamma(\theta^*) = \Gamma(\theta)$ for all $\theta \overset{o.e.}{\sim} \theta^*$ under $\Theta$.*

As in the canonical model with complete data, identification is defined under conditional independence restrictions *CI-1,2*. In contrast with the previous scenarios, observational equivalence in the data scenario considered in this section only requires parameters to generate identical static, conditional probabilities of unanimous agreements $F_{D|X}$ only. This is because now neither the orders of moves nor the agreed allocations of the cake are observed in data. Our point of departure in discussion of identification in this section is that the model is correctly specified under *CI-1,2* for some parameters $(\beta, c, F_{\epsilon|X}, \tilde{L}_{\rho|S})$. Thus it is implicitly required that the distributions observed necessarily satisfy testable restrictions of these assumptions (i.e. $F_{D_{t+1},X_{t+1}|D^t,X^t} = F_{D_{t+1}|X_{t+1}} G_{X_{t+1}|X_t}$, with $F, G$ time-homogenous) so that the identification region is not vacuously empty. Also note *CI-3* (conditional independence of $\epsilon$ from $\rho$ given $X$) is dropped as the order of moves is not observed in the current scenario.

## 5.1 Identifying the cake function with known USV distribution

We start by examining what can be learned about model primitives in the simplest case where the distribution of USV is known to econometricians. Throughout this section, we maintain the *AS* assumption that the cake function is additively separable in states observed and USV, i.e. $c(s) = \tilde{c}(x) - \varepsilon$ for some unknown function $\tilde{c}$. The following lemma shows that at least some scale and location normalizations of $F_{\epsilon|X}$ are required to identify $\tilde{c}$, and such normalizations are innocuous for predicting counterfactual agreement probabilities.

**Lemma 4** *Under AS, CI-1,2, the location and scale of $F_{\epsilon|X}$ cannot be identified jointly with $\tilde{c}$.*

This lemma is proved by showing that if the model is correctly specified for some $(\tilde{c}, F_{\epsilon|X})$, then any affine transformations $(a\tilde{c}(x)+b, F_{\epsilon|X}(\frac{\varepsilon-b}{a}|x))$ can also generate the same agreement probability $\Pr(D = 1|X)$ a.e. on $\Omega_X$. The lemma implies that scale and locational normalizations of $F_{\epsilon|X}$ are innocuous for predicting agreement probabilities under counterfactual changes in $\beta$ or $G_{X'|X}$. To see this, simply note the affine transformations above are independent of the discount rate and the transitions between observable states. Thus for any given $(\beta, G)$, the same affine transformation of the true parameters $(\tilde{c}, F_{\epsilon|X})$ is observationally equivalent to the truth both in the data-generating process (DGP) and in the counterfactual context of interest.

This also suggests the common assumption in empirical work that $F_{\epsilon|X}$ is known to researchers is less restrictive than it seems. Consider the cases where distributions of unobserved states are known to be independent of $X$ and belong to the normal family. Then restricting $F_{\epsilon|X}$ to be $N(0,1)$ in estimation is equivalent to introducing a local and scale normalization that is innocuous. This is also true with other parametric families characterized by location and scale parameters only. Berry and Tamer (2006) showed a positive result in identifying optimal-stopping models when the USV distribution $F_{\epsilon|X}$ is known to econometricians. Their result also applies in our context of canonical stochastic bargaining models.

**Proposition 4** *(Berry and Tamer (2006)) Suppose $c(s) = \tilde{c}(x) - \varepsilon$ for some unknown function $\tilde{c}$, and suppose $\beta$ is known, AS, CI-1,2 hold and $F_{\epsilon|X=x}$ is known and strictly increasing for all $x \in \Omega_X$. Then (i) $\tilde{c}$ is identified for $x \in \Omega_X$. (ii) If $\tilde{c}(\bar{x})$ is known for some $\bar{x} \in \Omega_X$, both $\beta$ and $\tilde{c}$ are identified.*

As shown in Section 2, theory suggests a "separation principle" where occurrence of a unanimous agreement only depends on states $S$ realized but not on the identity of the proposer under SSPE. A bargaining game ends when the realized cake size exceeds the ex ante total continuation payoff. Berry and Tamer (2006) showed with knowledge of the discount factor and the distribution of the USV, the optimal decision rule in dynamic stopping problems can also be fully recovered from conditional probabilities of agreements. This in turn helps identify the cake function. We adapt their proof in our context in the appendix.

## 5.2 Rationalizable counterfactual outcomes when USV distribution is unknown

When the unobservable state variable (USV) distribution is not known to belong to certain location-scale parametric family, imposing a specific form of $F_{\epsilon|X}$ that deviates from the truth can imply incorrect results in counterfactual outcomes. (See the example in the next subsection.) On the other hand, economic theories often suggest the structural elements of the model have to satisfy certain nonparametric shape or stochastic restrictions (such as monotonicity or concavity of the cake function or independence of $\epsilon$ from $X$). This naturally raises the question: how can econometricians exploit such exogenously given restrictions to infer counterfactual outcomes without imposing parametric assumptions on the structure? We propose a simple, novel algorithm that helps recover the complete set of all rationalizable probabilities for reaching an agreement in counterfactual bargaining contexts where transitions between states are perturbed.

For the rest of this subsection, we maintain that $\epsilon$ is statistically independent of $X$ (Assumption *SI*). We begin by noting a pair of parameters $(\tilde{c}, F_\epsilon)$ is observationally equivalent to the true parameters if and only if the following equation is satisfied:

$$q(p(x); F_\epsilon) \equiv F_\epsilon^{-1}(p(x)) = \tilde{c}(x) - \beta\pi_w(x; p, F_\epsilon)$$

where $p(x)$ is the probability for reaching an agreement conditional on $x$, and $\pi_w(x; p, F_\epsilon)$ is the ex ante total continuation payoffs in SSPE and solves the fixed point equation:

$$\begin{aligned} \pi_w(x) &= \int \max\{\tilde{c}(x') - \varepsilon', \beta\pi_w(x')\} dF(\varepsilon'|x') G(x'|x) \\ &= \int \beta\pi_w(x') + \phi(p(x'); F_\epsilon) dG(x'|x) \end{aligned} \tag{4}$$

with

$$\phi(p; F_\epsilon) \equiv \int_{-\infty}^{F_\epsilon^{-1}(p)} F_\epsilon^{-1}(p) - \varepsilon dF_\epsilon(\varepsilon)$$

The latter will be referred to as the "*conditional surplus function*" (CSF) hereafter. Note the second equality in (4) uses the fact that $p(x) = F_\epsilon(\tilde{c}(x) - \beta\pi_w(x))$ For the rest of this subsection, we will focus on the case where the support of $X$ is finite.

*DS (Discrete support) The support of $X$ (denoted $\Omega_X$) is finite with $M$ elements $\{x_1, x_2, ., x_M\}$.*

In discretized notations, a pair of parameters $(\tilde{c}, F_{\epsilon|X})$ is observationally equivalent to the true parameters underlying the DGP if and only if the following system of $M$ linear equations holds:

$$Q = \tilde{C} - \beta\Pi \tag{5}$$

where $Q$, $\tilde{C}$, $\Pi$ are $M$-vectors with $Q_m \equiv F_\epsilon^{-1}(p(x_m))$, $\tilde{C}_m \equiv \tilde{c}(x_m)$, and $\Pi$ solves

$$\Pi = G(\beta\Pi + \Phi) \tag{6}$$

where $\Phi$ is a $M$-vector with the $m$-th coordinate defined as $\Phi_m \equiv \phi(p(x_m))$, and $G$ is the $M$-by-$M$ transition matrix with the $(m, n)$-th entry defined as $G_{mn} \equiv \Pr(X' = x_n | X = x_m)$. Note the probabilities of reaching an agreement $p \equiv (p(x_1), ., p(x_M))$ enters the system defining observational equivalence through $\Phi$. in both $Q$ and $\Pi$. We shall normalize $q(1/2) = 0$ and $\phi(1/2) = \bar{\phi}$ for some strictly positive constant $\bar{\phi}$. (Note from our discussions in the previous subsection, such scale and location normalizations are innocuous for identifying counterfactual agreement probabilities when $\tilde{c}$ or $G$ are perturbed.) Under this normalization, $\tilde{c}(x)$ is the median cake size given state $x$.

In some empirical contexts, econometricians know $\tilde{C}$ satisfies certain shape restrictions that are derived exogenously from economic theory or common sense. Such restrictions can often be represented as a system of linear restrictions on the vector $\tilde{C}$. For example, if the state $x$ has three possible values $x_1 < x_2 < x_3$ and $\tilde{C}$ is known to be strictly increasing and concave. Then $\tilde{C}$ is known to satisfy the linear restriction $A\tilde{C} > 0$ with

$$A \equiv \begin{bmatrix} -1 & 2 & -1 \\ -1 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix}$$

Similar matrices of coefficients can be constructed if $\tilde{c}$ is only known to be increasing or concave in one of the coordinates in a multivariate $x$. Besides, any ranking of a subset of the states by $\tilde{c}(x)$ can also be represented as linear restrictions on $\tilde{C}$. For example, if a certain state $x_m$ is known to lead to the strictly smallest median size of the cake, this will lead to additional $M-1$ strict inequalities.

Given a set of exogenously given restrictions on the structure $\tilde{C}$ and $F_{\epsilon|X}$, a vector of agreement probabilities $p$ observed is *rationalized* if there exist $\tilde{C}, F_{\epsilon|X}$ such that (i) $\tilde{C}, F_{\epsilon|X}$ satisfy these restrictions and (ii) $p$ is generated as players make dynamic rational choices in SSPE given $\tilde{C}, F_{\epsilon|X}$ (i.e. $p$ satisfies (5), (6) under $\tilde{C}, F_{\epsilon|X}$). The next lemma gives sufficient and necessary conditions for a vector $p$ to be rationalized when $\tilde{C}$ satisfies some linear restrictions and $\epsilon$ is independent of $X$. Let $V_{(m)}$ denote the $m$-th smallest element in a generic vector $V$. Let $\bar{\phi}$ be any strictly positive constant.

**Lemma 5** *Suppose AS, CI-1,2 and DS, SI hold, $\beta$ is known and $c(S) = \tilde{c}(X) - \epsilon$ with $Median(\epsilon) = 0$ for an unknown function $\tilde{c}$ that satisfies a set of linear restrictions $A\tilde{C} > 0$. Then a vector of the probability of agreements $p$ observed in the DGP is rationalized if and only if the following linear system holds for some vectors $Q, \Phi$:*

$$A\left[Q + \beta(I - \beta G)^{-1} G\Phi\right] > 0 \tag{7}$$

$$Q_m \leq Q_n \Leftrightarrow p_m \leq p_n, \ \forall m, n \in \{1, ., M\} \tag{8}$$

$$p_{(m)}(Q_{(m+1)} - Q_{(m)}) \leq \Phi_{(m+1)} - \Phi_{(m)} \leq p_{(m+1)}(Q_{(m+1)} - Q_{(m)}), \ \forall m \in \{1, ., M-1\} \tag{9}$$

$$Q_m \leq 0 \Leftrightarrow p_m \leq 1/2 \ \text{and} \ \tfrac{1}{2}Q_m \leq \Phi_m - \bar{\phi} \leq p_m Q_m, \, \forall m \in \{1, ., M\} \tag{10}$$

$$\Phi_m > 0 \ \text{for} \ m \in \{1, ., M\} \tag{11}$$

Two remarks are in order. First, the feasibility of the linear system is *not only necessary but also sufficient* for a vector of agreement probabilities $p$ to be rationalized in a model of stochastic bargaining. Sufficiency follows from the fact that when the linear system holds, a pair of structural elements $(\tilde{C}, F_\epsilon)$ can be constructed to rationalize $p$ under *CI-1,2*, *DS* and *SI* as probability for agreements in SSPE. More specifically, such a $F_\epsilon$ can be constructed through interpolations between $Q$ (with its CSF satisfying constraints in (9)), and $\tilde{C} = Q + \beta(I - \beta G)^{-1}G\Phi$. Second, without the shape restrictions $A\tilde{C} > 0$, the lemma would be vacuous in the sense that we would always be able to define $Q, \Phi$ recursively from any $p \in (0,1)^M$ and the linear system (8)-(11), and then define $\tilde{C}$ as above so that $p$ is rationalized by $(\tilde{C}, F_\epsilon)$ under *CI-1,2*, *DS* and *SI*. In other words, without the shape restrictions $A\tilde{C} > 0$, any $p$ in $(0,1)^M$ must be rationalized by some $(\tilde{C}, F_\epsilon)$.

