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 10-008

"Identification of Stochastic Sequential Bargaining Models"
Second Version

by

Antonio Merlo and Xun Tang

http://ssrn.com/abstract=1571298

Identification of Stochastic Sequential Bargaining Models[1]

Antonio Merlo

University of Pennsylvania

Xun Tang

University of Pennsylvania

March 11, 2010

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 an agreement. In this paper, we present new results in nonparametric identification of such models under different scenarios of data availability. First, we give conditions for an observed distribution of players' decisions and agreed allocations of the surplus, or the "cake", to be rationalized by a sequential bargaining model. We show the common discount rate is identified, provided the surplus is monotonic in unobservable states (USV) given observed ones (OSV). Then the mapping from states to surplus, or the "cake function", is also recovered under appropriate normalizations. Second, when the cake is only observed under agreements, the discount rate and the impact of observable states on the cake can be identified, if the distribution of USV satisfies some exclusion restrictions and the cake is additively separable in OSV and USV. Third, if data only report when an agreement is reached but never report the size of the cake, we propose a simple algorithm that exploits shape restrictions on the cake function and the independence of USV to recover *all* rationalizable probabilities for agreements 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 Steven Berry, 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 past 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 led the way to more systematic attempts to take strategic bargaining models to data.

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). It also 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 variety of data sets.[2] However, little is known about whether the structural elements of these models or the bargaining outcomes in counterfactual environments 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 structural elements of these models in estimations, as in most cases such specifications 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 elements of the model.[3]

Empirical contexts of stochastic bargaining games may differ in what the econometricians observe in data. These differences 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 size of the surplus to be allocated, or "the cake", in each period regardless of whether an agreement is reached), "incomplete data with censored cakes" (where econometricians only observe the size of the cake in the period when an agreement is reached), and "incomplete data with unobservable cakes" (where econometricians only observe the timing of agreement, but never observe the size of the surplus). In all three scenarios, econometricians observe the evolution of a subset of the states that affect the total 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

[2]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.

[3]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, 2005), 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.

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 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).

For the case with complete data on surplus and players' decisions, we derive conditions for a joint distribution of observed states, surplus, agreements, and divisions of the cake to be rationalized by a class of stochastic sequential bargaining models where the transition of states and total surplus are independent of the bargaining protocol. We show how to recover the common discount factor from such rationalizable distributions, when the total surplus is monotone in USV. We also characterize the identified set for the mapping from states to surplus (i.e. the "cake function") and the bargaining protocol (i.e. the distribution of orders of moves among players), and show they can be recovered under appropriate normalizations. For the case with data on censored cakes, we show when the total surplus is additively separable in OSV and USV, then the impact of OSV on surplus is identified, provided the USV distribution satisfies some exclusion restrictions or has multiplicative heterogeneity.

In the data scenario with unobserved cakes, earlier results in Berry and Tamer (2006) on identifying optimal stopping problems also apply in the context of sequential bargaining under the assumption that the USV distribution is known to econometricians. Our contribution in this scenario is to relax the assumption of a known USV distribution, and show partial identification of counterfactual outcomes (i.e. probability for reaching an agreement conditional on observed states) is possible under nonparametric shape restrictions of the cake function and independence of USV. Our approach is motivated by the fact that the cake function is often known to satisfy certain shape restrictions derived 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 outcome space, with the aid of nonparametric restrictions such as independence of USV. To our knowledge, this is the first positive result in identifying counterfactuals in optimal stopping models without assuming knowledge of the USV distribution. We propose a simple algorithm to recover the identified set of rationalizable counterfactual outcomes (*ISRCO*), which are defined as conditional agreement probabilities consistent with players' dynamic rationality, the shape and stochastic restrictions of the model elements, as well as the actual outcomes observed in data. We use numerical examples to show the *ISRCO* recovered can be informative, and small relative to the whole outcome space.

We also address the identification of two extensions of the canonical model of stochastic bargaining where the players evaluate the surplus according to a concave utility function, or the discount factors are heterogeneous across players.[4] We show that if players across all bargaining games observed in data are known to follow strategies with the same stationary subgame perfect equilibrium (SSPE) payoffs, then heterogenous discount rates and utility functions can both be identified from complete data under fairly weak restrictions on players' risk attitudes.[5]

The rest 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 our main 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.

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

[5]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 (e.g. Bajari, Hong, Krainer and Nekipelov (2008) and Tang (2009)). Such assumptions allow econometricians to link model structures to observable distributions using theoretical characterizations of Bayesian Nash equilibria (BNE) or SSPE payoffs, while remaining agnostic about which BNE or SSPE payoff is realized in the data-generating process.

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 Ω_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. $R^t \equiv \{R_0, R_1, R_2, ., R_t\}$.) The set of feasible utility vectors to be allocated in period t with realized state s_t is given by $C(s_t) \equiv \{u \in \mathbb{R}^K : \sum_{i=1}^K u_i \leq c(s_t)\}$, where $c(s_t) : \Omega_S \rightarrow \mathbb{R}^1_+$ is the "cake function".[6] In each period t, the order of moves among players is a permutation of $\{1, 2, ., K\}$, and is denoted by a K-vector ρ_t, whose i-th coordinate $\rho_{t,(i)}$ is the identity of the player who makes the i-th move at time t. Let Ω_ρ denote the set of all possible permutations of the K-vector. Let $\kappa_t \equiv \rho_{t,(1)}$ denote the proposer in period t. Throughout the paper, we maintain the following restriction on the transition of states and the order of moves.

CI-1 (Conditional independence of histories) Conditional on S_t, (i) ρ_t is independent of past states and orders of moves $\{S^{t-1}, \rho^{t-1}\}$, and the conditional distribution over Ω_ρ, denoted by $\tilde{L}_t(\rho_t|S_t)$, is time-homogeneous; and (ii) S_{t+1} is independent of $\{S^{t-1}, \rho^t\}$, and transitions of states $H_t(S_{t+1}|S_t)$ are time-homogeneous.

The conditions in *CI-1* imply that the order of moves among players does not reveal any information to players about future states or surplus to be shared in the current period in addition to what they already see in S_t. Such a restriction is justified in empirical contexts where the order of moves does not affect the transition of states S_t or the evolution of total surplus. It accommodates the special case where the order of moves in period t is determined by the current state S_t alone. It can also allow ρ_t to be given by a function $\rho(S_t, \zeta_t)$, where ζ_t consists of some noises excluded from S_t and unobservable to players and econometricians. In this case, *CI-1* holds if, conditional on any s_t, $\zeta_t \perp \{S^{t-1}, \zeta^{t-1}\}$ and $S_{t+1} \perp \{S^{t-1}, \zeta^t\}$. Under *CI-1*, the transition between information variables is reduced to

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

where the time subscripts are dropped from $\tilde{L}_t(.|.)$ and $H_t(.|.)$ because they are both time-

[6] 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.

homogeneous under *CI-1.* Unless the timing is not clear from the context, we use (R, R') to denote random vectors in the current and the next period respectively.

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 ρ 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 ρ' 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 (τ, η_τ) consists of a stopping time τ and a random K-vector η_τ 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 $\tilde{s}_t \equiv (s_t, \rho_t)$, τ denotes the period in which a proposal is accepted by all players, and η_τ denotes the proposed allocation which is accepted in state s_τ when the order of moves is ρ_τ. For a game starting with state s and order of moves ρ, an outcome (τ, η_τ) 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 there 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 by all players. Using property (iii), we let $v^\mu(s, \rho) \equiv E[\beta^\tau \mu(\tilde{S}_\tau) | \tilde{S}_0 = (s, \rho)]$ denote the vector of individual von-Neumann-Morgenstern payoffs given initial state and order of moves (s, ρ). 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 *(Characterization of SSPE) Suppose CI-1 holds. Then (a) $v \in F^K$ is a unique SSPE payoff if and only if $A(v) = v$, with coordinates of A defined for all $(s, \rho) \in \Omega_{S,\rho}$ as:*

$$A_i(v)(s,\rho) \equiv \max\{c(s) - \beta E[\textstyle\sum_{j\neq i} v_j(S',\rho')|S=s], \beta E[v_i(S',\rho')|S=s]\} \text{ if } \rho_{(1)} = i \quad (2)$$

$$A_j(v)(s,\rho) \equiv \beta E[v_j(S',\rho')|S=s], \text{ if } \rho_{(1)} \neq j \quad (3)$$

(b) the SSPE total payoff w is independent of ρ given s, and solves

$$w(s) = \max\{c(s), \beta E[w(S')|S=s]\} \quad (4)$$

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

The proof of Lemma 1 uses results in Theorems 1-3 in Merlo and Wilson (1998). It exploits conditions in *CI-1* to show that the total payoff in SSPE and the occurrence of an agreement only depend on the current state S, but not the order of moves ρ. This property of SSPE, known as the "Separation Principle", is instrumental for some of our identification strategies below. In contrast, the individual SSPE payoffs $(v_i)_{i\in K}$ still depend on the order of moves. Namely, only $i = \rho_{(1)}$ can claim the additional "gain to the proposer", i.e. $c(S) - \beta E[\sum_{k=1}^K v_k(S',\rho')|S=s]$, while all others just get their individual continuation payoffs $\beta E[v_j(S',\rho')|S]$.

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, the identity of the proposer and observable states X in every period, but not ϵ. 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. As is common in inference of structural models, we posit 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 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 observed distributions of (X, ρ) on the vector (K, Z^K) only for the sake of notational simplicity.

3 Empirical Content of the Model with Complete Data

In this section we consider the empirical content of the canonical stochastic bargaining model when econometricians observe a complete history of *(i)* states X_t and sizes of the cake $Y_t = c(X_t, \epsilon_t)$ (but not ϵ_t); *(ii)* whether an 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 $\kappa_t \equiv \rho_{t,(1)}$) 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 τ is the termination period when an agreement is reached), but may not observe details of proposals in any other period.

Below we first derive necessary and sufficient conditions for an observed joint distribution of states and decisions to be rationalized in SSPE under the conditional independence of state transitions and the monotonicity of the cake function. Next we provide a constructive proof of identification of the common discount rate β. We also characterize the identified set of the cake function, the USV distribution, and the proposer-choosing mechanism, and show how these parameters are recovered under the appropriate normalizations. For 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 Ω_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}, ϵ_{t+1} is independent of (X_t, ϵ_t) for all t; and (ii) conditional on X_t, X_{t+1} is independent of ϵ_t for all t.

The condition *CI-2* requires dynamics between the current and the next period's states S and S' to be captured by persistence between 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}, ρ_{t+1}) are independent of ϵ_t since $F(s', \rho'|s) = \tilde{L}(\rho'|s')F(\varepsilon'|x')G(x'|x)$. 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 transition of observable states $G_{X'|X}$ and the distribution of initial states X_0 can be directly recovered from data. We say a joint distribution of the stopping time τ, the agreed allocations $\eta_\tau \equiv (\eta_{\tau,i})_{i=1}^K$ and the history $(X^\tau, Y^\tau, \rho^\tau)$ is *rationalized* by some θ if it is the distribution that arises in some SSPE of the bargaining game characterized by θ. Define a feature $\Gamma(.)$ as a mapping from a vector of parameters θ to some space of features. For example, $\Gamma(\theta)$ could be θ itself, or a subvector of θ (such as β or c), or some functional of θ such as the location (median) or the scale (variance) of ϵ given X.