A standard approach for structural analyses of probabilities for agreements under counterfactual environments (such as perturbations in the state transitions or the cake function) would take two steps. First, identify and estimate the cake function $\tilde{c}$ and the USV distribution $F_\epsilon$ using observable distributions from the DGP, and second, use the identified parameters to predict the conditional probability of reaching agreements in the counterfactual contexts. Unfortunately, when the USV distribution is not restricted to take any known parametric form, $\tilde{c}$ and $F_\epsilon$ may not be uniquely recovered from observables, and the first step fails. This non-identification is obvious from the lemma above, as the vector $(Q, \Phi)$ (i.e. the percentiles and the CSF corresponding to the USV distribution) that satisfies this linear system is not unique. For the rest of this section, we shall argue that, despite this non-identification result, a simple algorithm can be used to recover *all* rationalizable conditional agreement probabilities under counterfactual contexts. These are counterfactual agreement probabilities consistent with the model structure and restrictions (including shape restrictions on $\tilde{c}$ such as monotonicity, and stochastic restrictions on USV distribution such as independence of $\epsilon$ from $X$).

Below we formally define rationalizable counterfactual outcomes. Suppose the data-generating process (DGP) is characterized by true parameters $\tilde{c}^0, F_\epsilon^0, \beta, G_{X'|X}^0$ which generate conditional probabilities of agreement $p^0 \in [0,1]^M$ observed in SSPE. We are interested in the probability for agreements under two types of counterfactual environments: (a) the transition between observable states is perturbed from $G_{X'|X}^0$ to $G_{X'|X}^1$ while $\beta, \tilde{c}^0, F_\epsilon^0$ are fixed; or (b) the cake function is changed to $\tilde{c}^1(x) \equiv \tilde{c}^0(x)\alpha(x)$ (where $\alpha(x) \in \mathbb{R}_{++}^1$ denotes percentage changes in the (median) cake size given state $x$), while $G_{X'|X}^0, F_\epsilon^0$ remain the same. Suppose *CI-1,2* hold, $c(x, \varepsilon) = \tilde{c}(x) - \varepsilon$ , and $\beta$ is fixed and known.

**Definition 3** *Given certain restrictions on unknown structure $\tilde{c}, F_{\epsilon|X}$, the identified set of rationalizable counterfactual outcomes (ISRCO) consists of all conditional probabilities for agreement $p^1 \in [0,1]^M$ such that $(p^0, p^1)$ are jointly rationalized by some $\tilde{c}, F_{\epsilon|X}$ that satisfies the restrictions.*

The next proposition introduces a simple, new algorithm that recovers *ISRCO*. The basic idea extends the preceding lemma by synthesizing two linear systems characterizing rationalizability in two bargaining environments respectively (one observed and one counterfactual). This synthesis exploits the fact that the nuisance parameter $F_\epsilon$ is fixed under both contexts. The observed and (unknown) counterfactual outcomes $(p^0, p^1)$ enter the coefficient matrix of the "synthesized" linear system. Thus a $\tilde{p} \in [0,1]^M$ belongs to *ISRCO if and only if* the $2M$-vector $(p^0, \tilde{p})$ makes the synthesized linear system feasible with solutions in unknown parameters. The consistency of a linear system can be checked through standard linear programming algorithms. For four $2M$-vectors $(Q^j, \Phi^j)_{j=0,1}$, let $\tilde{p}^{01}$, $\tilde{Q}^{01}$, $\tilde{\Phi}^{01}$ denote $(2M+1)$-vectors that are defined as $[p^0, p^1, 1/2]$, $[Q^0, Q^1, 0]$, $[\Phi^0, \Phi^1, \bar{\phi}]$ respectively for some positive constant $\bar{\phi}$. Let $G^1$ denote the counterfactual state transitions of interests in (a) above, and let $\alpha$ denote a $M$-by-$M$ diagonal matrix with the $m$-th diagonal entry being $\alpha(x_m)$ as in (b) above.

**Proposition 5** *Suppose the assumptions in Lemma 5 all hold. Then the ISRCO in (a) is the set* *of all $p^1$ such that the following linear system holds for some $(Q^j, \Phi^j)_{j=0,1}$:*

$$Q^0 + \beta(I - \beta G^0)^{-1} G^0 \Phi^0 = Q^1 + \beta(I - \beta G^1)^{-1} G^1 \Phi^1 \tag{12}$$

$$A\left[Q^0 + \beta(I - \beta G^0)^{-1} G^0 \Phi^0\right] > 0 \tag{13}$$

$$Q^j_m \leq Q^k_n \Leftrightarrow p^j_m \leq p^k_n \ \forall m, n \in \{1, ., M\},\ j, k \in \{0, 1\} \tag{14}$$

$$\tilde{p}^{01}_{(m)}(\tilde{Q}^{01}_{(m+1)} - \tilde{Q}^{01}_{(m)}) \leq \tilde{\Phi}^{01}_{(m+1)} - \tilde{\Phi}^{01}_{(m)} \leq \tilde{p}^{01}_{(m+1)}(\tilde{Q}^{01}_{(m+1)} - \tilde{Q}^{01}_{(m)}) \ \textit{for}\ 1 \leq m \leq 2M \tag{15}$$

$$Q^j_m \leq 0 \Leftrightarrow p^j_m \leq 1/2\ ;\ \Phi^j_m > 0 \ \forall m, n \in \{1, ., M\},\ j, k \in \{0, 1\} \tag{16}$$

*And the ISRCO in (b) is the set of all $p^1$ such that a similar linear system (13)-(16) and (17) is feasible with solutions $(Q^j, \Phi^j)_{j=0,1}$, where (17) is defined as*

$$Q^1 + \beta(I - \beta G^0)^{-1} G^0 \Phi^1 = \alpha\left(Q^0 + \beta(I - \beta G^0)^{-1} G^0 \Phi^0\right) \tag{17}$$

Thus recovering *ISRCO* amounts to collecting all $p^1$ in $[0,1]^M$ such that $[p^0, p^1]$ makes the linear system feasible with solutions in $\{Q^j, \Phi^j\}_{j=0,1}$. Remarkably, this approach does not require any parametric assumption on the cake function or the USV distribution. On the other hand, it fully exploits the independence of $\epsilon$ and $X$ and exogenously given shape restrictions "$A\tilde{C} > 0$". By construction, the *ISRCO* consists of all possible counterfactual outcomes that could be rationalized under the model restrictions (i.e. dynamic rationality, independence of $\epsilon$ of $X$ and $A\tilde{C} > 0$).

We conclude this section by emphasizing that the *ISRCO* is interesting in its own right, regardless of its size relative to the outcome space $[0,1]^M$. This is because our approach efficiently exhausts *all information* about the counterfactuals that can be extracted from known restrictions on the model. Thus the set reveals the limit of what can be learned about the probability for agreements under counterfactual contexts, if econometricians choose to remain agnostic about the functional form of the structural elements. In the next section, we illustrate the algorithm in a simple numeric example. The *ISRCO* recovered there is small relative to the outcome space and quite informative.

### 5.3 A simple numeric example

In this subsection, we use a simple numeric example to illustrate the consequence of normalizations (locational and scale) and misspecifications of the USV distributions on counterfactual analyses. We also use the example to illustrate the algorithm proposed in Proposition 5 for recovering the set of rationalizable counterfactual outcomes.

*(Counterfactual outcomes when the true distribution of USV is uniform and known)* Suppose $M = 3$ and $\epsilon$ is independent of $X$ with a true USV distribution $F_\epsilon$ that is uniform on $[-5, 5]$. Thus $q(p_k; F_\epsilon) = 10p_k - 5$ and $\phi(p_k; F_\epsilon) = 5p_k^2$ for $p_k \in [0,1]$. For any $p = [p_1, p_2, p_3] \in [0,1]^3$, let

$Q^{unif}(p), \Phi^{unif}(p)$ denote $\mathbb{R}^3$-vectors with $k$-th coordinate being $q(p_k; F_\epsilon)$ and $\phi(p_k; F_\epsilon)$ respectively. Let the discount rate $\beta$ be 4/5, and the observed transition $G^0$ and the counterfactual transition $G^1$ be respectively defined as

$$G^0 \equiv \begin{bmatrix} 28/73 & 67/219 & 68/219 \\ 13/43 & 83/172 & 37/172 \\ 5/26 & 1/104 & 83/104 \end{bmatrix} \; ; \; G^1 \equiv \begin{bmatrix} 25/74 & 15/74 & 17/37 \\ 35/59 & 9/59 & 15/59 \\ 42/115 & 19/115 & 54/115 \end{bmatrix}$$

(These specifications are chosen randomly.) Suppose the true cake function (i.e. median cake sizes conditional on observable states) is

$$\tilde{C}_{unif} = [\ \tfrac{717442573}{165078240}, \ \tfrac{97368349}{132062592}, \ \tfrac{330851369}{264125184}] \approx [4.3461, \ 0.7373, \ 1.2526]$$

while the actual conditional probability for reaching agreement observed in the DGP is $p^0 = [\frac{3}{5}, \frac{1}{4}, \frac{5}{16}]$.[13] A counterfactual outcome under $G^1$ is a vector in $[0,1]^3$ (denoted by $p^1_{unif}$) with the $k$-th coordinate being Pr(an agreement is reached $|x_k$). The subscript is a reminder that the counterfactual outcome is calculated using the assumed knowledge that the USV is uniform on $[-5, 5]$. By definition, $p^1_{unif}$ solves a system of quadratic equations

$$Q^{unif}(p^1_{unif}) + \beta(I - \beta G^1)^{-1} G^1 \Phi^{unif}(p^1_{unif}) = \tilde{C}_{unif} \tag{18}$$

The solution is found to be $p^1_{unif} \approx [0.5880 \ , \ 0.2004 \ ; \ 0.2760]$.[14]

*(Innocuous location and scale normalizations)* Now suppose econometricians only know USV is uniformly distributed, but mis-specify the support (scale and location) of $\hat{F}_\epsilon$ as $[b - a, b + a]$ for some constants $a \in \mathbb{R}^1_+$, $b \in \mathbb{R}^1$ and $a \neq 5$, $b \neq 0$. Thus $q(p_k; \hat{F}_\epsilon) = b - a + 2ap_k$ and $\phi(p_k; \hat{F}_\epsilon) = ap_k^2$. For any $p \in [0,1]^3$, let $\hat{Q}^{unif}(p), \hat{\Phi}^{unif}(p)$ denote vectors with $k$-th coordinate being $q(p_k; \hat{F}_\epsilon)$ and $\phi(p_k; \hat{F}_\epsilon)$ respectively (i.e. quantile and conditional surplus functions calculated based on the wrong assumption $\hat{F}_\epsilon$). Then $\tilde{C}$ would be recovered (incorrectly) as

$$\hat{C}_{unif} = \hat{Q}^{unif}(p^0) + \beta(I - \beta G^1)^{-1} G^1 \hat{\Phi}^{unif}(p^0) \tag{19}$$

Straightforward substitutions show this misspecification still leads to the same system of nonlinear equations in $p^1_{unif}$ as (18). In other words, even though $Q^{unif}(.)$, $\Phi^{unif}(.)$ and $\tilde{C}_{unif}$ have different forms now due to the misspecification of $F_\epsilon$, the structure of the model is such that the differences cancel out and yield the same system of nonlinear equations in (18). (See the Appendix for algebraic details.) This verifies our remarks earlier (following Lemma 4) that the scale and locational normalizations of the USV distribution is innocuous for recovering counterfactuals.

[13] While choosing specifications of the example, we actually let $p^0$ be fixed at $[\frac{3}{5}, \frac{1}{4}, \frac{5}{16}]$ first, and then solve for $\tilde{C}_{unif}$ backwards by substituting $p^0$ into $\tilde{C}_{unif} = Q^{unif}(p^0) + \beta(I - \beta G^0)^{-1} G^0 \Phi^{unif}(p^0)$, where the functional forms of $Q^{unif}, \Phi^{unif}$ are defined above.

[14] See the Appendix for analytical close forms of the system of nonlinear equations. We use the "fmincon" function to solve for $p^1_{unif}$. The solution must be unique because given $\tilde{c}, F_{\epsilon|X}, G_{X'|X}$, the ex ante total continuation payoff $\pi_w$ is unique.