Definition 1 *Let Θ denote a set of parameters satisfying certain restrictions. Two parameters θ, θ' are observationally equivalent (denoted $\theta \overset{o.e.}{\sim} \theta'$) under restrictions Θ if $\theta, \theta' \in \Theta$ and both rationalize the same joint distribution of $\{\tau, \eta_\tau, X^\tau, Y^\tau, \rho^\tau\}$. A feature of the true parameter θ_0 (denoted $\Gamma(\theta_0)$) is identified under Θ if $\Gamma(\theta_0) = \Gamma(\theta)$ for all $\theta \overset{o.e.}{\sim} \theta_0$ in Θ.*

Any feature of the true parameter $\Gamma(\theta_0)$ that can be expressed in terms of observable distributions of $\{\tau, \eta_\tau, X^\tau, Y^\tau, \rho^\tau\}$ is identified. Also note identification is defined under *CI-1,2* as well as all additional restrictions on θ captured by Θ.[7] For any θ in Θ, let $\pi_i(S, \rho; \theta) \equiv E[v_i(S', \rho'; \theta)|S, \rho]$ and $\pi_w(S; \theta) \equiv E[w(S'; \theta)|S]$ denote respectively the individual and total continuation payoffs induced by θ in SSPE, where v_i and w are respectively the individual

[7] The role of *CI-1,2* in this definition of identification is to reduce the parameter space from that of $\{\beta, c, \tilde{H}_t(S_{t+1}, \rho_{t+1}|S^t, \rho^t)\}$ to that of $\theta \equiv \{\beta, c, \tilde{L}_{\rho|S}, F_{\epsilon|X}\}$.

and total SSPE payoffs given by (2)-(3) and (4) in Lemma 1. Throughout the section, we maintain the following two restrictions on the parameter space Θ.

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

ND (Non-degeneracy) For all i and x, $\Pr\{c(S) - \beta\pi_w(S;\theta) \geq 0\ ,\ \rho_{(1)} \neq i \mid X = x\} \in (0, 1)$.

Under *MT*, there exists a one-to-one mapping between cake sizes and unobserved states given any x. *ND* requires there is enough variation in unobserved states for all states x, so that an agreement may occur or not occur with positive probabilities. It helps rule out uninteresting cases where there is no uncertainty about reaching an agreement once observable states are realized. It also states that for each player i there is always positive chance that an agreement is reached on someone else's proposal. For instance, *ND* is satisfied if $\Pr\{c(S) - \beta\pi_w(S;\theta) \geq 0|X = x\} \in (0, 1)$ and $\Pr\{\rho_{(1)} \neq i|s\} > 0$ for all i and s. Note it can be tested using observable joint distributions.

The starting point for discussing identification is that the data-generating process (DGP) is correctly specified under *CI-1,2* and MT, ND. That is, the distribution of states and actions observed can indeed be rationalized by some $\theta \in \Theta$ under *CI-1,2*, where Θ is the set of all θ that satisfy MT and ND. The following lemma gives conditions for an observed joint distribution of $\{\tau,\ \eta_\tau,\ X^\tau,\ Y^\tau,\ \rho^\tau\}$ to be rationalizable under *CI-1,2*. If such conditions are not satisfied, the model must be misspecified and discussions of identification based on the observed distribution would be vacuous, as no θ in Θ can rationalize what the econometrician observes.

Lemma 2 *(Conditions for rationalizability) A joint distribution of $\{\tau, \eta_\tau, (X_t, Y_t, \rho_t)_{t=0}^{\tau}\}$ is rationalized by some $\theta \in \Theta$ under CI-1,2 if and only if: (i) if $\tau \geq 1$, then for all t and (x^t, y^t, d^t, ρ^t),*

$$F_{Y_{t+1},D_{t+1},\rho_{t+1},X_{t+1}|\sum_{s=0}^{t} D_s=0,y^t,\rho^t,x^t} = F_{Y_{t+1},D_{t+1},\rho_{t+1}|X_{t+1}} G_{X_{t+1}|x_t} \tag{5}$$

where $F_{Y,D,\rho|X}$ and $G_{X'|X}$ are time-homogeneous; (ii) for all x, $F_{Y|X=x}$ is strictly increasing and $p(x) \in (0, 1)$ where $p(x) \equiv \Pr\{D = 1|X = x\}$; (iii) $Y_\tau \geq \lambda^*(X_\tau)$, $Y_t < \lambda^*(X_t)$ for all $t < \tau$, where $\lambda^*(x) \equiv F_{Y|X=x}^{-1}(1 - p(x))$; and (iv) there exists $\alpha \in (0, 1)$ s.t. for all x,

$$\left(\int \max\{y', \lambda^*(x')\} dF_{Y',X'|X=x}\right)^{-1} \lambda^*(x) = \alpha \tag{6}$$

; and (v) $\Pr\{D = 1, \rho_{(1)} \neq i \mid X = x\} \in (0,1)$ *for all* i, x*, and there exist* K *functions* $(\lambda_i^*)_{i \in K}$ *s.t.* $\eta_{\tau,i} = \lambda_i^*(X_\tau)$ *for* $i \neq \rho_{\tau,(1)}$*,* $\eta_{\tau,i} = Y_\tau - \sum_{j \neq i} \lambda_j^*(X_\tau)$ *for* $i = \rho_{\tau,(1)}$*, and for all* i, x*,*

$$\lambda_i^*(x) = \alpha \int \lambda_i^*(x') + \int \max\{y' - \lambda^*(x'), 0\} 1\left(\rho'_{(1)} = i\right) dF_{Y',\rho'|X'} dG_{X'|X} \tag{7}$$

Conditions *(i)-(v)* all have intuitive interpretations. In SSPE, whether an agreement occurs only depends on current states. The conditional independence in (5) and time-homogeneity of $F_{Y,D,\rho|X}$ in *(i)* are due to $H_t(S_{t+1}, \rho_{t+1}|S^t, \rho^t) = \tilde{L}(\rho_{t+1}|s_{t+1})\ F_{\epsilon|X}(\varepsilon_{t+1}|x_{t+1})$ $G_{X'|X}(x_{t+1}|x_t)$ under *CI-1,2*. The strict monotonicity of the cake size distribution in *(ii)* results from the strict monotonicity of c and $F_{\epsilon|X}$ in *MT*. Under *CI-1,2*, subsequent states (S', ρ') are independent of (ϵ, ρ) given X, and therefore the total as well as individual continuation payoffs must be functions of X alone. By Lemma 1, this means an agreement is reached whenever cake sizes exceed some fixed threshold conditional on X. In addition, ND restricts the total continuation payoff to lie in the interior of the support of cake sizes, and condition *(iii)* simply relates discounted total continuation payoffs to the conditional quantile of the cake size under state x. Condition *(iv)* also builds on the same intuition, and relates the common discount factor to the distribution of observables. In SSPE, a player only receives his individual continuation payoffs if an agreement is reached on someone else's proposal, an event that occurs with positive probability under ND. Condition *(v)* then simply relates each individual's continuation payoffs to the physical shares that he is observed to receive as a non-proposer under agreements in any state x.

The rest of the section discusses identification of $\theta \equiv \{\beta, \sigma\}$ where $\sigma \equiv \{c, F_{\epsilon|X}, \tilde{L}_{\rho|S}\}$ when the joint distribution $\{\tau, \eta_\tau, X^\tau, Y^\tau, \rho^\tau\}$ observed from data is rationalizable, i.e. satisfying conditions in Lemma 2. Let $F^*_{\{D^\tau, X^\tau, Y^\tau, \rho^\tau\}}$ denote such an observed, rationalizable joint distribution. Let $F^*_{R_2|R_1}$ denote the observed time-homogeneous distribution of R_2 given R_1, where R_1, R_2 are subvectors of (D, Y, ρ, X). Let $\theta_0 \equiv \{\beta_0, \sigma_0\}$ where $\sigma_0 \equiv \{c_0, F^0_{\epsilon|X}, \tilde{L}^0_{\rho|S}\}$ denote true parameters in the actual data-generating process (DGP) that underlies the observed distribution $F^*_{\{D^\tau, X^\tau, Y^\tau, \rho^\tau\}}$.

Proposition 1 *(Identification of* β_0*) Under CI-1,2, MT and ND, the common discount rate is identified as*

$$\beta_0 = \left(\int \int \max\{y', \lambda^*(x')\} dF^*_{Y'|X'} dG_{X'|X=x} \right)^{-1} \lambda^*(x) \tag{8}$$

where λ^* *is defined in Lemma 2.*

The intuition of this identification strategy builds on two simple observations. First, under *CI-1,2*, the discounted total continuation payoff $\beta_0 \pi_w(s;\theta_0)$ must be a function of observable states x alone. Under ND, $\beta_0 \pi_w(x;\theta_0)$ must be in the interior of the support of surplus, and MT implies it can be directly recovered from the observed distribution of surplus as an appropriate conditional quantile, i.e. $\lambda^*(x)$. Second, changing variables between ϵ and Y under MT helps relate the discount total continuation payoff to observed distributions through a "quasi-structural" fixed-point equation:

$$\beta_0 \pi_w(x;\theta_0) = \beta_0 \int \int \max\{y', \beta_0 \pi_w(x';\theta_0)\} dF^*_{Y'|X'} dG_{X'|X=x} \tag{9}$$

where the prefix "quasi-" highlights that σ_0 enters through the observed distribution of cake sizes $F^*_{Y|X}$ it induces. Substituting $\lambda^*(x)$ in place of $\beta_0 \pi_w(x;\theta_0)$ in the quasi-structural form gives (8).

With β_0 now identified, we give a sufficient condition for a combination of remaining parameters $\sigma \equiv \{c, F_{\epsilon|X}, \tilde{L}_{\rho|S}\}$ to be observationally equivalent to the truth σ_0. Knowing such a condition reveals the source of under-identification. Let $F^{(\sigma)}_{D,Y,\rho|X}$ denote the conditional distribution of (D, Y, ρ) induced by σ.