*(Consequence of misspecifying USV distributions)* Suppose $\beta, G^0, G^1$ are still defined as above, but now econometricians misspecify USV to be generalized log-logistic with a distribution function

$$F_\epsilon(t) = \left(1 + \left[1 + \xi(\tfrac{t-\mu}{\sigma})\right]^{-1/\xi}\right)^{-1}$$

with parameters $\mu = 0$ (location), $\sigma = 1$ (scale), $\xi = 1$ (shape). The distribution is positively skewed with support bounded below at $-1$. For any $p = [p_1, p_2, p_3] \in [0,1]^3$, let $Q^{GLL}(p), \Phi^{GLL}(p)$ denote $\mathbb{R}^3$-vectors with the $k$-th coordinate being the quantile and conditional surplus functions at $p_k$, i.e.

$$\begin{aligned} q(p_k; F_\epsilon) &\equiv q_k = \mu + \tfrac{\sigma}{\xi}\left((\tfrac{1-p_k}{p_k})^{-\xi} - 1\right) \\ \phi(p_k; F_\epsilon) &\equiv q_k + 1 - \log(q_k + 2) \end{aligned}$$

respectively. Thus the conditional median cake function is recovered (incorrectly) from the observed $p^0$ as follows:

$$\tilde{C}_{GLL} = Q^{GLL}(p^0) + \beta(I - \beta G^0)^{-1} G^0 \Phi^{GLL}(p^0)$$

Then the implied counterfactual outcome $p^1_{GLL}$ must solve

$$Q^{GLL}(p^1_{GLL}) + \beta(I - \beta G^1)^{-1} G^1 \Phi^{GLL}(p^1_{GLL}) = \tilde{C}_{GLL} \tag{20}$$

Solving (18) with the right-hand side given by $\tilde{C}_{GLL}$ yields an implied counterfactuals $p^1_{GLL} \approx [.5926, .1317, .2311]$, where the subscript $GLL$ emphasizes this is the counterfactual outcome predicted under the misspecification of the USV in structural estimation.[15] This implies misspecifying USV to be a general log-logistic while the truth in the DGP is uniform is not innocuous, as it induces discrepancies between the counterfactual outcomes it implies and the true counterfactual outcomes.

*(Robust identification of ISRCO without knowing the USV distribution)* Now let the true $\beta, G^0$ underlying the DGP be defined as above, with $F_\epsilon$ uniform on $[-5, 5]$ and the conditional median cake function $\tilde{C} = \tilde{C}_{unif}$. As before, the conditional agreement probability observed in data is $p^0 = [\frac{3}{5}, \frac{1}{4}, \frac{5}{16}]$. Econometricians do not know the USV distribution $F_\epsilon$ or the true $\tilde{C}$. They only observe $p^0$ and know $\beta, G^0$ in the DGP, and are interested in predicting the counterfactual probabilities for agreements when the transition between states is changed to $G^1$. Furthermore, econometricians correctly learn from outside the model that the second state yields the lowest static payoff, i.e. $\tilde{c}(x_2) < \min\{\tilde{c}(x_1), \tilde{c}(x_3)\}$. Then the algorithm proposed above can be used to recover the complete set of *ISRCO* by collecting all $p^1 \in [0,1]^3$ that make the linear system (12)-(16) feasible. (See the Appendix for details in implementing the algorithm.) Figure 1 depicts the set of *ISRCO* recovered is about 5.1% of the outcome space $[0,1]^3$.

[15] We use a built-in command "fmincon" in Matlab to solve the system of nonlinear equations, which may have multiple solutions in general. The solution reported here is robust to the choice of initial point for the algorithm.

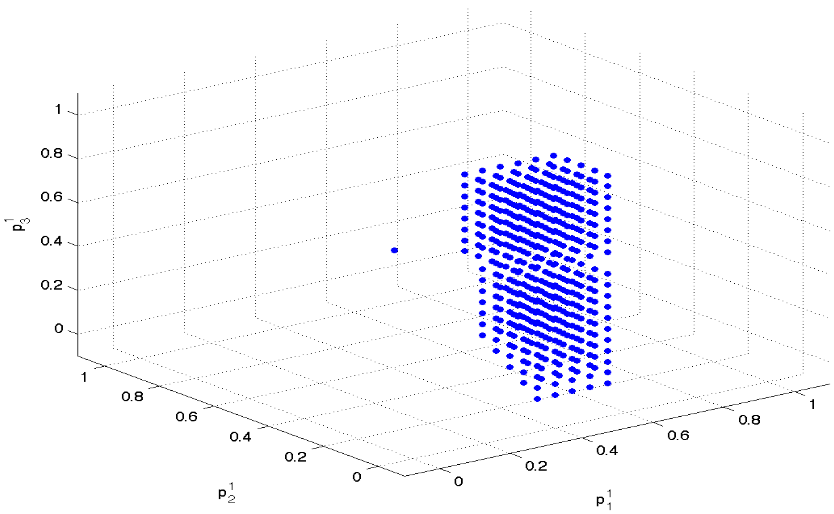


Figure 1: *ISRCO* with $\epsilon \perp X$, $\tilde{c}_2 < \min(\tilde{c}_1, \tilde{c}_3)$ and $p^0$=[3/5,1/4,5/16]

Our algorithm for recovering *ISRCO* only requires $\epsilon$ to be independent of $X$. The *ISRCO* is exhaustive and sharp in the following senses: (i) as long as the true USV distribution in the DGP satisfies this independence restriction and "$\tilde{C}_2 < \min\{\tilde{C}_1, \tilde{C}_3\}$", the true counterfactual outcomes under $G^1$ must lie in *ISRCO*; and (ii) any outcome vector in *ISRCO* is a rationalizable counterfactual outcomes corresponding to certain $F_\epsilon$ that satisfies independence from $X$ and $\tilde{C}$ such that $\tilde{C}_2 < \min\{\tilde{C}_1, \tilde{C}_3\}$. Also note in implementing the algorithm we have invoked a location normalization ($Median(\epsilon) = 0$) and a scale normalization ($\phi(1/2; F_\epsilon) = \bar{\phi} > 0$), which are known to be innocuous for counterfactual analyses from Lemma 4 above.

# 6 Extensions

So far we have focused on a canonical stochastic bargaining model where players' utilities are linear in the surplus they receive, and all players share the same discount rate throughout the game. Lemma 1 shows the payoffs in stationary subgame perfect equilibria is unique in such contexts. This section studies the identification when players evaluate the surplus received according to a concave utility function, or the discount rates are different across players. In either of these two cases, players' payoffs from SSPE are no longer unique in general. We shall show the utility function and the discount rates can be identified with complete data on the occurrence of agreements and divisions of the cake under unanimous agreements, provided players across all bargaining games observed in the DGP adopt strategies that yield the same profile of SSPE payoffs (which are functions of observable states).

*SE (Single equilibrium payoff) Players in all bargaining games observed follow stationary subgame perfect strategies that lead to the same profile of SSPE payoffs (as functions of states).*

This restriction is analogous to the "single-equilibrium" assumption used in the literature of estimating discrete games of incomplete information in the presence of multiple Bayesian Nash equilibria (e.g. Bajari, Hong, Krainer and Nekipelov (2008) and Tang (2009)). Such a "single-SSPE-payoff" restriction allows econometricians to exploit the characterization of SSPE payoffs in (5) and (6) to relate observable distributions to model primitives, without the need to specify which SSPE payoff is followed by players in the games observed.

## 6.1 Concave Utility Functions

In this subsection, we extend the basic model with complete information by relaxing restrictions of transferable utilities. The set of feasible allocations is now given by $C(s) = \{t \in \mathbb{R}^K : \sum_i u^{-1}(t_i) \leq c(s)$ for some von-Neumann Morgenstern utility function $u : \mathbb{R}^1_+ \to \mathbb{R}^1_+\}$. Econometricians observe the cake sizes, the identity of the proposer, and the physical shares of the cake for each player when a unanimous agreement occurs, but do not know the utility levels associated with these shares. The lemma below characterizes the SSPE payoffs in this model.

**Lemma 6** *Suppose CI-1 holds. Then (a) $v \in F^K$ is a SSPE payoff in the bargaining game with nontransferable utilities (NTU) if and only if $A(v) = v$ where for all $\tilde{s} = (s, \rho) \in \Omega_{\tilde{S}}$ and for all $i$,*

$$\begin{aligned} A_i(f)(\tilde{s}) &\equiv \max\left\{u\left(c(s) - \textstyle\sum_{j\neq i} u^{-1}(\beta E[v_j(\tilde{S}')|S=s])\right), \beta E[v_i(\tilde{S}')|S=s]\right\}, \text{ if } \rho_{(1)} = i \\ A_i(f)(\tilde{s}) &\equiv \beta E[v_i(\tilde{S}')|S=s], \text{ if } \rho_{(1)} \neq i \end{aligned}$$

*(b) For any SSPE payoff $v \in F^K$, a unanimous agreement occurs in state $s$ when the proposer is $i$ if and only if,*

$$u\left(c(s) - \textstyle\sum_{j\neq i} u^{-1}(\beta E[v_j(\tilde{S}')|S=s])\right) \geq \beta E[v_i(\tilde{S}')|S=s] \tag{21}$$

*When an agreement occurs, the offer made to a non-proposer $j$ is $u^{-1}(\beta E[v_j(\tilde{S}')|S=s])$, and player $j$ accepts if and only if the share offered is greater than $u^{-1}(\beta E[v_j(\tilde{S}')|S=s])$.*

The proof follows from Theorem 1 in Merlo and Wilson (1995) and uses conditions in *CI-1* to show the ex ante individual SSPE continuation payoffs are independent from the order of moves in the current period, i.e. $E[v_i(\tilde{S}')|\tilde{S}] = E[v_i(\tilde{S}')|S]$. Identification of the utility function is possible if we exploit observations of the division of cakes observed, and if the utility function is restricted to belong to a particular class of utility functions.

*PA (Possibility of agreement) For all $x \in \Omega_X$ and $\kappa \equiv \rho_{(1)} \in \{1, 2, ., K\}$, $\Pr(D = 1|x, \kappa) > 0$.*

*PS (Parameter space) The parameter space for utility function (denoted $\Theta_U$) is such that (i) $u' > 0$, $u(0) = 0$ for all $u \in \Theta_U$; and (ii) for all $u, \tilde{u} \in \Theta_U$, $\tilde{u} = g \circ u$ where $g$ is a strictly concave or convex function (possibly depending on $u, \tilde{u}$).*

Assumption *PA* ensures that for all values of $x$ there is positive probability of reaching an immediate agreement. This is important, as the link between utilities and physical shares exists only when a unanimous agreement occurs. Assumption *PS* allows us to use Jensen's inequality repeatedly to prove by contradiction that the observational equivalence of two utility functions $u, \tilde{u}$ fails under the assumptions above.

**Proposition 6** *Suppose $\beta$ is known and CI-1,2,3, SE, MT, PA, PS hold. Then $\nexists u \neq \tilde{u}$ in $\Theta_U$ such that $u \overset{o.e.}{\sim} \tilde{u}$.*

A corollary of the proposition is that $u$ is identified within the classes of increasing functions with either constant absolute risk aversions ($CARA$) or constant relative risk aversions ($CRRA$) respectively and $u(0) = 0$. To see this, suppose $u_1, u_2$ are both differentiable $CARA$ functions with $u_2 = g \circ u_1$. Let $R_a(h) \equiv -\frac{h''(x)}{h'(x)}$ denote the absolute risk aversion for a function $h$. Then algebra shows $R_a(g) = R_a(u_2) - R_a(u_1)$. Both $R_a(u_2)$ and $R_a(u_1)$ are constant by our supposition, and $g' > 0$ by condition (i) in *PS*. Hence $g''$ must be either strictly positive or strictly negative over its whole support. It follows the class of increasing $CARA$ functions with $u(0) = 0$ satisfies *PS*. Likewise, we can also show $u$ is identified within the class of increasing $CRRA$ functions with $u(0) = 0$.

### 6.2 Heterogenous Discount Factors

Now consider another extension where each player $i$ in the bargaining game has a different discount factor $\beta_i$. The lemma below characterizes the SSPE payoffs in this case.

**Lemma 7** *Suppose CI-1 holds. Then $f \in F^K$ is a SSPE payoff if and only if $A(f) = f$ where for all $(s, \rho) \in \Omega_{S,\rho}$,*

$$\begin{aligned} A_i(f)(s,\rho) &\equiv \max\{c(s) - E[\textstyle\sum_{j \neq i} \beta_j f_j(S', \rho') | S = s] \ , \ \beta_i E[f_i(S', \rho') | S = s]\}, \text{ if } \rho_{(1)} = i \\ A_j(f)(s,\rho) &\equiv \beta_j E[f_j(S', \rho') | S = s], \text{ if } \rho_{(1)} \neq j \end{aligned}$$

The proof of this lemma follows from similar arguments in Theorem 1 in Merlo and Wilson (1998), and is omitted for brevity. With heterogenous discount factors, additional information from observed divisions of cakes under agreements must be exploited to recover individual $\beta_i$.

Theory predicts in any SSPE, a non-proposer always receives a share that is equal to his individual ex ante continuation payoff when an agreement is reached. Analogous to Lemma 3, the basic idea underlying the identification of individual $\beta_i$ is to show there exists a strictly monotone mapping between individual discount rates and observed shares for a non-proposer, *once the observable distributions of* $(Y, D, X)$ *are controlled for.*

**Proposition 7** *Under CI-1,2,3, SE, MT and PA, the discount factors* $\{\beta_i\}_{i=1}^{K}$ *are identified.*

As before, *MT* ensures there exists a one-to-one mapping between $Y$ and $\epsilon$. *PA* ensures that for any $i$, his share of the cake when an agreement is reached under someone else's proposal can be observed as a function of $x$. *SE* ensures the observable distribution of $(Y, D)$ is rationalized by a single SSPE, rather than a mixture of distributions rationalized in each of the multiple SSPE due to heterogenous $\beta_i$. Then the probability of agreements and agreed shares of the cake can still be related to discount rates as theory predicts in Lemma 7.[16]

Lemma 3 and Proposition 7 differ in that the former only uses the distribution of total cake sizes and the probability of unanimous agreements while the latter also exploits actual allocations under agreement. This difference comports with the intuition that when discount rates are heterogenous among players, econometricians need to exploit more information from observables to identify the vector of individual $\beta_i$'s. In fact, the homogenous $\beta$ in Lemma 3 is over-identified in the sense that observing the total cake size alone is sufficient for identifying the single $\beta$, while econometricians get to observe a $K$-vector of agreed non-proposer shares conditional on $X$. Each coordinate in the $K$-vector contains enough information for identifying $\beta$.

## 7 Conclusion

In this paper we present positive results in the identification of structural elements and counterfactual outcomes in stochastic sequential bargaining models under various scenarios of data availability. A unifying theme throughout the paper is that, in the absence of parametric assumptions on model structures, the model and its counterfactuals can still be point- or informatively partially-identified under fairly weak nonparametric restrictions (such as shape restrictions on the cake function or stochastic restrictions on the unobservable states), depending on data availability.

We conclude by mentioning some interesting directions for future research. First, in this paper, we have not addressed the definition of estimators or their asymptotic properties. Second, our

[16] *SE* implies there should be no variation in the size of shares for a fixed non-proposer and $x$. This testable implication can be easily verified by observed data.

point of departure in this paper is a group of conditional independence restrictions *CI-1,2,3*. Under these assumptions and conditional on current observable states, the cake sizes are independent of histories of states, and the order of moves in each period reveals no information about unobserved states or cake sizes. These assumptions are instrumental to our discussion of identification, but also imply specific restrictions on observable distributions.[17] Directions for future research includes identification when these conditional independence restrictions are relaxed, so that cake sizes or the agreed allocations are allowed to be correlated with the order of moves given states observed.

# 8 Appendix:

## 8.1 Part A: Proofs of Lemmas and Propositions

*Proof of Lemma 1.* It follows from Theorem 1 in Merlo and Wilson (1998) that the individual SSPE payoff is characterized by

$$\begin{aligned} A_i(f)(s,\rho) &\equiv \max\{c(s) - \beta E[\textstyle\sum_{j\neq i} f_j(\tilde{S}')|\tilde{S}=(s,\rho)]\ ,\ \beta E[f_i(\tilde{S}')|\tilde{S}=(s,\rho)]\}, \text{ if } \rho_{(1)} = i \\ A_j(f)(s,\rho) &\equiv \beta E[f_j(\tilde{S}')|\tilde{S}=(s,\rho)], \text{ if } \rho_{(1)} \neq j \end{aligned}$$

and from Theorem 2 in Merlo and Wilson (1998) that the total SSPE payoff must satisfy the fixed point equation $w(s,\rho) = \max\{c(s)\ ,\ \beta E[w(\tilde{S}')|\tilde{S}=(s,\rho)]\}$ for all $\tilde{s}=(s,\rho)$, and that agreement occurs for $\tilde{s}$ if and only if $c(s) \geq \beta E[w(\tilde{S}')|\tilde{S}=(s,\rho)]$. Note under *CI-1*, for any function $h$ of $(S,\rho)$,

$$\begin{aligned} & E[h(\tilde{S}')|\tilde{S}=(s,\rho)] \\ = & \int E[h(S',\rho')|S'=s',\tilde{S}=(s,\rho)]dH_{S'|S,\rho}(s'|s,\rho) \\ = & \int E[h(S',\rho')|s']dH_{S'|S,\rho}(s'|s,\rho) \\ = & \int E[h(S',\rho')|s']dH_{S'|S}(s'|s) = E[h(\tilde{S}')|S=s] \end{aligned}$$

where the first equality follows from the law of total probability, the second follows from condition (i) in *CI-1*, and the third follows from condition (ii) in *CI-1*. Then (a), (b) and (c) in the lemma follows. The uniqueness of SSPE payoffs is shown Theorem 3 in Merlo and Wilson (1998). $\square$

[17] For example, under *CI-1,2,3*, the ex ante individual continuation payoffs $\pi_i$ in SSPE must be independent of $\epsilon,\rho$ given $X$ in the canonical model. This implies whenever an agreement occurs with $i$ being proposer and with a fixed $x$ realized, the player $i$ must always proposes the same profile of shares to each of the other players. This limits the model's applicability in contexts where we do observe variations in proposals made by certain player to his rivals conditional on $x$.

*Proof of Lemma 2.* *Part (a)*: That players must exhaust the complete cake under agreements in SSPE is trivial. Under *CI-1,2*, $F_{S_{t+1},\rho_{t+1}|X^t,\rho^t} = \tilde{L}_{\rho_{t+1}|S_{t+1}} F_{S_{t+1}|X^t,\rho^t} = \tilde{L}_{\rho_{t+1}|S_{t+1}} F_{S_{t+1}|X_t}$, where the first equality follows from the first condition in *CI-1* and the second follows from the second condition in *CI-1* and both conditions in *CI-2*.[18] Note $Y_t$ is determined by $S_t$ only and, due to the "separation principle", $D_t$ is also determined by $S_t$ alone in SSPE. Therefore in SSPE, $(Y_{t+1}, D_{t+1}, X_{t+1}, \rho_{t+1})$ is determined by $(\epsilon_{t+1}, X_{t+1}, \rho_{t+1})$ and must be jointly independent of $(\rho^t, X^{t-1})$ conditional on $X_t$, and $F_{Y_{t+1},D_{t+1},\rho_{t+1},X_{t+1}|X^t,\rho^t} = F_{Y_{t+1},D_{t+1},\rho_{t+1}|X_{t+1},X_t} G_{X_{t+1}|X_t}$. But note $F_{\epsilon_{t+1},\rho_{t+1}|X_{t+1},X_t} = \tilde{L}_{\rho_{t+1}|S_{t+1}} F_{\epsilon_{t+1}|X_{t+1}}$ under *CI-1,2*. Hence $(Y_{t+1}, D_{t+1}, \rho_{t+1})$ must be independent of $X_t$ given $X_{t+1}$, and (2) holds under SSPE. To show (3), it suffices to note the division of the cake under a unanimous agreement $\eta_{t+1}$ is completely determined by $S_{t+1}$, $\rho_{t+1}$ and the parameters in $\theta$. (In fact only the identity of the proposer $\rho_{t+1,(1)}$ matters.) The time-homogeneity of $F_{Y_t,D_t,\rho_t|X_t}$ and $F_{\eta_t,\rho_t|D_t=1,X_t}$ follows from the time-homogeneity of $\tilde{L}_{\rho|S}$ and $H_{S'|S} = F_{\epsilon'|X'} G_{X'|X}$, which we maintain throughout the paper.

*Part (b)*: Given that the initial distribution of $X_0$ is identified, any $\theta, \theta'$ that generate the same $F_{\rho,D|X}$, $F_{\eta|D=1,X,\rho}$ must also by definition generate the same joint distribution of observables with $\tau = 0$. Now consider the case $\tau = 1$. Then

$$
\begin{aligned}
& \Pr(D_0 = 0, D_1 = 1, \rho_0, \rho_1, Y_0, X_1, \eta_1 | X_0) \\
= \; & \Pr(D_1 = 1, \rho_1, \eta_1 | X_1, D_0 = 0, X_0, Y_0, \rho_0) \Pr(X_1 | D_0 = 0, Y_0, \rho_0, X_0) \Pr(D_0 = 0, Y_0, \rho_0 | X_0) \\
= \; & \Pr(D_1 = 1, \rho_1, \eta_1 | X_1) G(X_1 | X_0) \Pr(D_0 = 0, Y_0, \rho_0 | X_0)
\end{aligned}
$$

where the second equality follows from *CI-1,2* again. Recall $G_{X'|X}$ is directly recovered from observables, and by our supposition, $\theta, \theta'$ generate the same conditional distributions $F_{\eta_1,D_1,\rho_1|X_1}$ and $F_{Y_0,D_0,\rho_0|X_0}$. Hence $\theta, \theta'$ induce the same joint distribution of observables with $\tau = 1$. We then complete the proof through induction. Suppose $\theta, \theta'$ generate the same observable distribution for $\tau \leq t$. Now consider the distribution with $\tau = t + 1$,

$$
\begin{aligned}
& \Pr(\textstyle\prod_{s=0}^t D_s = 0, D_{t+1} = 1, \rho^{t+1}, Y^t, X^{t+1}, \eta_{t+1} | X_0) \\
= \; & \Pr(D_{t+1} = 1, \rho_{t+1}, \eta_{t+1} | X_{t+1}, \textstyle\prod_{s=0}^t D_s = 0, \rho^t, Y^t, X^t) \Pr(X_{t+1} | \textstyle\prod_{s=0}^t D_s = 0, \rho^t, Y^t, X^t) * \\
& \Pr(\textstyle\prod_{s=0}^t D_s = 0, \rho^t, Y^t, X^t | X_0) \\
= \; & \Pr(D_{t+1} = 1, \rho_{t+1}, \eta_{t+1} | X_{t+1}) G(X_{t+1} | X_t) \Pr(\textstyle\prod_{s=0}^t D_s = 0, \rho^t, Y^t, X^t | X_0)
\end{aligned}
$$

under *CI-1,2*. By our supposition at the beginning of this induction step, $\theta, \theta'$ generate the same first and last terms in the product. $\square$

*Proof of Lemma 3.* Note $\Pr(D = 1|x) = \Pr\{Y \geq \beta\pi_w(x)|x\} = 1 - F_{Y|X}(\beta\pi_w(x))$, where $\pi_w(s) \equiv E[w(S')|S = s]$ is the ex ante total continuation payoff for all players in SSPE, which obviously

[18] Note under the second condition in CI-1, $F_{S_{t+1}|X_t,\epsilon_t,X^{t-1},\rho^t} = H_{S_{t+1}|S_t}$. Then note under CI-2, $H_{S_{t+1}|S_t} = F_{\epsilon_{t+1}|X_{t+1}} G_{X_{t+1}|X_t}$. Hence $S_{t+1}$ is independent of $\{\epsilon_t, X^{t-1}, \rho^t\}$ conditional on $X_t$

depends on structural elements $(\beta, c, F_{\epsilon|X}, G_{X'|X})$. Under *CI-1,2*, $\pi_w$ must be a function of $X$ only. Let $p(x) \equiv \Pr(D=1|x)$. By construction, $\pi_w(x)$ solves the fixed point equation:

$$\pi_w(x) = \int \max\{c(x', \varepsilon'), \beta\pi_w(x')\} dF_{\epsilon|X=x'}(\varepsilon') G(x'|x) \tag{22}$$

$$= \int \beta\pi_w(x') + \phi(c(x', \varepsilon'), \pi_w(x'); \beta, F_{\epsilon|X=x'}) dG(x'|x) \tag{23}$$

where

$$\phi(c(x,\varepsilon), \pi_w(x); \beta, F_{\epsilon|X=x}) \equiv \int \max\{c(x,\varepsilon) - \beta\pi_w(x), 0\} dF_{\epsilon|X=x}(\varepsilon)$$