Lemma 3 *(Identified set of σ) Let the true discount factor β_0 be identified. A σ satisfying MT, ND can rationalize $F^*_{\{D^\tau, X^\tau, Y^\tau, \rho^\tau\}}$ under CI-1,2 if and only if $F^{(\sigma)}_{Y,\rho|X=x} = F^*_{Y,\rho|X=x}$ for all $x \in \Omega_X$.*

Necessity follows from the definition of rationalization. The intuition behind the proof of sufficiency is as follows. Since a rationalizable $F^*_{\{D^\tau, X^\tau, Y^\tau, \rho^\tau\}}$ must necessarily satisfy (5) with a time-homogeneous $F^*_{Y,\rho,D|X=x}$, it suffices to show *(i)* $F^{(\sigma)}_{Y,\rho,D|X=x} = F^*_{Y,\rho,D|X=x}$ for all x, and *(ii)* the individual continuation payoffs induced by σ coincides with observed shares received by a non-proposer in any state x. Recall for any $\theta \equiv \{\beta, \sigma\}$, the individual continuation payoffs in SSPE is given by:

$$\pi_i(x;\theta) = \int \beta \pi_i(x';\theta) + \int \max\{c(x', \varepsilon') - \beta \pi_w(x';\theta), 0\} 1\left(\rho'_{(1)} = i\right) dF_{\epsilon',\rho'|X'} dG_{X'|X=x} \tag{10}$$

Since $\pi_i(s, \rho;\theta)$ and $\pi_w(s;\theta)$ must depend on x alone under *CI-1,2*. With β_0 identified and changing variables between ϵ and Y, we can formulate (10) into the "quasi-structural" form:

$$\begin{aligned} & \beta_0 \pi_i(x;\beta_0,\sigma) \\ = \; & \beta_0 \int \beta_0 \pi_i(x') + \int \max\{y' - \beta_0 \pi_w(x'), 0\} 1\left(\rho'_{(1)} = i\right) dF^{(\sigma)}_{Y',\rho'|X'} dG_{X'|X=x} \end{aligned} \tag{11}$$

By supposition of Lemma 3, $F^{(\sigma)}_{Y,\rho|X} = F^{*}_{Y,\rho|X}$. In such a case, (9) implies $\beta_0\pi_w(x;\beta_0,\sigma) = \lambda^*(x)$ and (11) implies $\beta_0\pi_i(x;\beta_0,\sigma) = \lambda_i^*(x)$. Hence conditions *(i)* and *(ii)* above must hold, and the sufficient condition in Lemma 3 is established.

Since Lemma 3 has shown that $F^0_{\epsilon|X=x}$, c_0 cannot be jointly identified, one might think setting $F^0_{\epsilon|X=x}$ to some known distribution (say uniform$(0,1)$) in estimation is a necessary normalization in structural estimations. Unfortunately, in general such an arbitrary "normalization" can lead to errors in predicting the counterfactual distribution of (X,Y) when the mapping from states to the total surplus is changed. The only special case where such choices do not preclude correct counterfactual analyses is when unobserved states are known to be independent of observed ones. (See Appendix C for more details.) The following proposition shows how to identify the other structural elements in this case by normalizing F^0_ϵ to some known distribution.

Proposition 2 *Suppose CI-1,2, MT and ND hold, and ϵ is independent of X. With F^0_ϵ normalized to a known distribution, $\{c_0, \tilde{L}^0_{\rho|S}\}$ are identified as*

$$c_0(x,\varepsilon) = F^{*-1}_{Y|X=x}(F^0_\epsilon(\varepsilon)) \quad ; \quad \tilde{L}^0_{\rho|S=s} = F^*_{\rho|X=x,Y=c_0(x,\varepsilon)} \tag{12}$$

The proof follows from $F^*_{Y|X=x}(c_0(x,\varepsilon)) = F^0_\epsilon(\varepsilon)$ for all (x,ε) and $\tilde{L}^0_{\rho|X=x,\epsilon=c_0^{-1}(x,y)} = F^*_{\rho|X=x,Y=y}$ for all (x,y), where both equalities are due to strict monotonicity of Y in ϵ given all x under MT. One way to normalize F^0_ϵ is to let it be a uniform distribution on $[0,1]$ for all x. Then $c_0(x,\alpha)$ is identified as the α-th conditional quantile of Y given $X=x$, and $\tilde{L}^0_{\rho|S=(x,\alpha)} = F^*_{\rho|X=x,Y=c^0(x,\alpha)}$. An alternative normalization is to let $c_0(\bar{x},\varepsilon) = \varepsilon$ for some $\bar{x}$. Then $F^0_\epsilon(t)$ is identified as $F^*_{Y|X=\bar{x}}(t)$, and $c_0(x,\varepsilon)$ as $F^{*-1}_{Y|X=x}(F^*_{Y|X=\bar{x}}(\varepsilon))$. This is the normalization used in Matzkin (2003).[8]

4 Incomplete Data with Censored Cakes

In this section, we discuss identification of the canonical stochastic bargaining model when cake sizes and proposals are only observed in the event of an agreement. Thus the distribution of cake sizes observed is censored at the discounted total continuation payoffs in SSPE

[8] Matzkin (2003) also provided a slight generalization of the identification arguments by showing $c_0, F^0_{\epsilon|X}$ can be identified if $X \equiv (X_0, X_1)$ and ϵ is independent of X_1 conditional on X_0. In such cases, $c_0(x_0,x_1,\varepsilon)$ needs to be normalized for each x_0 as $c_0(x_0,\bar{x}_1,\varepsilon) = \varepsilon$ at some $\bar{x}_1$.

(i.e. $\beta_0 \pi_w(s;\theta_0) = \beta_0 E[w(S';\theta_0)|s]$). In such cases, the common discount rate can still be recovered from the distribution of observables under *CI-1,2*. The main theme in this section is that additive separability of the cake function in observed states X and the unobserved state ϵ, along with stochastic restrictions on $F^0_{\epsilon|X}$ such as multiplicative heterogeneity or strong exclusion restriction, are sufficient for identifying the cake function despite the censoring of observed cake sizes.

For rest of the paper, we continue to use subscripts and superscripts 0 to denote the true parameters in the data-generating process (e.g. $\beta_0, c_0, F^0_{\epsilon|X}$ and $\pi_w(.;\theta_0)$). We drop these subscripts or superscripts while referring to generic elements in the parameter space Θ. We continue to use F^* to denote the distribution observed from data.

4.1 Identification under multiplicative heterogeneity

We first consider models where ϵ is known to be in a "location-scale" family. This subsumes the case of normally distributed unobserved states (USV) with zero means and variances depending on observable states (OSV). Let $B(\Omega_X)$ denote the set of continuous, bounded functions defined over support of X, i.e. Ω_X. In this subsection, we maintain the following restrictions on the parameter space Θ.

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

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

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

RG (Regularity) For all $b \in B(\Omega_X)$, $\bar{b}(x) \equiv \int \max\{\tilde{c}(x) + \varepsilon, b(x)\} dF_{\epsilon|X=x} \in B(\Omega_X)$ *and* $\tilde{b}(x) \equiv \int b(x') dG_{X'|X=x} \in B(\Omega_X)$.

AS and *MH-(i)* together imply *MT* in the previous section. Under *MH*, $\tilde{c}(X)$ is the median of cake sizes in state X. Besides, *MH* implies ϵ_t is independent from history of states conditional on X_t, and hence the discounted total continuation payoffs $\pi_w(s;\theta)$ must be a function of x alone. Hence under *AS*, the *realized* "gains to the proposer", i.e. $Y^* \equiv \max\{Y - \beta\pi_w(X), 0\}$ can be represented as $\max\{\gamma(X) + \epsilon, 0\}$ with $\gamma(x) \equiv \tilde{c}(x) - \beta\pi_w(x)$ being the conditional median of the *potential* gains/losses to the proposer. Then the identification arguments in Chen, Dahl and Khan (2005) for nonparametric censored regressions can be applied to identify the true conditional median function. The zero median condition in MH is a location normalization.[9] The support condition in SG and the location-scale form in MH ensure that for any x there exists some α (possibly dependent upon x) close enough to 1 such that the conditional α-th quantile of the realized gains to the proposer must be linear in γ and σ (i.e. $q_\alpha(Y^*|x) = \gamma(x) + \sigma(x)c_\alpha$ where c_α is the α-th quantile of $\tilde{\epsilon}$). As explained below, this linearity is crucial for identifying the true conditional median of the potential gains/losses γ_0.Regularity conditions in RG guarantee the total continuation payoffs and the conditional median of potential gains to the proposer must both be bounded and continuous in the parameter spaces. Under AS and MH-(i), $\Pr\{D = 1|X = x\} = \Pr\{\gamma(X) + \epsilon \geq 0|X = x\} \in (0, 1)$ for all x. This guarantees that the true total continuation payoff $\beta_0\pi_w(.;\theta_0)$ can be recovered from $F^*_{Y|D=1,X}$ as infimum of the support of censored cake sizes, which is identical to $F^{*-1}_{Y|X}(1 - p(x))$ where $p(x) \equiv F^*_{D=1|X=x}$ is the conditional agreement probability $\Pr\{D = 1|X = x\}$ observed in data. Consequently, $\tilde{c}_0$ is recovered as the sum of γ_0 and $\beta_0\pi_w(.;\theta_0)$.

Proposition 3 *(i) Under CI-1, AS, MH, SG and RG, β_0 is identified, and both $\tilde{c}_0(X)$ and $\sigma_0(X)$ are identified over Ω_X from $F^*_{Y|D=1,X}$ and $F^*_{D=1|X}$; (ii) Under MH, the condition SG holds if and only if* $\Pr\{p(X) > \frac{1}{2}\} > 0$.

The identification of β_0 uses arguments similar to those in Proposition 1 in the case with complete data. The only difference is that, with cake sizes censored in data, the true total continuation payoff $\beta_0\pi_w(.;\theta_0)$ is recovered from $F^*_{Y|D=1,X}$ as the infimum of the censored support, instead of a quantile of the uncensored distribution $F^{*-1}_{Y|X=x}(p(x))$. SG guarantees the true γ_0 is identified directly as $q_{1/2}(Y^*|x)$ over Ω^+_X, which happens with a positive probability. RG ensures for any $x \in \Omega_X \backslash \Omega^+_X$, there exists a pair (α_1, α_2) strictly greater than $1/2$ and close enough to 1 (possibly dependent upon x) such that conditional quantiles $q_{\alpha_l}(Y^*|x)$ must be strictly positive for $l = 1, 2$. Under SG, we can recover the true values for such high quantiles of $\tilde{\epsilon}$ (denoted by $c^0_{\alpha_l}$) through a linear system that relates the conditional quantiles of cake sizes to $\gamma_0(x)$, $\sigma_0(x)$ and $c^0_{\alpha_l}$ for $x \in \Omega^+_X$. Knowledge of these $c^0_{\alpha_l}$ is then used to identify $\gamma_0(x)$ and $\sigma_0(x)$ for $x \notin \Omega^+_X$ by using the equation $q_{\alpha_l}(Y^*|x) = \gamma_0(x) + \sigma_0(x)c^0_{\alpha_l}$ for those α_l. The main identifying restriction SG is testable, since the event

[9] That the support of ϵ is unbounded is stronger than necessary. Identification only requires that conditional on all $x \in \Omega_X$, support of $\tilde{\epsilon}$ is large enough to induce positive gains to the proposer (and therefore a unanimous agreement) with positive probability.

"$\gamma_0(x) > 0$" is equivalent to "$\Pr\{\epsilon > -\gamma_0(x)|x\} > \frac{1}{2}$" when ϵ has strictly positive densities around 0 and $F^0_{\epsilon|X=x}(0) = 1/2$ for all $x \in \Omega_X$. Therefore, SG is equivalent to $\Pr\{p(X) > \frac{1}{2}\}$ > 0 and can be tested using observables.

4.2 Identification under exclusion restriction

Under MH, observed states only impact the scale of the distribution of unobserved disturbances, but not its shape. This restriction is easy to motivate only if the noise affecting total surplus is known to belong to some scale-location family. It rules out endogenous disturbances with unrestricted dependence on observed states. Besides, it implies stringent conditions on quantiles of the surplus given X. Namely, for *any* pair of observable states x, x' and *any* four percentiles $\{\alpha_j\}_{j=1}^4$ with $q_{\alpha_j}(Y^*|x), q_{\alpha_j}(Y^*|x') > 0$, the ratio between inter-quantile range of the total surplus must be independent of realized states. That is, for all x, x',

$$\chi(x) \equiv \frac{q_{\alpha_1}(Y^*|x) - q_{\alpha_2}(Y^*|x)}{q_{\alpha_3}(Y^*|x) - q_{\alpha_4}(Y^*|x)} = \frac{c_{\alpha_1} - c_{\alpha_2}}{c_{\alpha_3} - c_{\alpha_4}} = \chi(x')$$

for all $\{\alpha_j\}_{j=1}^4$. Below we show $\tilde{c}$ can be identified where the USV is correlated with observed states in ways that can accommodate general dependence of the shape of the USV distribution on X. In this subsection, we maintain the following additional restrictions on the parameter space Θ.

ER (Exclusion restriction) $X = (X_a, X_b)$, *and* ϵ *is independent of* X_b *given any* x_a *in each period.*

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

RS (Rich support) ϵ *has positive densities w.r.t. the Lebesgue measure over* $\mathbb{R}^1$*, and* $F_{\epsilon|X=x}(0) = 1/2$ *for all* x.

Our identification arguments extend immediately to cases where ER and SG' can hold after further conditioning on some available instruments X_c, which may or may not enter the function of median cake sizes $\tilde{c}$.

Proposition 4 *Under CI-1,2, AS, RG, ER, SG' and RS, the discount factor* β_0 *is identified and the cake function* $\tilde{c}_0$ *is identified on* Ω_X.

Under ER and for α large enough, the conditional α-th quantile of the potential gains/losses to the proposer is additively separable in γ_0 and the quantiles of the USV given observed states. That is, $q_\alpha(Y^*|x) = \gamma_0(x) + c^0_\alpha(x)$ for all $x \in \Omega_X$ and α close enough to 1, where $c^0_\alpha(x)$ is the true infinite dimensional nuisance parameter in the data-generating process. RG and RS guarantee in the parameter space that γ is bounded on Ω_X while the support of ϵ given any x is unbounded. Hence $q_\alpha(Y^*|x) > 0$ must hold even for $x \notin \Omega^+_X$ for some α greater than 1/2 and close enough to 1. ER and SG' allow us to fix any x_a and exploit variations in x_b alone to reach some state $\tilde{x} \equiv (x_a, \tilde{x}_b) \in \Omega^+_X$. Note ER implies that for all c_α in the parameter space, $c_\alpha(\tilde{x}) = c_\alpha(x_a)$ for all α (including those α greater than 1/2). Hence the true parameter $c^0_\alpha(\tilde{x})$ for any $\alpha > 1/2$ can be recovered as $q_\alpha(Y^*|\tilde{x}) - \gamma_0(\tilde{x})$, with both components observed directly for $\tilde{x} \in \Omega^+_X$. Now that the true nuisance parameter $c^0_\alpha(x_a)$ is identified for any $\alpha \geq 1/2$, $\gamma_0(x)$ can be identified for $x \notin \Omega^+_X$ as $q_{\bar{\alpha}}(Y^*|x) - c^0_{\bar{\alpha}}(x_a)$ for some $\bar{\alpha}$ greater than 1/2 and close enough to 1. Finally, the true median surplus function $\tilde{c}_0$ is recovered on Ω_X as the sum of γ_0 and $\beta_0\pi_w(x;\theta_0)$. Note the distribution of agreed proposals, though observed, are not involved in recovering $\tilde{c}_0$ from observed distributions, since assumptions AS, MH or ER have already introduced enough restrictions on $\{c, F_{\epsilon|X}\}$ in the parameter space to attain identification of the true parameter $\tilde{c}_0$.

5 Incomplete Data with Unobserved Cakes

In some other contexts, econometricians observe states X in all periods and when an agreement is reached, but observe neither the cake sizes, the order of moves, nor the agreed proposals in any period. Such a scenario arises when all parties in a game choose to keep details of negotiations and agreements confidential, and econometricians only get to observe the history of X and when the bargaining game ends. Econometricians seek to learn enough about the cake function and the USV distribution to predict probabilities for reaching an agreement in counterfactual contexts (such as when the transition between states are perturbed). Suppose conditions *CI-1,2* hold. Let Θ denote the set of generic restrictions on unknown parameters $\theta \equiv (\beta, c, F_{\epsilon|X})$. As in the previous section, we continue to use subscripts and superscripts 0 to denote true parameters (e.g. $\theta_0 = (\beta_0, c_0, F^0_{\epsilon|X})$), and drop these subscripts and superscripts while referring to generic elements in the parameter space. Note that now $\tilde{L}_{\rho|S}$ is dropped from θ, because the order of moves ρ is never observed and, by the "separation principle" in SSPE, does not affect the chance for reaching an agreement. Let $F_{D|X}(\theta)$ denote the conditional probability for agreements induced by a set of parameters θ. Let $F^*_{D|X} = F_{D|X}(\theta_0)$ denote the actual conditional agreement probability observed.

Definition 2 *(Identification with Unobserved Cakes)* $\theta \overset{o.e.}{\sim} \theta'$ *under* Θ *if* $\theta, \theta' \in \Theta$, $F_{D|X=x}(\theta) =$

$F_{D|X=x}(\theta')$ for all $x \in \Omega_X$. A feature of the truth θ_0 (denoted $\Gamma(\theta_0)$) is identified under Θ if $\Gamma(\theta_0) = \Gamma(\theta)$ for all $\theta \overset{o.e.}{\sim} \theta_0$ under Θ.

As in the case with complete data, identification is defined under conditional independence in *CI-1,2*. However, observational equivalence when the cake is unobserved by econometricians only requires parameters to induce the same static probabilities of agreements conditional on X only. This is because now neither the orders of moves nor the agreed proposals are observable. Our starting point for discussing identification is that the model is correctly specified under *CI-1,2* for some parameters $(\beta, c, F_{\epsilon|X})$. That is, the distribution observed necessarily satisfies restrictions implied under *CI-1,2* (i.e. $F^*_{D_{t+1},X_{t+1}|D^t,X^t} = F^*_{D_{t+1}|X_{t+1}} G^0_{X_{t+1}|X_t}$, with $F^*_{D|X}, G^0_{X'|X}$ observed and time-homogenous). Otherwise the set of θ's that are observationally equivalent to the true θ_0 would be vacuously empty. Throughout this section, we maintain that the true discount factor β_0 is known to econometricians. [10]

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 additive separability of the cake functions in the parameter space. That is, $c(s) = \tilde{c}(x) - \varepsilon$ for some unknown function $\tilde{c}$. At least some scale and location normalizations of $F_{\epsilon|X}$ are required to identify $\tilde{c}$. Such normalizations are innocuous in the sense that they do not affect the prediction of probabilities for reaching an agreement in counterfactual environments where the transition between states or the cake function is changed.

Proposition 5 *Suppose $c(s) = \tilde{c}(x) - \varepsilon$ for some unknown function $\tilde{c}$ and $\Pr(D = 1|X = x) \in (0, 1)$ for all $x \in \Omega_X$. Suppose CI-1,2 hold and the true USV distribution $F^0_{\epsilon|X=x}$ is known and strictly increasing for any $x \in \Omega_X$. (i) If the true discount factor β_0 is known, then $\tilde{c}_0$ is identified over Ω_X. (ii) If $\tilde{c}_0(\bar{x})$ is known for some $\bar{x} \in \Omega_X$, then β_0 is identified and $\tilde{c}_0$ is identified over Ω_X.*

The proof builds on results in identification of optimal stopping models in Berry and Tamer (2006). The assumption that the true distribution of unobservable states $F^0_{\epsilon|X}$ is

[10] Within the class of canonical models where players' utilities are transferrable, this assumption is often justified, as the discount rate can usually be recovered exogenously. For example, in some applications, the cake size is measured in monetary terms and the discount rate can be estimated as the interest rate that is relevant throughout the bargaining process.

known to researchers is less restrictive than it seems. Consider cases where distributions of unobserved states are known to be independent of X and belong to the normal family. Then restricting $F^0_{\epsilon|X}$ to be $N(0,1)$ in estimation is equivalent to an innocuous location and scale normalization. This is also true with other parametric families characterized only by location and scale parameters. The proof exploits the "separation principle" in SSPE to show that occurrence of agreements in the bargaining model is analogous to a collective decision to stop in an optimal stopping problem. Hence earlier results by Berry and Tamer (2006) apply to show identification.

5.2 Rationalizable counterfactual outcomes when USV distribution is unknown

When the unobservable state (USV) distribution is not known to belong to certain location-scale parametric family, imposing a specific form on the USV distribution that deviates from the truth can imply incorrect results in counterfactual outcomes. (See the example in Section 5.3 below.) On the other hand, economic theories often suggest the structural elements of the model have to satisfy certain shape or stochastic restrictions, such as monotonicity or concavity of the cake function or independence of ϵ from X. This raises the question: how can econometricians exploit such exogenously given restrictions to infer counterfactual outcomes without invoking parametric assumptions on the structure? We propose a simple, novel algorithm that helps recover the identified set of all rationalizable probabilities for reaching an agreement in counterfactual bargaining contexts, where transitions between states or the cake function are perturbed. We maintain the following assumption throughout this section.

SI (Statistical independence) ϵ is statistically independent of X.

Let G^0 denote the transition between states observed from data. We begin by noting a generic pair $(\tilde{c}, F_\epsilon)$ in the parameter space is observationally equivalent to the true parameters if and only if the following equation is satisfied:

$$q(p(x); F_\epsilon) \equiv F_\epsilon^{-1}(p(x)) = \tilde{c}(x) - \beta_0 \pi_w(x)$$

where $p(x) \equiv \Pr\{D = 1|X = x\} = F^*_{D=1|X}$ is the probability for reaching an agreement given x observed from data, and $\pi_w(x;\, \beta_0,\, G^0,\, p,\, F_\epsilon)$ is the total continuation payoff under SSPE. By definition,

$$\pi_w(x) = \int \beta_0 \pi_w(x') + \phi(p(x'); F_\epsilon) dG^0(x'|x) \tag{13}$$

where for any $\alpha \in (0,1)$,

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

and will be referred to as the "*conditional surplus function*" (CSF) hereafter. Note the equality in (13) follows from $p(x) = F_\epsilon(\tilde{c}(x) - \beta_0 \pi_w(x))$ for all 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 observed states Ω_X is finite with M elements $\{x_1, x_2, ., x_M\}$.