But note $\beta\pi_w(x) = F^{-1}_{Y|X=x}(1-p(x))$ in SSPE since $F_{Y|X=x}$ is strictly increasing for all $x$ due to *MT*. Hence the function $\phi$ can be expressed in terms of observable distributions as

$$\tilde{\phi}(x; p, F_{Y|X}) \equiv \int_{F^{-1}_{Y|X=x}(1-p(x))} y - F^{-1}_{Y|X=x}(1-p(x)) dF_{Y|X=x}(y)$$

This in turn implies the true $\pi_w(x)$ can be alternatively represented as a solution to the following "quasi-structural" fixed-point equation: $f = T(f; \beta, p, F_{Y|X})$, where

$$T(f; \beta, p, F_{Y|X}) \equiv \int \beta f(x') + \tilde{\phi}(x'; p, F_{Y|X}) dG(x'|x) \tag{24}$$

The "quasi-" prefix is intended to highlight that structural elements $(\beta, c, F_{\epsilon|X}, G_{X'X})$ all enter (24) indirectly through observed $p$, $F_{Y|X}$. It is easy to see that with $\tilde{\phi}$ bounded and continuous, (i) $T(f+c) \leq Tf + \beta c$ for $\beta < 1$ and all $c \in \mathbb{R}^1_{++}$; and (ii) $f_1 \geq f_2$ implies $T(f_1) \geq T(f_2)$. Hence the operator $T$ is a contraction mapping when $\beta < 1$ with $\tilde{\phi}$. More generally, let $\hat{\pi}_w$ denote the unique solution for $\hat{\pi}_w = T(\hat{\pi}_w; \hat{\beta}, p, F_{Y|X})$, with $p, F_{Y|X}$ being observed from the data generating process (DGP) but $\hat{\beta}$ being any generic discount rate. That is, $\hat{\beta}$ may be any element in $(0,1)$ that differs from the true discount rate $\beta$. Recursive substitutions of (24) show that $\hat{\pi}_w(x; \hat{\beta}, p, F_{Y|X})$ must be strictly increasing in $\hat{\beta}$ with $p, F_{Y|X}$ fixed from data observed.[19] Hence $\hat{\beta}\hat{\pi}_w(x; \hat{\beta}, p, F_{Y|X})$ must also be strictly increasing in $\hat{\beta}$ for all $x$. It then follows that the true discount rate $\beta$ is identified as

$$\inf\{\hat{\beta} : \hat{\beta}\hat{\pi}_w(x; \hat{\beta}, p, F_{Y|X}) \geq F^{-1}_{Y|X}(1-p(x))\}$$

where we have also used the fact that $F_{Y|X}$ is strictly increasing in $y$ for all $x$ under *MT*. $\square$

[19] To see this, note (24) can be written as

$$\hat{\pi}_w = (I - \hat{\beta}G)^{-1} G\tilde{\phi}$$

where $Gf \equiv \int f(x') dG_{X'|X}$ and $If = f$ for bounded continuous $f$. That $I - \hat{\beta}G$ is invertible follows from the fact that $G$ is a Markov transition and $\hat{\beta} \in (0,1)$. Then

$$\hat{\pi}_w = \left(I + \lim_{N\to+\infty} \sum_{n=1}^{N} \hat{\beta}^n G^n\right) G\tilde{\phi}$$

and $\hat{\pi}_w$ must be increasing in $\hat{\beta}$ with $\tilde{\phi}$ fixed from observable distributions and known to be positive by definition.

*Proof of Proposition 1.* Necessity follows from the definition of observational equivalence. For sufficiency, results from Lemma 2 suggests it suffices to show $F_{Y,D|X}(\sigma) = F_{Y,D|X}(\sigma')$ and $F_{\eta|D=1,X,\rho}(\sigma) = F_{\eta|D=1,X,\rho}(\sigma')$ a.e. on $\Omega_X$ if $F_{Y|X}(\sigma) = F_{Y|X}(\sigma')$ a.e. on $\Omega_X$. (Recall $\rho$ is independent of $\epsilon$ conditional on $X$, and $L_{\rho|X}$ is directly identified from data.) Let $\pi \equiv (\pi_1, ., \pi_K)$ denote the vector of ex ante continuation payoffs for players $(1, ., K)$ in SSPE by

$$\pi_i(\tilde{s};\sigma) \equiv E[v_i(\tilde{S}')|\tilde{S} = (s,\rho)] =$$
$$= \left\{ \begin{array}{c} E(v_i(\tilde{S}')|\tilde{S}' \in \tilde{S}(\mu), \kappa' = i, s, \rho) \Pr(\tilde{S}' \in \tilde{S}(\mu), \kappa' = i|s,\rho) + \\ E(v_i(\tilde{S}')|\tilde{S}' \notin \tilde{S}(\mu), \kappa' = i, s, \rho) \Pr(\tilde{S}' \notin \tilde{S}(\mu), \kappa' = i|s,\rho) + \\ E(v_i(\tilde{S}')|\kappa' \neq i, s, \rho) \Pr(\kappa' \neq i|s,\rho) \end{array} \right\} \tag{25}$$

where $v_i$ is player $i$'s SSPE payoff (with its dependence on $\sigma$ suppressed for notational ease), $\kappa'$ is the identity of the proposer in the next period and $\tilde{S}(\mu)$ denotes the subset of $\Omega_{S,\rho}$ such that $c(s) \geq \beta \sum_i \pi_i(s,\rho)$. Our proof of sufficiency uses Lemma A1 below.

*Lemma A1: Suppose CI-1,2,3 hold. Then (i) the vector of ex ante continuation payoff in SSPE $\pi$ is independent of unobserved states $\epsilon$ and the order of moves $\rho$ given $x$ (i.e. $\pi(s,\rho;\sigma) = \pi(x;\sigma)$) for all $s \in \Omega_S$ and any $\sigma$; (ii) for any $\sigma, \sigma'$ such that $F_{Y|X}(\sigma) = F_{Y|X}(\sigma')$ a.e. on $\Omega_X$, $\pi(x;\sigma) = \pi(x;\sigma')$ a.e. on $\Omega_X$.*

*Proof:* To prove (i), note $\pi$ is the solution to the fixed point equation $\pi = \Psi(\pi;\sigma)$, where the $i$-th coordinate of $\Psi$ is given by

$$\Psi_i(s,\rho;\pi) \equiv \left\{ \begin{array}{c} E\left[\left(c(S') - \beta \sum_{j\neq i} \pi_j(\tilde{S}')\right) 1\left(\tilde{S}' \in \tilde{S}(\mu), \kappa' = i\right) |s,\rho\right] + \\ E[\beta\pi_i(\tilde{S}')1(\tilde{S}' \notin \tilde{S}(\mu), \kappa' = i)|s,\rho] + E(\beta\pi_i(\tilde{S}')1(\kappa' \neq i)|s,\rho) \end{array} \right\} \tag{26}$$

where $1(\Sigma)$ is an indicator function that is equal to 1 if and only if the event $\Sigma$ is true. Under *CI-1,2*, $(S',\rho')$ is independent of $\epsilon$ and $\rho$ conditional on $X$.[20] Then it follows for any function $h$ of $(S,\rho)$,

$$E\left[h(S',\rho')|S = s,\rho\right] = E\left[h(S',\rho')|X = x\right]$$

Hence $\Psi$ only maps into the space of functions that are independent of $\epsilon$ and the order of moves $\rho$ conditional on $x$. The usual fixed point arguments show solutions to $\pi = \Psi(\pi;\sigma)$ exist and must be independent of $\epsilon$ given $x$ (i.e. $\pi$ must be a function of $x$ but does not involve $\epsilon$ or $\rho$).

To prove (ii), let $q_i(x) \equiv \Pr(\kappa = i|x)$ (which is directly identified from observables) and note

[20] To see this, it suffices to note under *CI-1,2*, we have (with a slight abuse of notation):

$$F(S',\rho'|S,\rho) = F(\rho'|S',S,\rho)F(S'|S,\rho) = \tilde{L}(\rho'|S')F(\epsilon'|X')G(X'|X)$$

the result in (i) implies components in (26) can be represented as

$$
\begin{aligned}
& E\left[\left(c(S') - \beta \textstyle\sum_{j\neq i} \pi_j(\tilde{S}')\right) 1\left(c(S') \geq \beta \textstyle\sum_{j=1}^{K} \pi_j(\tilde{S}'), \kappa' = i\right) |s\right] \\
= \; & E\left[\left(Y' - \beta \textstyle\sum_{j\neq i} \pi_j(X')\right) 1(Y' \geq \beta \textstyle\sum_{j=1}^{K} \pi_j(X'))1\left(\kappa' = i\right) |x\right] \\
= \; & \int_{\Omega_{X'|x}} q_i(x') \int_{\Omega_{Y'|x'}} 1\left(y' \geq \beta \textstyle\sum_{j=1}^{K} \pi_j(x')\right)\left(y' - \beta \textstyle\sum_{j\neq i} \pi_j(x')\right) dF_{Y|X}(y'|x')dG(x'|x)
\end{aligned}
$$

where the first equality follows from $Y = c(S)$ and *CI-1,2* and *MT* as before and the second follows from *CI-3.* Besides,

$$
\begin{aligned}
& E\left[\beta\pi_i(\tilde{S}')1\left(c(S') < \beta \textstyle\sum_{j=1}^{K} \pi_j(\tilde{S}'), \kappa' = i\right) |s\right] \\
= \; & \beta E\left[\pi_i(X')1\left(Y' < \beta \textstyle\sum_{j=1}^{K} \pi_j(X')\right) 1\left(\kappa' = i\right) |x\right] \\
= \; & \beta \int_{\Omega_{X'|x}} \pi_i(x')q_i(x') \int_{\Omega_{Y'|x'}} 1\left(Y' < \beta \textstyle\sum_{j=1}^{K} \pi_j(X')\right) dF_{Y|X}(y'|x')dG(x'|x)
\end{aligned}
$$

where the first equality follows from *CI-1,2* as above and the second from *CI-3.* Finally,

$$
E(\beta\pi_i(\tilde{S}')1(\kappa' \neq i)|s) = \beta \int_{\Omega_{X'|x}} \pi_i(x')(1 - q_i(x'))dG(x'|x)
$$

Hence given any $\beta$ and with $G_{X'|X}, L_{\rho|X}$ (and therefore $\{q_i(.)\}_{i=1}^{k}$) directly recovered from observables in data, the ex ante continuation payoffs $\pi$ depend on $c$ and $F_{\epsilon|X}$ only through $F_{Y|X}$, as is seen in the derivation of the fixed point equation $\pi = \Psi(\pi; \beta, G_{X'|X}, F_{Y|X}, L_{\rho|X})$ above. Therefore for any pair $\sigma, \sigma'$ such that $F_{Y|X}(\sigma) = F_{Y|X}(\sigma')$ a.e. on $\Omega_X$, we have

$$
\pi(x; \sigma) = \pi(x; \sigma')
$$

a.e. on $\Omega_X$. This completes the proof of *Lemma A1.* *Q.E.D.*

Recall for any $y \in \mathbb{R}^1_+$, $X = x$,

$$
\begin{aligned}
& \Pr(D = 1|Y \leq y, X = x) \\
= \; & \Pr(Y \geq \beta \textstyle\sum_{j=1}^{K} \pi_j(X; \beta, G_{X'|X}, F_{Y|X}, L_{\rho|X})|Y \leq y, X = x)
\end{aligned}
$$

Hence this conditional probability of reaching an agreement is a functional of $(\beta, G_{X'|X}, F_{Y|X}, L_{\rho|X})$ only. Therefore for any $\sigma, \sigma'$ such that $F_{Y|X}(\sigma) = F_{Y|X}(\sigma')$ a.e. on $\Omega_X$, we also have $F_{Y,D|X}(\sigma) = F_{Y,D|X}(\sigma')$ a.e. on $\Omega_X$. It only remains to show $F_{\eta,Y|D=1,X,\rho}(\sigma) = F_{\eta,Y|D=1,X,\rho}(\sigma')$ a.e. on $\Omega_X$. But whenever agreements occur, the proposer $i$ offers $\beta E[v_j(S')|S = s] \equiv \beta\pi_j(x)$ to others, while claiming $c(s) - \beta \sum_{j\neq i} \pi_j(x)$ to himself. Therefore by Lemma A1, for any fixed $x$ and with $L_{\rho|X}$ observed directly from data, the joint distribution of the cake size and agreed allocations to all players are the same for any $\sigma$ and $\sigma'$ such that $F_{Y|X}(\sigma) = F_{Y|X}(\sigma')$ a.e. on $\Omega_X$. $\square$

*Proof of Proposition 2.* That $\beta$ is identified follows from Lemma 3 and the fact that *AS, MH* implies *CI-2, MT*. Since *MH* implies conditional independence of $\epsilon$ from histories given $X$, we have

$$\pi_w(s) \equiv E[w(S')|S=s] = E[w(S')|x] = \int \tilde{w}(x')dG(x'|x) \tag{27}$$

where $\tilde{w}(x) \equiv E[w(S)|X=x]$. Note $G_{X'|X}$ is directly recovered from observables. With knowledge of $\beta$, $\tilde{w}(x)$ can also be identified from data as $E[\beta^\tau \sum_i \eta_{\tau,i}|X=x]$, where $\tau$ denotes the number of periods it takes to reach an agreement after state $x$ is realized, and $\eta_{\tau,i} = v_i(s_\tau, \rho_\tau)$ denotes the proposed allocation for player $i$ accepted in state $s_\tau$. Note $\sum_i v_i(s,\rho) = c(s)$ when agreement occurs under $(s,\rho)$, since agreed proposals always exhaust the cake size. Therefore the right-hand side of (27) is identified.

Since cake sizes are not reported when proposals are rejected, only censored values of $c(s)$, denoted as $y = \max\{\beta E[\tilde{w}(X')|X=x], c(s)\}$, is observed. Then the gains to the proposer in SSPE is:

$$\begin{aligned} y^* &\equiv y - \beta E[\tilde{w}(X')|X=x] \\ &= \max\{\tilde{c}(x) - \beta E[\tilde{w}(X')|X=x] + \varepsilon, 0\} \end{aligned}$$

Note $\pi_w(s)$ is a function of $x$ but not $\varepsilon$ under *CI-1* and *MH*. Therefore the ex ante total continuation payoff $\pi_w$ is the unique fixed point of a contraction mapping

$$\pi_w(x) = \int \max\{\tilde{c}(x') + \varepsilon', \beta\pi_w(x')\} dF_{\epsilon|X}(\varepsilon'|x')dG(x'|x)$$

And it maps from the space of continuous, bounded functions to itself. Hence $\pi_w$ must also be continuous and bounded over $\Omega_X$, and so is $\gamma$. The rest of proof of identification follows from Chen and Khan (2005), which relies on SG. To see SG is testable using observed distributions, note the event "$\gamma(x) > 0$" is equivalent to "$\Pr\{\epsilon > -\gamma(x)|x\} > \frac{1}{2}$" since $Median(\epsilon|x) = 0$ for all $x \in \Omega_X$. Hence SG is equivalent to $\Pr\{\Pr(D=1|X) > \frac{1}{2}\} > 0$. $\square$

*Proof of Proposition 3.* Let $q_\alpha(Y^*|x)$ denote the $\alpha$-th quantile of $Y^*$ given $x$. For $x \in \Omega_X^+$, $\gamma(x) = q_{1/2}(Y^*|x)$ is identified. Consider any $x = (x_0, x_1)$ s.t. $x \notin \Omega_X^+$. By RG and RS, $\gamma$ is bounded over $\Omega_X$ while support of unobservable states $\Omega_\epsilon$ given $x_0$ is unbounded. Then $\exists \bar{\alpha} > 1/2$ and close enough to 1 s.t. $q_{\bar{\alpha}}(Y^*|x) = \gamma(x) + c_{\bar{\alpha}}(x) > 0$ is observed, where $c_{\bar{\alpha}}(x) = c_{\bar{\alpha}}(x_0)$ denotes the $\bar{\alpha}$-th quantile of $\epsilon$ conditional on $x \equiv (x_0, x_1)$, and is independent of $x_1$ by *ER*. Now pick $\tilde{x} = (x_0, \tilde{x}_1)$ such that $q_{1/2}(Y^*|\tilde{x}) = \gamma(\tilde{x}) > 0$ is observable. Such a choice is possible because of SG2. Hence $\gamma(\tilde{x})$ is identified. Since $\bar{\alpha} > 1/2$ and $c_{\bar{\alpha}}(x) = c_{\bar{\alpha}}(\tilde{x})$ under ER, $q_{\bar{\alpha}}(Y^*|\tilde{x})$ must also be positive, observable, and equal to $\gamma(\tilde{x})+ c_{\bar{\alpha}}(x_0)$. Hence (with a slight abuse of notation) $c_{\bar{\alpha}}(x_0) = c_{\bar{\alpha}}(\tilde{x})$ is recovered as $q_{\bar{\alpha}}(Y^*|\tilde{x})- \gamma(\tilde{x})$. This implies $\gamma(x)$ can then be recovered as $q_{\bar{\alpha}}(Y^*|x) - c_{\bar{\alpha}}(\tilde{x})$ for any $x \notin \Omega_X$. $\square$

*Proof of Lemma 4.* Suppose the model is correctly specified for some $(\tilde{c}_1(x), F^1_{\epsilon|X})$. Consider an alternative specification $(\tilde{c}_2(x), F^2_{\epsilon|X})$ where $\tilde{c}_2(x) = a\tilde{c}_1(x) + b$ and $F^2_{\epsilon|X}(t) = F^1_{\epsilon|X}(\frac{t-b}{a})$ for some constants $a > 0$ and $b$. Lemma 1 showed total payoff $w$ in SSPE must satisfy: $w(s) = \max\{c(s), \beta E[w(S')|s]\}$ and an agreement is reached in state $s$ if and only if $c(s) \geq \beta E[w(S')|s]$. Let $\pi_w(s) \equiv \sum_i \pi_i(s) = E[w(S')|s]$. By *CI-1,2*, $\pi_w$ must be independent of $\varepsilon$ for all $x$, and

$$\pi_w(x) \equiv \int \max\{c(s'), \beta E[w(S'')|s']\} dF_{\epsilon|X}(\varepsilon'|x')dG(x'|x)$$

Under our supposition of the model specification $c(s) = \tilde{c}_1(x) - \varepsilon$ and $F^1_{\epsilon|X}$, this suggests SSPE continuational payoff under the first pair of specifications, denoted $\pi_{w,1}$, must solve the fixed point equation in $f$:

$$f(x) = \int \max\{\tilde{c}_1(x') - \varepsilon', \beta f(x')\} dF^1_{\epsilon|X}(\varepsilon'|x')dG(x'|x)$$

Some algebra using change-of-variables shows the fixed point solutions under the alternative specification $(\tilde{c}_2, F^2_{\epsilon|X})$ is $\pi_{w,2}(x) = a\pi_{w,1}(x)$. Therefore

$$\Pr(D = 1|x; \tilde{c}_2, F^2_{\epsilon|X}) = F^2_{\epsilon|X}(\tilde{c}_2(x) - \beta\pi_{w,2}(x)|x) = F^1_{\epsilon|X}(\tilde{c}_1(x) - \beta\pi_{w,1}(x)|x) = \Pr(D = 1|x; \tilde{c}_1, F^1_{\epsilon|X})$$

□

*Proof of Proposition 4.* To prove (i), first recall by Lemma 1, the total SSPE payoff $w$ must satisfy: $w(s) = \max\{c(s), \beta E[w(S')|s]\}$ and an agreement is reached in state $s$ if and only if $c(s) \geq \beta E[w(S')|s]$. Let $\pi_w(s) \equiv \sum_i \pi_i(s) = E[w(S')|s]$ as before. By conditional independence, $\pi_w$ must be independent of $\epsilon$ for all $x$. Note

$$\pi_w(x) \equiv \int \max\{c(s'), \beta E[w(S'')|s']\} dF_{\epsilon|X}(\varepsilon'|x')dG(x'|x)$$

Under *AS*, $c(s) = \tilde{c}(x) - \varepsilon$, and the formula can be written as

$$\begin{aligned}\pi_w(x) &\equiv \int \max\{\tilde{c}(x') - \varepsilon', \beta\pi_w(x')\} dF_{\epsilon|X}(\varepsilon'|x')dG(x'|x) \\ &= \int \beta\pi_w(x') + \int \max\{\tilde{c}(x') - \beta\pi_w(x') - \varepsilon', 0\} dF_{\epsilon|X}(\varepsilon'|x')dG(x'|x)\end{aligned}$$

Define $p(x) \equiv \Pr(D = 1|x) = \Pr(c(S) \geq \beta\pi_w(X)|x) = F_{\epsilon|X=x}(\tilde{c}(x) - \beta\pi_w(x))$. Hence $\tilde{c}(x) - \beta\pi_w(x) = F^{-1}_{\epsilon|X=x}(p(x))$. Define the mapping

$$\phi(x; p, F_{\epsilon|X}) \equiv \int \max\{F^{-1}_{\epsilon|X=x}(p(x)) - \varepsilon, 0\} dF_{\epsilon|X}(\varepsilon|x)$$

Then

$$\pi_w(x) = \int \beta\pi_w(x') + \phi(x'; p, F_{\epsilon|X})dG(x'|x) \tag{28}$$

For any given $F_{\epsilon|X}$ and $p$ observed, the right hand side of (28) is a contraction mapping by standard arguments in the BS Theorem. Therefore with knowledge of $F_{\epsilon|X}$, $p$ and $G$, $\pi_w$ is uniquely recovered as the solution to the fixed point equation in (28). Hence $\tilde{c}$ is identified with knowledge of $F_{\epsilon|X}$. The proof of (ii) follows almost immediately. When $\beta$ is unknown, $\pi_w$ can be written as a monotone function of $\beta$. We can solve the following equation given knowledge of $F_{\epsilon|X}$, $p$ and $\tilde{c}(\bar{x})$,

$$\beta\pi_w(\bar{x};\beta) = \tilde{c}(\bar{x}) - F^{-1}_{\epsilon|X=\bar{x}}(p(\bar{x}))$$

It then follows from the proof of (ii) that with knowledge of $\beta$, the component in cake function $\tilde{c}$ can be identified. $\square$

*Proof of Lemma 5.* (*Necessity*) Suppose a vector $p$ is generated by some true parameters $(\tilde{c}, F_\epsilon)$ underlying the DGP such that $A\tilde{C} > 0$ and $\epsilon$ is independent of $X$ with median 0. Then let $\tilde{Q}_m = F_\epsilon^{-1}(p_m)$ and $\tilde{\Phi}_m = \phi(p(x_m); F_\epsilon)$. It follows immediately from the substitution of (6) into (5), the independence of $\epsilon$ from $X$ and the monotonicity of $F_\epsilon$ that (7) and (8) must hold for $\tilde{Q}, \tilde{\Phi}$. The definition of $\phi$ and some straightforward algebra (involving the Leibniz rule for differentiating integrals) suggest for any $m, n$,

$$\phi(p_m) - \phi(p_n) = \int_{\tilde{Q}_n}^{\tilde{Q}_m} F_\epsilon(\varepsilon)d\varepsilon$$

which must be bounded between $p_n(\tilde{Q}_m - \tilde{Q}_n)$ and $p_m(\tilde{Q}_m - \tilde{Q}_n)$. Hence (9) holds for $\tilde{Q}, \tilde{\Phi}$. Note (11) holds for $\tilde{\Phi}$ by definition of the CSF $\phi$, and (10) holds for $\tilde{Q}, \tilde{\Phi}$ if $\bar{\phi}$ is equal to the true CSF at $\frac{1}{2}$, i.e. $\tilde{\phi} \equiv \phi(1/2; F_\epsilon)$. More generally, if $\bar{\phi} \neq \tilde{\phi}$, the system (7)-(11) still holds for the scale multiplications $(\bar{\phi}/\tilde{\phi})\tilde{Q}$ and $(\bar{\phi}/\tilde{\phi})\tilde{\Phi}$. *(Sufficiency)* We need to show that if (7)-(11) holds for some $Q, \Phi$ then there must be a pair $(\tilde{c}, F_\epsilon)$ such that (i) $\epsilon$ is independent of $X$ and $\tilde{c}$ satisfies the shape restrictions; and (ii) $(\tilde{c}, F_\epsilon)$ generates $p$ as the decision maker's dynamic rational choice probabilities. By supposition the linear system is feasible. Hence we can find such a $F_\epsilon$ by choosing the $p_m$-percentile $F_\epsilon^{-1}(p_m)$ to be the solutions $Q_m$ and choosing $\phi(p(x_m))$ by first setting $\phi(1/2) = \bar{\phi}$ and then interpolating between $F_\epsilon^{-1}(p_m)$ so that $\phi(p(x_m))$ is equal to the solution $\Phi_m$. This is possible because the inequality restrictions (9) and (10) are satisfied. A distribution constructed this way naturally satisfies the independence of $X$ and $Median(\epsilon) = 0$ due to the definition of the linear system (7)-(11). Then define $\tilde{C} = Q + \beta(I - \beta G)^{-1}G\Phi$ and the pair $(\tilde{c}, F_\epsilon)$ satisfies both requirements (i) and (ii) above. $\square$

*Proof of Proposition 5.* The distribution of unobservable states $F_\epsilon^0$ is fixed in both the observed and the counterfactual environments. It suffices to note that in type (a) counterfactual exercise, $\tilde{C} = Q^j + \beta(I - \beta G^j)^{-1}G^j\Phi^j$ for $j = 0, 1$. And in type (b) counterfactual exercise, $\tilde{C} = Q^0 + \beta(I - \beta G^0)^{-1}G^0\Phi^0$ while $\alpha\tilde{C} = Q^1 + \beta(I - \beta G^1)^{-1}G^1\Phi^1$. The rest of the proof follows from similar arguments in Lemma 5 and is omitted for brevity. $\square$

*Proof of Proposition 6.* For any player $i$, let $v_i(s, \rho)$ denote his SSPE payoffs conditional on $s, \rho$.