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

$$Q = \tilde{C} - \beta_0 \Pi \tag{14}$$

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

$$\Pi = G^0(\beta_0 \Pi + \Phi) \tag{15}$$

where Φ is a M-vector with the m-th coordinate defined as $\Phi_m \equiv \phi(p(x_m); F_\epsilon)$, and G^0 is the observed M-by-M transition matrix with the (m,n)-th entry defined as $G^0_{mn} \equiv \Pr(X' = x_n | X = x_m)$. Note the probabilities of reaching an agreement $p \equiv (p(x_1), ., p(x_M))$ enters the system defining observational equivalence through Φ in both Q and Π. We normalize the true median and the true CSF at $\alpha = 1/2$ respectively as $q(1/2; F_\epsilon^0) = 0$ and $\phi(1/2; F_\epsilon^0) = \bar{\phi}$ for some strictly positive constant $\bar{\phi}$. Under this normalization, $\tilde{c}_0(x)$ is the median cake size given state x.

In some empirical contexts, econometricians can reduce the parameter space for $\tilde{C}$ with the help of shape restrictions that are implied by economic theory or common sense. Often such shape restrictions can be represented as linear inequalities on $\tilde{C}$. For example, if the state x has three possible values $x_1 < x_2 < x_3$ and the conditional median cake function $\tilde{c}(x)$ is known to be strictly increasing, then the parameter space for $\tilde{C}$ is reduced to a subset satisfying $A\tilde{C} > 0$ with $A \equiv [-1, 1, 0; 0, -1, 1]$. Ranking of a subset of the states by $\tilde{c}(.)$ can also be represented as linear restrictions on $\tilde{C}$. For instance, knowing that the median cake size x_m is the smallest gives $M - 1$ strict inequalities. Similar matrices of coefficients can be constructed if $\tilde{c}$ is known to be increasing or additively separable or supermodular in some of the coordinates when x is multivariate.

Given a set of restrictions on the cake functions and the USV distributions, a vector of agreement probabilities p observed is *rationalized* under such restrictions if there exist some $\tilde{C}, F_{\epsilon|X}$ that satisfy these restrictions and can induce p in SSPE (i.e. p satisfies (14), (15) for such $\tilde{C}, F_{\epsilon|X}$). The following lemma gives conditions for a vector p to be rationalized under certain linear shape restrictions $A\tilde{C} > 0$ and the independence of ϵ from X. Let $V_{(m)}$ denote the m-th smallest element in a generic vector V. Let $\bar{\phi}$ be any strictly positive constant.

Lemma 4 *Suppose $c(s) = \tilde{c}(x) - \varepsilon$ for some unknown $\tilde{c}$ that satisfies $A\tilde{C} > 0$ with $q(1/2; F_\epsilon) = 0$. Suppose CI-1,2, DS and SI hold and β_0 is known. A vector of agreement probabilities p observed in the DGP is rationalized if and only if the following linear system holds for some vectors Q, Φ:*

$$A\left[Q + \beta_0(I - \beta_0 G^0)^{-1} G^0 \Phi\right] > 0 \tag{16}$$

$$Q_m \leq Q_n \Leftrightarrow p_m \leq p_n, \ \forall m, n \in \{1, ., M\} \tag{17}$$

$$p_{(m)}(Q_{(m+1)} - Q_{(m)}) \leq \Phi_{(m+1)} - \Phi_{(m)} \leq p_{(m+1)}(Q_{(m+1)} - Q_{(m)}), \ \forall m \in \{1,.,M\text{-}1\} \tag{18}$$

$$Q_m \leq 0 \Leftrightarrow p_m \leq 1/2 \ \textit{and} \ \tfrac{1}{2} Q_m \leq \Phi_m - \bar{\phi} \leq p_m Q_m, \ \forall m \in \{1, ., M\} \tag{19}$$

$$\Phi_m > 0 \ \textit{for} \ m \in \{1, ., M\} \tag{20}$$

Two remarks are in order. First, feasibility of the linear system is *not only necessary but also sufficient* for rationalizability of p. Sufficiency follows from the fact that when the linear system holds, a pair $(\tilde{C}, F_\epsilon)$ can be constructed to rationalize p under *CI-1,2*, *DS*, *SI* and the shape restrictions. Namely, such a F_ϵ can be constructed through interpolations between Q (with the *CSF* of such a F_ϵ satisfying (18)), and $\tilde{C} = Q + \beta_0(I - \beta_0 G^0)^{-1} G^0 \Phi$. Second, without shape restrictions $A\tilde{C} > 0$, the lemma would be vacuous, as *any* p in $(0,1)^M$ could be rationalized by some $(\tilde{C}, F_\epsilon)$. To see this, note it is always possible to construct Q, Φ recursively from any $p \in (0,1)^M$ and (17)-(20) and then define $\tilde{C}$ as $Q + \beta_0(I - \beta_0 G^0)^{-1} G^0 \Phi$.

Conventional structural analyses of probabilities for agreements in counterfactual contexts (such as perturbations in state transitions or the cake function) would take two steps. First, identify and estimate the true cake function $\tilde{c}_0$ and the USV distribution F_ϵ^0 using observable distributions; and second, use the identified parameters to predict counterfactual agreement probabilities induced in SSPE. Unfortunately, when the USV distribution is not restricted to have parametric forms, $\tilde{c}_0$ and F_ϵ^0 may not be uniquely recovered from observables, and the first step fails. Below we argue that nonetheless a simple algorithm can be used to recover *all* rationalizable counterfactual agreement probabilities. These are counterfactual conditional probabilities for agreements that are consistent with the model restrictions (including shape restrictions on $\tilde{c}$ and the independence of ϵ from X).

We first define rationalizable counterfactual outcomes. Suppose the data-generating process (DGP) is characterized by true parameters $\tilde{c}^0, F^0_\epsilon, \beta_0, G^0_{X'|X}$ which induce the conditional probabilities of agreement $p^0 \in [0,1]^M$ observed under SSPE. We are interested in predicting the probability for agreements under two types of counterfactual environments: (a) the transition between observable states is perturbed from $G^0_{X'|X}$ to $G^1_{X'|X}$ while $\beta_0, \tilde{c}^0, F^0_\epsilon$ 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 $G^0_{X'|X}, F^0_\epsilon$ remain the same. Suppose *CI-1,2* hold and the true discount rate β_0 is known.

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

The next proposition introduces a simple algorithm that recovers the *ISRCO*. The basic idea extends the preceding lemma by synthesizing two linear systems that characterize respectively the rationalizability in observed and counterfactual contexts. Such a synthesis exploits the fact that the nuisance F_ϵ is held fixed in both contexts. The observed p^0 and (unknown) counterfactual p^1 both enter the coefficient matrix of the "synthesized" linear system. Thus a $\tilde{p} \in [0,1]^M$ belongs to the *ISRCO if and only if* the $2M$-vector $(p^0, \tilde{p})$ makes the synthesized linear system feasible with solutions in the 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 α denote a M-by-M diagonal matrix with the m-th diagonal entry being $\alpha(x_m)$ as in (b) above.

Proposition 6 *Suppose the assumptions in Lemma 4 all hold. Then the ISRCO in (a) is the set of all p^1 such that the following linear system holds for some $(Q^j, \Phi^j)_{j=0,1}$:*

$$Q^0 + \beta_0(I - \beta_0 G^0)^{-1} G^0 \Phi^0 = Q^1 + \beta_0(I - \beta_0 G^1)^{-1} G^1 \Phi^1 \tag{21}$$

$$A\left[Q^0 + \beta_0(I - \beta_0 G^0)^{-1} G^0 \Phi^0\right] > 0 \tag{22}$$

$$Q^j_m \leq Q^k_n \Leftrightarrow p^j_m \leq p^k_n \;\forall m, n \in \{1, ., M\},\; j, k \in \{0, 1\} \tag{23}$$

$$\tilde{p}^{01}_{(m)}(\tilde{Q}^{01}_{(m+1)} - \tilde{Q}^{01}_{(m)}) \leq \tilde{\Phi}^{01}_{(m+1)} - \tilde{\Phi}^{01}_{(m)} \leq \tilde{p}^{01}_{(m+1)}(\tilde{Q}^{01}_{(m+1)} - \tilde{Q}^{01}_{(m)}) \textit{ for } 1 \leq m \leq 2M \tag{24}$$

$$Q^j_m \leq 0 \Leftrightarrow p^j_m \leq 1/2 \;;\; \Phi^j_m > 0 \;\forall m, n \in \{1, ., M\},\; j \in \{0, 1\} \tag{25}$$

And the ISRCO in (b) is the set of all p^1 *such that a similar linear system (22)-(25) and (26) is feasible with solutions* $(Q^j, \Phi^j)_{j=0,1}$*, where (26) is defined as*

$$Q^1 + \beta_0(I - \beta_0 G^0)^{-1} G^0 \Phi^1 = \alpha \left(Q^0 + \beta_0(I - \beta_0 G^0)^{-1} G^0 \Phi^0 \right) \tag{26}$$

Thus recovering the *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 ϵ and X and exogenously given shape restrictions. By construction, the *ISRCO* consists of all possible counterfactual outcomes that could be rationalized under the model restrictions (i.e. independence of ϵ of X and the shape restrictions).

We conclude this subsection 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 the known shape restrictions on the model. Thus the set reveals the limit of what can be learned about the counterfactual probability for agreements, if econometricians choose to remain agnostic about the functional form of the structural elements. In the following subsections, 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 ϵ is independent of X with a true USV distribution F_ϵ^0 that is uniform on $[-5, 5]$. Thus $q(p_k; F_\epsilon^0) = 10p_k - 5$ and $\phi(p_k; F_\epsilon^0) = 5p_k^2$ 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 $q(p_k; F_\epsilon^0)$ and $\phi(p_k; F_\epsilon^0)$ respectively. Let the discount rate β 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

$$\tilde{C}_0 \equiv \tilde{C}_{unif} = [\ \tfrac{717442573}{165078240},\ \tfrac{97368349}{132062592},\ \tfrac{330851369}{264125184}] \approx [4.3461,\ 0.7373,\ 1.2526]$$

while the actual conditional probability for reaching agreement observed in the DGP is $p^0 = [\frac{3}{5}, \frac{1}{4}, \frac{5}{16}]$.[11] 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

$$Q^{unif}(p^1_{unif}) + \beta(I - \beta G^1)^{-1} G^1 \Phi^{unif}(p^1_{unif}) = \tilde{C}_{unif} \tag{27}$$

The solution is found to be $p^1_{unif} \approx [0.5880\ ,\ 0.2004\ ;\ 0.2760]$.[12]

(Innocuous location and scale normalizations) Now suppose econometricians only know USV is uniformly distributed, but misspecify 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

$$\hat{C}_{unif} = \hat{Q}^{unif}(p^0) + \beta(I - \beta G^1)^{-1} G^1 \hat{\Phi}^{unif}(p^0) \tag{28}$$

Straightforward substitutions show this misspecification still leads to the same system of nonlinear equations in p^1_{unif} as (27). In other words, even though $Q^{unif}(.)$, $\Phi^{unif}(.)$ and $\tilde{C}_{unif}$ have different forms now due to the misspecification of F_ϵ, the structure of the model is such that the differences cancel out and yield the same system of nonlinear equations in (27). (See the Appendix for algebraic details.) This verifies our remarks earlier (following

[11] 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}, Φ^{unif} are defined above.