The *ex ante* individual continuation payoff for $i$ in SSPE is:

$$\begin{aligned} & \pi_i(s,\rho) \equiv E[v_i(\tilde{S}')|\tilde{S}=(s,\rho)] \\ = & \left\{ \begin{array}{l} E[v_i(S',\rho')1(D'=1,\kappa'=i)|s,\rho]+ \\ E[v_i(S',\rho')1(D'=0,\kappa'=i)|s,\rho]+E[v_i(S',\rho')1(\kappa'\neq i)|s,\rho] \end{array} \right\} \end{aligned} \tag{29}$$

where $D_t$ is the dummy for an agreement in period $t$ and it is a function of $S_t$ but not $\rho_t$. Under *CI-1,2*,

$$F(S',\rho'|S,\rho)=\tilde{L}(\rho'|S')F(\epsilon'|X')G(X'|X)$$

That is, $S',\rho'$ are independent of $\epsilon,\rho$ conditional on $X$. Therefore $\pi$ is a function of $X$ only. Hence (29) can be written as

$$\begin{aligned} \pi_i(x) &= E\left[u\left(c(S')-\textstyle\sum_{j\neq i}u^{-1}(\beta\pi_j(X'))\right)1(D'=1,\kappa'=i)|x\right]+ \\ & \quad E\left[\beta\pi_i(X')1(D'=0,\kappa'=i)|x\right]+E\left[\beta\pi_i(X')1(\kappa'\neq i)|x\right] \end{aligned}$$

Let $q_i(x)\equiv \Pr(\kappa=i|x)$. Note the first term on the right-hand side can be written as

$$\begin{aligned} & \int u\left(c(S')-\textstyle\sum_{j\neq i}u^{-1}(\beta\pi_j(X'))\right)1(D'=1,\kappa'=i)dF_{S',\kappa'|X=x} \\ = & \int\int u\left(c(S')-\textstyle\sum_{j\neq i}u^{-1}(\beta\pi_j(X'))\right)1\left(c(S')\geq\sum_{j=1}^{K}u^{-1}(\beta\pi_j(X'))\right)1(\kappa'=i)dF_{\epsilon',\kappa'|X',X}dG_{X'|X=x} \\ = & \int q_i(X')\int u\left(c(S')-\textstyle\sum_{j\neq i}u^{-1}(\beta\pi_j(X'))\right)1\left(c(S')\geq\sum_{j=1}^{K}u^{-1}(\beta\pi_j(X'))\right)dF_{\epsilon'|X'}dG_{X'|X=x} \end{aligned}$$

where the first equality follows from the law of iterated expectations and the second equality follows from the fact that under *CI-1,2,3:*

$$F_{\epsilon',\kappa'|X',X}=F_{\kappa'|S',X}F_{\epsilon'|X',X}=F_{\kappa'|X'}F_{\epsilon'|X'}$$

Furthermore, under MT, $Y$ is increasing in $\epsilon$ given $X$. Hence

$$\begin{aligned} & \int u\left(c(S)-\textstyle\sum_{j\neq i}u^{-1}(\beta\pi_j(X))\right)1\left(c(S)\geq\sum_{j=1}^{K}u^{-1}(\beta\pi_j(X))\right)dF_{\epsilon|X} \\ = & \int u\left(Y-\textstyle\sum_{j\neq i}u^{-1}(\beta\pi_j(X))\right)1\left(Y\geq\sum_{j=1}^{K}u^{-1}(\beta\pi_j(X))\right)dF_{Y|X} \\ = & \int_{\sum_i u^{-1}(\beta\pi_i(X))}u\left(Y-\textstyle\sum_{j\neq i}u^{-1}(\beta\pi_j(X))\right)dF_{Y|X} \\ \equiv & \ \lambda(X;u,\pi_{-i},F_{Y|X}) \end{aligned}$$

Likewise, the second term can be written as

$$\begin{aligned} & \int\beta\pi_i(X')1(D'=0,\kappa'=i)dF_{S',\kappa'|X=x} \\ = & \int\beta\pi_i(X')\int 1(D'=0,\kappa'=i)dF_{\epsilon',\kappa'|X'}dG_{X'|X=x} \\ = & \int\beta\pi_i(X')q_i(X')\int 1\{Y'<\textstyle\sum_i u^{-1}(\beta\pi_i(X'))\}dF_{Y'|X'}dG_{X'|X=x} \end{aligned}$$

And the third term is

$$E\left[\beta\pi_i(X')1(\kappa' \neq i)|x\right] = \int \beta\pi_i(X')(1-q_i(X'))dG_{X'|X=x}$$

Hence we can write $\pi = \Psi(\pi; u, \beta, G_{X'|X}, F_{Y|X})$ where $\Psi$ is a $\mathbb{R}^k$-valued function with the $i$-th coordinate $\Psi_i$ defined as

$$\begin{aligned}\Psi_i(x;\pi) \quad \equiv \quad & \int q_i(x')\lambda(x';u,\pi_{-i})dG(x'|x) + \int(1-q_i(x'))\beta\pi_i(x')dG(x'|x) + \\ & \int q_i(x')\beta\pi_i(x')\int 1\{y' < \textstyle\sum_i u^{-1}(\beta\pi_i(x'))\}dF_{Y|X}(y'|x')dG(x'|x)\end{aligned} \tag{30}$$

For notational ease, we suppress dependence of the fixed point equation on $(\beta, G_{X'|X}, F_{Y|X})$. Define the physical share of the cake for a non-proposer $i$ when an agreement occurs in state $x$ as $\psi_i(x) = u^{-1}(\beta\pi_i(x))$. The assumption PA implies that for each individual $i$ and observable state $x$, there is positive probability that an agreement is reached when $i$ is not the proposer. Hence for each player $i$, $\psi_i(x)$ is observed over the support $\Omega_X$ as the physical shares for player $i$ when agreements occur and $\kappa \neq i$ in state $s = (x, \varepsilon)$.

Define $y_i^* \equiv y - \sum_{j\neq i}\psi_j(x)$ for all $i$. Alternatively, this can be written as

$$\beta\pi_i(x) = \beta\int q_i(x')p_1(x')\bar{\lambda}_i(x') + \left[q_i(x')p_0(x') + 1 - q_i(x')\right]\beta\pi_i(x')dG(x'|x) \tag{31}$$

where $p_1(x) \equiv \Pr(D=1|x)$ and $p_0(x) \equiv 1 - p_1(x)$ and

$$\bar{\lambda}_i(x;u) \equiv \int_{\psi_i(x)}^{+\infty} u(t)dF_{Y_i^*|X,D=1}(t|x) = E[u(Y_i^*)|D=1,x]$$

We refer to (31) as a "quasi-fixed-point equation" for $\beta\pi_i(x)$. Compared with (30), (31) differs in that it explicitly expresses how $\{\psi_i\}_{i=1}^k$, $p_1$ and $F_{Y^*|X,D=1}$ enter the fixed point equation. Though dependent upon the unknown true utility function $u$, these three functions are observable from data and therefore are held fixed in identification arguments.

We prove by contradiction. Suppose there exists $u \neq \tilde{u}$ in $\Theta_U$ and $u \overset{o.e.}{\sim} \tilde{u}$. Let $\pi, \tilde{\pi}$ denote solutions to fixed point equations corresponding to $u, \tilde{u}$ respectively

$$\pi = \Psi(\pi; u)\ ;\ \tilde{\pi} = \Psi(\tilde{\pi}; \tilde{u})$$

By supposition of observational equivalence of $u$ and $\tilde{u}$, we have for all $i$ and almost everywhere on $\Omega_X$,

$$\psi_i(x;u) \equiv u^{-1}(\beta\pi_i(x;u)) = \tilde{u}^{-1}(\beta\tilde{\pi}_i(x;\tilde{u})) \equiv \psi_i(x;\tilde{u}) \tag{32}$$

$$p_1(x;u) \equiv \Pr(D=1|x;u) = \Pr(D=1|x;\tilde{u}) \equiv p_1(x;\tilde{u}) \tag{33}$$

It follows that for the distribution of cake size $F_{Y|X}$ observed, the same conditional distribution $F_{Y^*|X,D}$ is induced by both $u, \tilde{u}$. Suppose $\tilde{u} = g \circ u$ for some *strictly* concave function $g : \mathbb{R}^1_+ \to \mathbb{R}^1_+$. Then $\bar{\lambda}_i(x; \tilde{u}) = \bar{\lambda}_i(x; g \circ u) < g \circ \bar{\lambda}_i(x; u)$ by concavity of $g$ and the Jensen's Inequality. Also note $\psi_i(x; u) = u^{-1}(\beta \pi_i(x; u))$. Therefore for $u \overset{o.e.}{\sim} \tilde{u}$,

$$\begin{aligned}
& \tilde{u}(\psi_i(x;\tilde{u})) \\
= \;& \beta \int q_i(x')p_1(x';\tilde{u})\bar{\lambda}_i(x';\tilde{u}) + \left[q_i(x')p_0(x';\tilde{u}) + 1 - q_i(x')\right] \tilde{u}(\psi_i(x';\tilde{u}))dG(x'|x) \\
< \;& \beta \int q_i(x')p_1(x';u) g \circ \bar{\lambda}_i(x';u) + \left[q_i(x')p_0(x';u) + 1 - q_i(x')\right] g \circ u(\psi_i(x';u))dG(x'|x) \\
< \;& \beta \int g\left[q_i(x')p_1(x';u)\bar{\lambda}_i(x';u) + \left(q_i(x')p_0(x';u) + 1 - q_i(x')\right) u(\psi_i(x';u))\right] dG(x'|x) \\
< \;& \beta g\left(\int q_i(x')p_1(x';u)\bar{\lambda}_i(x';u) + \left(q_i(x')p_0(x';u) + 1 - q_i(x')\right) u(\psi_i(x';u))dG(x'|x)\right) \\
< \;& g\left(\beta \int q_i(x')p_1(x';u)\bar{\lambda}_i(x';u) + \left(q_i(x')p_0(x';u) + 1 - q_i(x')\right) u(\psi_i(x';u))dG(x'|x)\right) \\
= \;& g \circ u(\psi_i(x;u)) = \tilde{u}(\psi_i(x;u)) = \tilde{u}(\psi_i(x;\tilde{u}))
\end{aligned}$$

where the inequalities all follow from concavity of $g$ and applications of Jensen's Inequality as well as (32) and (33). In addition, the last inequality also uses $g(0) = 0$ (implied by $u(0) = 0$ for all $u \in \Theta_U$). This constitutes a contraction. The proof for the case with $\tilde{u} = h \circ u$ for some strictly convex function $h$ follows from symmetric arguments and is omitted for brevity. Hence $\nexists u \neq \tilde{u}$ in $\Theta_u$ such that $u \overset{o.e.}{\sim} \tilde{u}$. $\square$

*Proof of Proposition 7.* Let $\overrightarrow{R}$ denote a generic vector in $\mathbb{R}^K$, with the $i$-th coordinate $R_i$ corresponding to player $i$. Let $\overrightarrow{v} \equiv (v_j)_{j=1}^K$ be the vector of individual SSPE payoffs that solve the fixed-point equation in Lemma 7. Note under *CI-1,2*,

$$E[v_i(S', \rho')|S = (x, \varepsilon)] = E[v_i(S', \rho')|X = x]$$

Under *PA*, for all $i$, there is positive probability that $i$ is not the proposer and an agreement is reached. This implies for all $i$ and $x$, the discounted individual ex ante continuation payoff $\beta_i \pi_i(x) \equiv \beta_i E[v_i(S', \rho')|X = x]$ can be observed as $\psi_i(x)$, i.e. the physical division of the cake received by $i$ when $i$ is not the proposer and an agreement is reached under $x$. By definition,

$$\begin{aligned}
\pi_i(x) \;&\equiv\; \int \begin{pmatrix} E[v_i(S', \rho')|\kappa' = i, x'] \Pr(\kappa' = i|x') + \\ E[v_i(S', \rho')|\kappa' = i, x'] \Pr(\kappa' \neq i|x') \end{pmatrix} dG(x'|x) \\
&=\; \int \begin{pmatrix} q_i(x') \left[\beta_i \pi_i(x') + \int \max\{c(x', \varepsilon') - \sum_j \beta_j \pi_j(x'), 0\} dF_{\epsilon|X=x'}(\varepsilon')\right] + \\ (1 - q_i(x'))\beta_i \pi_i(x') \end{pmatrix} dG(x'|x) \qquad (34)
\end{aligned}$$

with $\kappa'$ being the identity of the proposer in the next period, $q_i(x) \equiv \Pr(\kappa = i|x)$, and the second equality following from Lemma 7. Hence (34) implies $\overrightarrow{\pi}$ solves the structural fixed-point equation

$\overrightarrow{\pi} = T(\overrightarrow{\pi}; \overrightarrow{\beta}, c, L_{\rho|X}, F_{\epsilon|X})$, where the $i$-th component of $T$ is defined as

$$\begin{aligned} & T_i(\overrightarrow{f}; \overrightarrow{\beta}, c, L_{\rho|X}, F_{\epsilon|X}) \\ \equiv \quad & \int \beta_i f_i(x') + q_i(x') \int \max\{c(x', \varepsilon') - \textstyle\sum_j \beta_j f_j(x'), 0\} dF_{\epsilon|X=x'}(\varepsilon') dG(x'|x) \end{aligned}$$

This in turn implies under $MT$, $\pi_i(x)$ can be alternatively represented as the solution to the following "quasi-structural" fixed-point equation: $\pi_i = \tilde{T}(\pi_i; \beta_i, \overrightarrow{\psi}, F_{Y|X}, L_{\rho|X})$, where

$$\tilde{T}(f_i; \beta_i, \overrightarrow{\psi}, F_{Y|X}, L_{\rho|X}) = \int \beta_i f_i(x') + \tilde{\phi}_i(x'; \overrightarrow{\psi}, F_{Y|X}, L_{\rho|X}) dG(x'|x)$$

with $\tilde{\phi}_i(x; \overrightarrow{\psi}, F_{Y|X}, L_{\rho|X}) \equiv q_i(x) \int \max\{y - \sum_{j=1}^{K} \psi_j(x), 0\} dF_{Y|X=x}(y)$. It is easy to show that, with $\overrightarrow{\psi}, F_{Y|X}, L_{\rho|X}$ observed and fixed from data, the "quasi-structural" mapping $\tilde{T}$ is a contraction for any generic $\hat{\beta}_i \in (0,1)$. Let $\hat{\pi}_i$ denote the solution $\hat{\pi}_i = \tilde{T}(\hat{\pi}_i; \hat{\beta}_i, \overrightarrow{\psi}, F_{Y|X})$ for a generic $\hat{\beta}_i$. Arguments similar to Lemma 3 show that, with $\tilde{\phi}_i$ positive for all $i, x$ by definition, $\hat{\beta}_i \hat{\pi}_i(x; \hat{\beta}_i, \overrightarrow{\psi}, F_{Y|X}, L_{\rho|X})$ must be strictly increasing in $\hat{\beta}_i$ for all $x$. Therefore a true individual discount rate $\beta_i$ is identified as

$$\inf\{\hat{\beta}_i : \hat{\beta}_i \hat{\pi}_i(x; \hat{\beta}_i, \overrightarrow{\psi}, F_{Y|X}, L_{\rho|X}) \geq \psi_i(x)\}$$

This completes the proof for all $i$. $\square$

## 8.2 Part B: Details of the example in Section 5.3

(*Counterfactual outcomes when the true distribution of USV is uniform and known*) The closed form for the system of nonlinear equations in (18) is:

$$\begin{aligned} \tfrac{49567930}{6444299} p_1^2 + \tfrac{23508550}{6444299} p_2^2 + \tfrac{55809500}{6444299} p_3^2 + 10p_1 - 5 &= \tfrac{717442573}{165078240} \\ \tfrac{55814500}{6444299} p_1^2 + \tfrac{22413980}{6444299} p_2^2 + \tfrac{50657500}{6444299} p_3^2 + 10p_2 - 5 &= \tfrac{97368349}{132062592} \\ \tfrac{50090600}{6444299} p_1^2 + \tfrac{22567500}{6444299} p_2^2 + \tfrac{56227880}{6444299} p_3^2 + 10p_3 - 5 &= \tfrac{330851369}{264125184} \end{aligned} \tag{35}$$

*(Innocuous location and scale normalizations)* For example, suppose $a = 3, b = 2$. Then the nonlinear system in (19) is

$$\begin{aligned} \tfrac{29740758}{6444299} p_1^2 + \tfrac{14105130}{6444299} p_2^2 + \tfrac{33485700}{6444299} p_3^2 + 6p_1 - 1 &= \tfrac{1267703373}{275130400} \\ \tfrac{33488700}{6444299} p_1^2 + \tfrac{13448388}{6444299} p_2^2 + \tfrac{30394500}{6444299} p_3^2 + 6p_2 - 1 &= \tfrac{537576989}{220104320} \\ \tfrac{30054360}{6444299} p_1^2 + \tfrac{13540500}{6444299} p_2^2 + \tfrac{33736728}{6444299} p_3^2 + 6p_3 - 1 &= \tfrac{1211268649}{440208640} \end{aligned} \tag{36}$$

which is the same system as (36). Such an equivalence holds for all $a \neq 5$ and $b \neq 0$ in general.

*(Robust identification of ISRCO without knowing the USV distribution)* For any candidate counterfactual $p^1$ considered and the actual $p^0$ observed in the DGP, rewrite the linear system (12)-(16) as :

$$M_I V > 0 \quad (37)$$
$$M_E V = d \quad (38)$$

where $V \equiv [Q^0, Q^1, \Phi^0, \Phi^1]$ is the vector of unknown distributional parameters from $F_\epsilon$. Then substitute out a subvector of $V$ in (37) using the equalities in (38). This give a system of strict inequalities in the form

$$\tilde{M}_I \tilde{V} > b$$

We want to check if $(p^0, p^1)$ makes this linear system feasible with at least one solution $\tilde{V} = \tilde{v}$. We exploit the fact that this is equivalent to

$$\begin{aligned} -\tilde{M}_I \tilde{v} + b < & \ 0 \text{ for some } \tilde{v} \\ \Longleftrightarrow & \ \text{solution to " } \min_{(\tilde{v},t)} t \text{ s.t. } -\tilde{M}_I \tilde{v} + b \leq \mathbf{1}'t \text{ " is strictly negative} \\ \Longleftrightarrow & \ \text{solution to " } \min_{(\tilde{v},t)} t \text{ s.t. } -\tilde{M}_I \tilde{v} - \mathbf{1}'t \leq -b \text{ " is strictly negative} \end{aligned}$$

Standard linear programming algorithms can be used for checking the feasibility of the system. For the $p^0$ observed, collecting all $p^1$ that makes the system feasible gives the *ISRCO*.

# References

[1] Athey, S. and P. Haile (2002), "Identification of Standard Auction Models," Econometrica 70:2107-2140

[2] Athey, S. and P. Haile (2005), "Nonparametric Approaches to Auctions," Handbook of Econometrics, Vol.6, forthcoming

[3] Bajari, P., Hong, H., Krainer, J. and D. Nekipelov, "Estimating Static Models of Strategic Interactions," forthcoming in the Journal of Business and Economic Statistics.

[4] Benjamin, D. and M. Wright, "Recovery Before Redemption? A Theory of Delays in Sovereign Debt Renegotiations," Working Paper, UCLA, April 2008.

[5] Berry, S. and E. Tamer, "Identification in Models of Oligopoly Entry," Advances in Economics and Econometrics, 2006, Vol 2, Ninth World Congress, 46-85, Cambridge University Press.

[6] Bi, R., "'Beneficial' Delays in Debt Restructuring Negotiations," Working Paper 08/38, IMF, February 2008.

[7] Binmore, K., Osborne, M. and A. Rubinstein, “Noncooperative Models of Bargaining” in Aumann, R. and S. Hart (Eds.), Handbook of Game Theory with Economic Applications, Vol. I, ch. 7, pp. 179-225, Amsterdam: Elsevier, 1992.

[8] Chen, S., Dahl, G., and S. Khan, “Nonparametric Identification and Estimation of a Censored Location-Scale Regression Model,” Journal of the American Statistical Association, March 2005, 100, 212-221.

[9] Chiappori, P. and O. Donni, “Learning from a Piece of Pie: The Empirical Content of Nash Bargaining,” Working Paper 2128, IZA, May 2006.

[10] Diaz-Moreno, C. and J. Galdon-Sanchez, Collective Bargaining Under Complete Information,” Working Paper TE/2000/401, LSE STICERD, July 2000.

[11] Diermeier, D., Keane, M. and A. Merlo, “A Political Economy Model of Congressional Careers,” American Economic Review, March 2005, 95, 347-373.

[12] Diermeier, D., Eraslan, H. and A. Merlo, “A Structural Model of Government Formation,” Econometrica, January 2003, 71, 27-70.

[13] Eraslan, H., “Corporate Bankruptcy Reorganizations: Estimates from a Bargaining Model,” International Economic Review, May 2008, 49, 659-681.

[14] Guerre, E., I. Perrigne and Q. Vuong (2000), "Optimal Nonparametric Estimation of First-Price Auctions," Econometrica 68:525-574

[15] Haile, P. and E. Tamer (2003), "Inference with an Incomplete Model of English Auctions," Journal of Political Economy, vol.111, no.1

[16] Haile, P., H. Hong and M. Shum (2004) "Nonparametric Tests for Common Values in First-Price Sealed Bid Auctions," mimeo, Yale University

[17] Hendricks, K., J. Pinkse and R. Porter (2003), "Empirical Implications of Equilibrium Bidding in First-Price, Symmetric, Common Value Auctions," Review of Economic Studies 70:115-145

[18] Laffont, J.J. and Q. Vuong, “Structural Analysis of Auction Data,” American Economic Review, May 1996, 36, 414-420.

[19] Merlo, A., “Bargaining over Governments in a Stochastic Environment,” Journal of Political Economy, February 1997, 105, 101-131.

[20] Merlo, A., Ortalo-Magne, F. and J. Rust, “The Home Selling Problem: Theory and Evidence,” Working Paper, University of Maryland, March 2009.

[21] Merlo, A. and C. Wilson, “A Stochastic Model of Sequential Bargaining with Complete Information,” Econometrica, March 1995, 63, 371-399.

[22] Merlo, A. and C. Wilson, "Efficient Delays in a Stochastic Model of Bargaining," Economic Theory, January 1998, 11, 39-55.

[23] Muthoo, A., Bargaining Theory with Applications, Cambridge: Cambridge University Press, 1999.

[24] Osborne, M. and A. Rubinstein, Bargaining and Markets, San Diego, CA: Academic Press, 1990.

[25] Rubinstein, A., "Perfect Equilibrium in a Bargaining Model," Econometrica, January 1982, 50, 97-109.

[26] Sieg, H., "Estimating a Bargaining Model with Asymmetric Information: Evidence from Medical Malpractice Disputes," Journal of Political Economy, October 2000, 105, 1006-1021.

[27] Simcoe, T., "Standard Setting Committees," Working Paper, University of Toronto, December 2008.

[28] Stahl, I., Bargaining Theory, Stockholm: Economic Research Institute, Stockholm School of Economics, 1972.

[29] Tang, X., "Estimating Simultaneous Games with Incomplete Information under Median Restrictions," Working Paper, University of Pennsylvania, April 2009.

[30] Watanabe, Y., "Learning and Bargaining in Dispute Resolution: Theory and Evidence from Medical Malpractice Litigation," Working Paper, Northwestern University, August 2006.