[12] 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 π_w is unique.

Lemma 4) that the scale and locational normalizations of the USV distribution is innocuous for recovering counterfactuals.

(Consequence of misspecifying USV distributions) Suppose β, 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.

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

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

$$Q^{GLL}(p^1_{GLL}) + \beta(I - \beta G^1)^{-1} G^1 \Phi^{GLL}(p^1_{GLL}) = \tilde{C}_{GLL} \tag{29}$$

Solving (27) 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.[13] 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.

(Recovering the ISRCO when the USV distribution is unknown) Now let the true β_0, G^0 underlying the DGP be defined as above, with F^0_ϵ uniform on $[-5, 5]$ and the conditional median cake function $\tilde{C}_0 = \tilde{C}_{unif}$. 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^0_ϵ or the true $\tilde{C}_0$. They only observe p^0 and know β_0, G^0 in the DGP, and are interested in predicting the

[13] 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.

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. $\tilde{c}_0(x_2) < \min\{\tilde{c}_0(x_1), \tilde{c}_0(x_3)\}$. Then the algorithm proposed above can be used to recover the *ISRCO* by collecting all $p^1 \in [0,1]^3$ that make the linear system (21)-(25) feasible. (See the Appendix for details in implementing the algorithm.) Figure 1 depicts the *ISRCO* recovered is about 5.1% of the outcome space $[0,1]^3$.

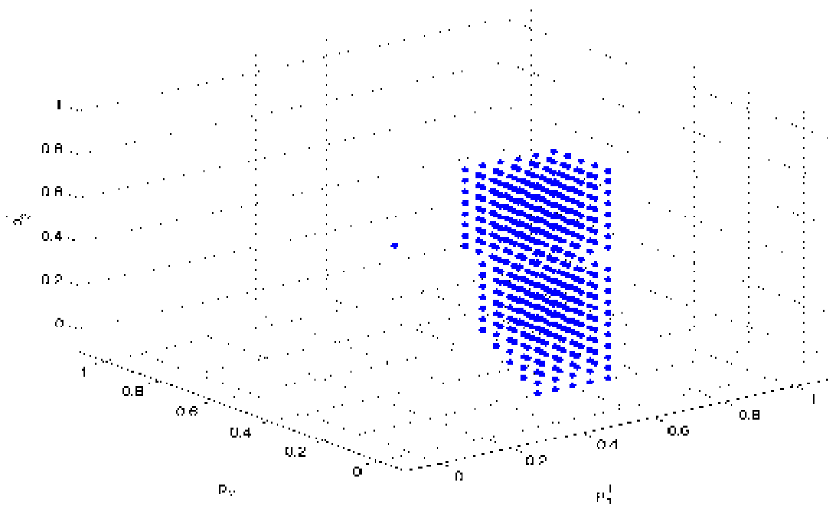

Figure 1: *ISRCO* with $\epsilon \perp X$, $\tilde{c}_{02} < \min(\tilde{c}_{01}, \tilde{c}_{03})$ and p^0=[3/5,1/4,5/16] (where $\tilde{c}_{0k} \equiv \tilde{c}_0(x_k)$)

Our algorithm for recovering the *ISRCO* only requires ϵ 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 the *ISRCO*; and (ii) any outcome vector in the *ISRCO* is a rationalizable counterfactual outcomes corresponding to certain F_ϵ that satisfies independence from X and some $\tilde{C}$ such that $\tilde{C}_2 < \min\{\tilde{C}_1, \tilde{C}_3\}$. Also note in implementing the algorithm we have invoked a location normalization ($q(1/2; F_\epsilon^0) = 0$) and a scale normalization ($\phi(1/2; F_\epsilon^0) = \bar{\phi} > 0$), which are known to be innocuous for counterfactual analyses.

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 factor. Lemma 1 shows the payoffs in stationary SPE is unique under these restrictions. In this section we study the identification when players evaluate the surplus according to a concave utility function, or the discount rates differ across players. In either case, players' payoffs from SSPE are not unique in general. We shall show the utility function and the discount rates can be identified from complete data on states, agreements and agreed allocations under an additional assumption.

SE (Single equilibrium payoff) Players in observed bargaining games observed all adopt strategies that lead to the same profile of SSPE payoffs.

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)). It allows econometricians to relate observable distributions to model primitives through (14) and (15), without having to specify which equilibrium is followed in the presence of multiple SSPE.

6.1 Concave Utility Functions

In this subsection, we extend the basic model with complete information by relaxing restrictions of linear utilities. The set of feasible allocations is now given by $C(s) = \{m \in \mathbb{R}^K : \sum_i u^{-1}(m_i) \leq c(s)$ 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 an agreement occurs, but do not know the utility levels associated with these shares. The lemma below characterizes the SSPE payoffs in this model.

Lemma 5 *Suppose CI-1 holds. Then (a) $v \in F^K$ is a SSPE payoff in the bargaining game with general utilities if and only if $A(v) = v$ where for all $\tilde{s} = (s, \rho) \in \Omega_{\tilde{S}}$ and for all i,*

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

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

$$u\left(c(s) - \sum_{j \neq i} u^{-1}(\beta E[v_j(\tilde{S}')|S = s])\right) \geq \beta E[v_i(\tilde{S}')|S = s] \tag{30}$$

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.

ND' $\Pr\{c(S) - \sum_{j=1}^{K} u^{-1}(\beta\pi_j(X)) \geq 0$, $\rho_{(1)} \neq i \mid X = x\} > 0$ *for all* x.

PS (Parameter space) The parameter space for utility function (denoted Θ_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 $u, \tilde{u}$).

CI-3 (Conditional independence of order of moves) For all t, the order of moves ρ_t is independent of ϵ_t given any x_t.

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. The role of *ND'* is analogous to condition (ii) in *ND* — it makes sure that for i, the physical divisions of the cake that he receives in any state x can always be observed in data with positive probability. Under *CI-3*, the order of moves is independent from the size of the cake given x.

Proposition 7 *Suppose β is fixed and CI-1,2,3, SE, MT, ND', PS hold. Then $\nexists u \neq \tilde{u}$ in Θ_U such that $u \overset{o.e.}{\sim} \tilde{u}$.*

A corollary of the proposition is that u is identified within the class of increasing functions with either constant absolute risk aversions ($CARA$) or constant relative risk aversions ($CRRA$) and $u(0) = 0$. 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 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 β_i. The lemma below characterizes the SSPE payoffs in this case.

Lemma 6 *Suppose CI-1 holds. Then $f \in F^K$ is a SSPE payoff if and only if $A(f) = f$ where for all $(s, \rho) \in \Omega_{S,\rho}$,*

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

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 β_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 the case with complete data, the basic idea underlying the identification of individual β_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 8 *Under CI-1,2, SE, MT and ND, the discount factors $\{\beta_i\}_{i=1}^K$ are identified.*

As before, *MT* ensures there exists a one-to-one mapping between Y and ϵ. *ND* 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 β_i. Then the probability of agreements and agreed shares of the cake can still be related to discount rates as theory predicts in Lemma 6.[14]

[14] *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.

Proposition 1 and Proposition 8 differ in that the former only uses the distribution of total cake sizes and the probability of agreements while the latter also exploits actual allocations under agreement. This difference is consistent 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 β_i's. In fact, the homogenous β_0 in Proposition 1 is over-identified in the sense that observing the total cake size alone is sufficient for identifying the single β_0, 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 β_0.

7 Conclusion

In this paper we have presented 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 starting point in this paper is a group of conditional independence restrictions *CI-1,2*. 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.[15] 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.

[15] For example, under *CI-1,2,3*, the ex ante individual continuation payoffs π_i in SSPE must be independent of ϵ, ρ 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.

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,ρ),

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

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). □

Proof of Lemma 2. *(Necessity)* Suppose $\exists\{\beta, c, \tilde{L}_{\rho|S}, F_{\epsilon|X}\}$ that satisfies MT and ND and rationalizes the distribution of $\{\tau, \eta_\tau, Y^\tau, X^\tau, \rho^\tau\}$. Recall $Y_t = c(X_t,\epsilon_t)$ and by Lemma 1, $D_t = 1$ if and only if $Y_t \geq E(w(S_{t+1})|s_t)$ in any SSPE. Hence under MT, the equality in (5) is implied by $F_{\epsilon_{t+1},\rho_{t+1},X_{t+1}|\varepsilon^t,x^t,\rho^t} = F_{\epsilon_{t+1},\rho_{t+1},X_{t+1}|X_{t+1}}\ G_{X_{t+1}|x_t}$, which follows immediately from *CI-1,2*. The time-homogeneity of $F_{Y_t,D_t,\rho_t|X_t}$ follows from time-homogeneity of $\tilde{L}_{\rho|S}$ and $H_{S'|S}$ in *CI-1*. Under *CI-2*, $E(w(S')|s) = E(w(S')|x)$ and hence $p(x) = \Pr\{Y \geq \beta E(w(S')|x) \mid x\} = 1 - F_{Y|X=x}(\beta E(w(S')|x))$. Under MT, $F_{Y|X=x}(t) = \Pr\{c(X,\epsilon) \leq t \mid X = x\} = F_{\epsilon|X=x}(c^{-1}(x,t))$ is strictly increasing in t on the support of Y conditional on x (where $c^{-1}(x,.)$ is the inverse of $c(x,.)$ given x). Under ND, $p(x) \in (0,1)$

for all x, and $\beta E(w(S')|x) = \lambda^*(x)$. Then by Lemma 1, conditions *(ii)* and *(iii)* must hold in any SSPE. Next note by definition

$$\begin{aligned}
\beta E\left[w(S')|x\right] &= \beta \int\int \max\{c(x',\varepsilon'), \beta E\left[w(S'')|x'\right]\} dF_{\epsilon'|X'} dG_{X'|X=x} \\
\Leftrightarrow \lambda^*(x) = & \quad \beta \int\int \max\{c(x',\varepsilon'), \lambda^*(x')\} dF_{\epsilon'|X'} dG_{X'|X=x} \qquad (31)\\
&= \beta \int\int \max\{y', \lambda^*(x')\} dF_{Y'|X'} dG_{X'|X=x} \qquad (32)
\end{aligned}$$

where the equality in (31) follows from substituting $\beta E\left[w(S')|x\right]$ with $\lambda^*(x)$. The equality in (32) follows from from a change-of-variable between y and ε given x which is feasible under MT. Since $\beta \in (0,1)$, condition *(iv)* must hold. Furthermore, the assumption ND implies $\Pr\{D = 1, \rho_{(1)} \neq i \mid X = x\} \in (0,1)$ in any SSPE. Note the individual's continuation payoffs in SSPE, i.e. $\pi_i(s,\rho) \equiv \beta E[v_i(S',\rho')|s,\rho]$, are defined by the unique solution to the fixed point equation:[16]

$$\pi_i(s,\rho) = \beta \int \pi_i(s',\rho') + 1\{\rho'_{(1)} = i\} \max\{c(s') - \beta \textstyle\sum_i \pi_i(s',\rho'), 0\} dF_{S',\rho'|s,\rho} \qquad (33)$$

Under *CI-1,2*, (S',ρ') is independent of (ϵ,ρ) given X, and π_i, as a solution to the fixed-point equation in (33), must be a function of x alone. Thus, (33) can be written as

$$\pi_i(x) = \beta \int \pi_i(x') + \int 1\{\rho'_{(1)} = i\} \max\{y' - \beta \textstyle\sum_i \pi_i(x'), 0\} dF_{Y',\rho'|X'} dG_{X'|X=x} \qquad (34)$$

using change-of-variables between Y' and ϵ', where a one-to-one mapping exists given X'. Under any agreement in SSPE, a non-proposer is offered his discounted continuation payoff $\pi_i(x)$. Hence condition *(v)* follows from (34) with λ_i playing the role of π_i and $E[w(S')|s] \equiv \sum_i \pi_i(x) = \lambda^*(x)$ for all x.

(Sufficiency) Suppose a joint distribution of $\{\tau, \eta_\tau, Y^\tau, X^\tau, \rho^\tau\}$ satisfies conditions *(i)-(v)*. We need to find a set of primitive elements $\{\beta, c, \tilde{L}_{\rho|S}, F_{\epsilon|X}\}$ that satisfy MT and ND, and could rationalize this joint distribution under *CI-1,2*. First, choose *any* strictly increasing distribution $F_{\epsilon|X=x}$ and define $c(x,\varepsilon) \equiv F^{-1}_{Y|X=x}(F_{\epsilon|X=x}(\varepsilon))$. By condition *(ii)*, $F_{Y|X=x}(t)$ is increasing in t on the support of Y conditional on x, and therefore $c(x,\varepsilon)$ is increasing in ε on the support of ϵ conditional on x, and MT is satisfied. Furthermore,

$$\begin{aligned}
&\Pr\{c(X,\epsilon) \leq y|X = x\} \equiv \Pr\{F^{-1}_{Y|X=x}(F_{\epsilon|X=x}(\epsilon)) \leq y|X = x\} \\
= &\Pr\{F_{\epsilon|X=x}(\epsilon) \leq F_{Y|X=x}(y)|X = x\} = F_{Y|X=x}(y)
\end{aligned}$$

[16] See Lemma 2 of Merlo and Wilson (1998) for details.

for all y on the support conditional on x. The equalities follow from the condition on $F_{Y|X=x}$ in *(ii)*, and the fact that $F_{\epsilon|X=x}(\epsilon)$ is uniform on $[0,1]$ conditional on $X=x$. Next, define $\beta \equiv \left(\int \max\{y', \lambda^*(x')\}dF_{Y',X'|X=x}\right)^{-1}\lambda(x)$. Under condition *(iv)*, β is a well-defined discount factor between $(0,1)$, and $\tilde{L}_{\rho|x,\varepsilon}$ is defined as $\Pr\{\textit{the order of moves is } \rho \mid X=x, Y=c(x,\varepsilon)\}$. By construction, $\lambda^*(x)$ is the *unique* solution for the following fixed point equation:

$$f(x) = \beta \int\int max\{y' \ , \ f(x')\}dF_{Y'|X'}dG_{X'|X=x} \tag{35}$$

where the r.h.s. of (35) is a contraction mapping. Using change-of-variables between Y' and ϵ' conditional on X', (35) can be written as

$$f(x) = \beta \int\int \max\{c(x',\varepsilon') \ , \ f(x')\}dF_{\epsilon'|X'}dG_{X'|X=x}$$

which also has a unique solution in f. Hence $\lambda^*(x) = \beta E[w(S')|s]$. Then:

$$\Pr\{Y \geq \beta E[w(S')|X] \mid X=x\} = \ \Pr\{Y \geq \lambda^*(x) \mid X=x\} = \ \Pr\{D=1 \mid X=x\} \equiv p(x)$$

where the second equality follows from condition *(iii)*. Next define $\tilde{L}_{\rho|S=(x,\varepsilon)} = F_{\rho|X=x,Y=c(x,\varepsilon)}$. The restrictions in ND are satisfied because $p(x) \in (0,1)$ in condition *(iii)*, and because $\Pr\{D=1, \rho_{(1)} \neq i \mid X=x\} \in (0,1)$ for all i,x in condition *(v)*. Also under MT, condition *(v)* can be written as

$$\lambda_i^*(x) = \beta \int \lambda_i^*(x') + \max\{c(x',\varepsilon') - \lambda^*(x'), 0\}1\left(\rho'_{(1)} = i\right) dF_{\epsilon',\rho'|X'}dG_{X'|X}$$

and hence $\lambda_i^*(x)$ is the unique solution of the fixed point equation that defines individual continuation payoffs $\beta E[v_i(S',\rho')|s,\rho]$ under *CI-1,2* and MT. Hence conditions *(iii)*, *(iv)* and *(v)* ensure time-homogeneous conditional distributions $F_{\rho,D,Y|X}$ and $F_{\eta_\tau|\rho_\tau,D_\tau=1,X_\tau}$ observed from the data can be rationalized by the set of $\{\beta, c, \tilde{L}_{\rho|S}, F_{\epsilon|X}\}$ chosen above. Finally, construct the full transition of states subject to restrictions in *CI-1,2* by defining for all $t \geq 0$:

$$\tilde{H}_t(s_{t+1},\rho_{t+1}|s^t,\rho^t) \equiv \ \tilde{L}_{\rho|S}(\rho_{t+1}|s_{t+1})F_{\epsilon|X}(\varepsilon_{t+1}|x_{t+1})G_{X'|X}(x_{t+1}|x_t)$$

Since (5) holds, and $F_{Y,\rho,D|X}$, $G_{X'|X}$ are time-homogenous by condition *(i)*, straightforward inductive arguments show $\{\beta, c, \tilde{H}_t\}$ rationalizes $\{\tau, \eta_\tau, X^\tau, Y^\tau, \rho^\tau\}$ as long as $\{\beta, c, \tilde{L}_{\rho|S}, F_{\epsilon|X}\}$ rationalizes $F_{Y,D,\rho|X}$ and $F_{\eta_\tau|\rho_\tau,D_\tau=1,X_\tau}$. This completes the proof. $\square$

Proof of Proposition 1. By Lemma 1, $p(x) \equiv \Pr(D=1|X=x) = \Pr\{Y \geq \beta\pi_w(S)|\ X=x\}$ $= 1 - F_{Y|X=x}(\beta\pi_w(x))$, where $\pi_w(S) \equiv E[w(S')|S]$ is the total continuation payoff for all players, and must be a function of X alone under *CI-1,2*. By construction, π_w solves the fixed point equation:

$$\beta\pi_w(x) = \beta\int \max\{c(x',\varepsilon'),\beta\pi_w(x')\}dF_{\epsilon|X=x'}(\varepsilon')G(x'|x) \tag{36}$$

$$= \beta\int \beta\pi_w(x')+\phi(x')dG(x'|x) \tag{37}$$

where

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

Since $F_{Y|X=x}$ is increasing under MT and $p(x)\in(0,1)$ under *ND*, $\beta\pi_w(x)=F^{-1}_{Y|X=x}(1-p(x))$ $\equiv\lambda^*(x)$, which is a conditional quantile of the observed distribution of surplus as defined in Lemma 2. Under MT, y is increasing in ε given x. This implies ϕ can be expressed in terms of observable distributions by changing variables between y and ε:

$$\tilde{\phi}(x;\lambda^*,F_{Y|X}) \equiv \int_{\lambda^*(x)} y-\lambda^*(x)dF_{Y|X=x}$$

Thus (37) can be written as

$$\lambda^*(x)=\beta\int \lambda^*(x')+\tilde{\phi}(x';\lambda^*,F_{Y|X})dG(x'|x)$$

and β is identified as $\left(\int \max\{y',\lambda^*(x')\}dF_{Y',X'|X=x}\right)^{-1}\lambda^*(x)$. $\square$

Proof of Lemma 3. Necessity follows from the definition of rationalization. It has been shown in the proof of Lemma 2 that the unique total and individual continuation payoffs must be functions of x alone under *CI-1,2*. Under these assumptions and MT, the discounted total continuation payoff $\beta E[w(S')|s]$ in SSPE can be expressed in terms of observables as the unique solution for the fixed-point equation:

$$f(x)=\beta\int \max\{y'\ ,\ f(x')\}dF_{Y',X'|X=x} \tag{38}$$

Individual continuation payoffs can also be expressed in terms of observables as the unique solution in f_i for the fixed-point equation:

$$f_i(x)=\beta\int f_i(x')+\int \max\{y'-f(x'),0\}1\left(\rho'_{(1)}=i\right)dF_{Y',\rho'|X'}dG_{X'|X=x} \tag{39}$$

with f given in (38) above. Note both (38) and (39) are completely defined by $F_{Y,\rho|X}$, β and $G_{X'|X}$. Recall that the true discount rate β_0 is identified by Proposition 1. Thus if $F^{(\sigma)}_{Y,\rho|X}=F^*_{Y,\rho|X}$, then σ must generate the same total and individual continuation payoff functions as those implied by $F^*_{Y,\rho|X}$ (i.e. λ^* and λ_i in Lemma 2). It is then straightforward to show that $F^{(\sigma)}_{Y,D,\rho|X} = F^{*}_{Y,D,\rho|X}$ a.e. on Ω_X, and that the distribution of $F_{\eta|\rho,D=1,X}$ implied by σ also matches the $F^{*}_{\eta_\tau|\rho,D=1,X}$ observed in $F^{*}_{\{\tau,\eta_\tau,X^\tau,Y^\tau,D^\tau\}}$.

The rest of the proof uses inductive arguments. With the initial distribution of X_0 observed, any σ that rationalizes the time-homogeneous distribution $F^{*}_{Y,\rho,D|X}$ must also by definition rationalize the joint distribution of observables with $\tau = 0$. Now consider the case $\tau = 1$. Then the observed distribution can be written as

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

where the second equality follows from the necessary conditions for $F^{*}_{\{\tau,\eta_\tau,X^\tau,Y^\tau,D^\tau\}}$ to be rationalizable under *CI-1,2*. Recall $G_{X'|X}$ is directly observed from data, and σ rationalizes $F^{*}_{\eta_1|D_1=1,\rho_1,X_1}$, $F^{*}_{Y_1,D_1,\rho_1|X_1}$and $F^{*}_{Y_0,D_0,\rho_0|X_0}$. Hence σ also rationalizes the joint distribution of observables with $\tau = 1$.

Now suppose σ rationalizes the observable distribution for $\tau \leq t$. Consider the case with $\tau = t + 1$,

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

where the second equality again follows from necessary conditions for rationalizability under *CI-1,2*. By supposition at the beginning of this induction step, σ rationalizes the first and last terms in the product. $\square$

Proof of Proposition 3. Under *CI-1* and *MH*, the total continuation payoff must be a function of x alone. Let $\theta_0 = (\beta_0, \tilde{c}_0, F^0_{\epsilon|X}, L^0_{\rho|S})$ denote the true parameter underlying the DGP. By definition, the true discounted total continuation payoff $\beta_0\pi_w(.;\theta_0)$ is the unique fixed-point of a contraction mapping:

$$\beta_0\pi_w(x;\theta_0) = \beta_0 \int \max\{\tilde{c}_0(x') + \varepsilon', \beta_0\pi_w(x';\theta_0)\} dF^0_{\epsilon'|X'} dG_{X'|X=x}$$

Under RG, $\beta_0\pi_w(.;\theta_0) \in B(\Omega_X)$. Hence MH implies $\Pr\{\tilde{c}_0(X) - \beta_0\pi_w(X;\theta_0) + \epsilon \geq 0|X = x\} \in (0, 1)$, and $\beta_0\pi_w(x;\theta_0)$ is identified for all x as the infimum of support of the cake size under agreements. Then by arguments similar to Proposition 1, the common discount rate β_0 is identified as in

$$\beta_0 = \left(\int_{\Omega_{X'}} \int_{\Omega_{Y'|X'}} D'y'dF^*_{Y'|X'} + (1-D')\tilde{\lambda}(x')dG_{X'|X=x} \right)^{-1} \lambda^*(x)$$

where $D = 1$ if an agreement is reached, and $\tilde{\lambda}(x)$ now denotes the infimum of conditional support of $F^*_{Y|D=1,X=x}$. The realized "gains to the proposer" is defined as:

$$y^* \equiv y - \beta_0 \pi_w(x;\theta_0) = \max\{\tilde{c}_0(x) - \beta_0 \pi_w(x;\theta_0) + \varepsilon, 0\}$$

and $\gamma_0(x) \equiv \tilde{c}_0(x) - \beta_0 \pi_w(x;\theta_0)$ must be continuous and bounded under the additive separability assumption. The identification of $\gamma_0(x)$ on Ω_X under SG follows from Chen, Dahl and Khan (2005). Then $\tilde{c}_0$ is recovered as $\tilde{c}_0(.) = \gamma_0(.) + \beta_0 \pi_w(.;\theta_0)$ on Ω_X. To prove (ii), note "$\gamma_0(x) > 0$" is equivalent to "$\Pr\{\epsilon > -\gamma_0(x)|x\} > \frac{1}{2}$" since $q_{0.5}(\epsilon|x) = 0$ for all x. Hence SG is equivalent to $\Pr\{\Pr(D=1|X) > \frac{1}{2}\} > 0$. $\square$

Proof of Proposition 4. Under *CI-1,2* and *AS*, $\beta_0 \pi_w(x;\theta_0)$ is identified over Ω_X using arguments as in Proposition 3. Let $q_\alpha(Y^*|x)$ denote the α-th quantile of the observable, realized gains to the proposer $Y^* \equiv \max\{Y - \beta_0 \pi_w(X;\theta_0)\ ,\ 0\}$ given $X = x$. For $x \in \Omega_X^+$, $\gamma_0(x) = q_{0.5}(Y^*|x)$ is identified. Consider any $x = (x_a, x_b)$ s.t. $x \notin \Omega_X^+$. By RG and RS, γ_0 is bounded over Ω_X while support of ϵ given x_a is unbounded. Then $\exists \bar{\alpha} > 1/2$ (possibly dependent upon x) and close enough to 1 s.t. $q_{\bar{\alpha}}(Y^*|x) = \gamma_0(x) + c^0_{\bar{\alpha}}(x) > 0$ is observed, where $c^0_{\bar{\alpha}}(x) = c^0_{\bar{\alpha}}(x_a)$ denotes the true $\bar{\alpha}$-th quantile of ϵ conditional on $x \equiv (x_a, x_b)$, and is independent of x_b by ER. Now pick $\tilde{x} = (x_a, \tilde{x}_b)$ such that $q_{1/2}(Y^*|\tilde{x}) = \gamma_0(\tilde{x}) > 0$ is observable. Such a choice of $\tilde{x}$ is possible because of SG'. Hence $\gamma_0(\tilde{x})$ is identified. Since $\bar{\alpha} > 1/2$ and $c^0_{\bar{\alpha}}(x) = c^0_{\bar{\alpha}}(\tilde{x})$ under ER, $q_{\bar{\alpha}}(Y^*|\tilde{x})$ must also be positive, observable, and equal to $\gamma_0(\tilde{x})+ c^0_{\bar{\alpha}}(x_a)$. Hence (with a slight abuse of notation) $c^0_{\bar{\alpha}}(x_a) = c^0_{\bar{\alpha}}(\tilde{x})$ is recovered as $q_{\bar{\alpha}}(Y^*|\tilde{x})-\ \gamma_0(\tilde{x})$. This implies $\gamma_0(x)$ can then be recovered as $q_{\bar{\alpha}}(Y^*|x) - c^0_{\bar{\alpha}}(x_a)$ for any $x \notin \Omega_X$. $\square$

Proof of Proposition 5. The total continuation payoff satisfies

$$\pi_w(x) = \int \beta \pi_w(x') + \int \max\{\tilde{c}(x') - \beta \pi_w(x') - \varepsilon', 0\} dF_{\epsilon'|X'} dG_{X'|X}$$

Define $p(x) \equiv \Pr(D=1|x) = F_{\epsilon|X=x}(\tilde{c}(x) - \beta \pi_w(x))$. Hence $\tilde{c}(x) - \beta \pi_w(x) = F^{-1}_{\epsilon|X=x}(p(x))$, as $\Pr(D=1|x) \in (0,1)$ for all x. Define

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

Then

$$\beta\pi_w(x) = \beta \int \phi(x'; p, F_{\epsilon|X}) + \beta\pi_w(x')dG(x'|x) \tag{40}$$

For any $F_{\epsilon|X}$ given and p observed, the right hand side of (40) is a contraction mapping. Therefore with $F_{\epsilon|X}$, β known and p, $G_{X'|X}$ observed, the discounted total continuation payoff $\beta\pi_w$ is uniquely recovered as the solution to the fixed point equation of (40). Hence $\tilde{c}$ is identified with knowledge of $F_{\epsilon|X}$.

To prove (ii), note with $F_{\epsilon|X}$ known and p observed, the discounted total continuation payoff $\beta\pi_w$ is the unique solution of the fixed point equation in (40). With $\phi(x; p, F_{\epsilon|X})$ fixed (because $p, F_{\epsilon|X}$ are known), denote the solution to (40) for a generic $\hat{\beta}$ as $\hat{\beta}\pi_w(.; \hat{\beta})$, which must be increasing in $\hat{\beta}$ for all x. We can then solve the equation:

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

where the r.h.s. can be calculated from observables with the assumed knowledge of $\tilde{c}$ at $\bar{x}$. Once β is identified, so is $\tilde{c}$ using arguments in the proof of (i). $\square$

Proof of Lemma 4. (*Necessity*) Suppose a vector p is generated by some true parameters $(\tilde{c}_0, F^0_\epsilon)$ underlying the DGP such that $A\tilde{C}_0 > 0$ and ϵ is independent of X with median 0. Then let $\tilde{Q}_m = F^{0,-1}_\epsilon(p_m) = q(p_m; F^0_\epsilon)$ and $\tilde{\Phi}_m = \phi(p(x_m); F^0_\epsilon)$. It follows immediately from the substitution of (15) into (14), the independence of ϵ from X and the monotonicity of F^0_ϵ that (16) and (17) must hold for $\tilde{Q}, \tilde{\Phi}$. The definition of ϕ and some straightforward algebra (involving the Leibniz rule for differentiating integrals) suggest for any m, n,

$$\phi(p_m) - \phi(p_n) = \int_{\tilde{Q}_n}^{\tilde{Q}_m} F^0_\epsilon(\varepsilon)d\varepsilon$$

which must be bounded between $p_n(\tilde{Q}_m - \tilde{Q}_n)$ and $p_m(\tilde{Q}_m - \tilde{Q}_n)$. Hence (18) holds for $\tilde{Q}, \tilde{\Phi}$. Note (20) holds for $\tilde{\Phi}$ by definition of the CSF ϕ, and (19) holds for $\tilde{Q}, \tilde{\Phi}$ if $\bar{\phi}$ is equal to the true CSF at $\frac{1}{2}$, i.e. $\tilde{\phi} \equiv \phi(1/2; F^0_\epsilon)$. More generally, if $\bar{\phi} \neq \tilde{\phi}$, the system (16)-(20) 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 (16)-(20) holds for some Q, Φ then there must be a pair $(\tilde{c}, F_\epsilon)$ such that (i) ϵ 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_ϵ by choosing the p_m-percentile $F^{-1}_\epsilon(p_m)$ to be the solutions Q_m and choosing $\phi(p(x_m))$ by first setting $\phi(1/2) = \bar{\phi}$ and then interpolating between $F^{-1}_\epsilon(p_m)$ so that $\phi(p(x_m))$ is equal to the solution Φ_m. This is possible because the inequality restrictions (18) and (19) 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 (16)- (20). 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. □

Proof of Proposition 6. The distribution of unobservable states F^0_ϵ 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 4 and is omitted for brevity. □

Proof of Proposition 7. Arguments similar to above show that under *CI-1,2*, the individual continuation payoff in SSPE, given parameters $u, F_{\epsilon|X}, L_{\rho|X}, G_{X'|X}$, can be written as

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

where $\kappa \equiv \rho_{(1)}$ is the identity of the proposer. Let $q_i(x) \equiv \Pr(\kappa = i|x)$. Note the first term on the right-hand side can be written as

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

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*, $F_{\epsilon',\kappa'|X',X} = F_{\kappa'|X'}F_{\epsilon'|X'}$. Furthermore, under *MT*, Y is increasing in ϵ given X and $\tilde{q}_i(x,\varepsilon) = q_i(x, c(x,\varepsilon))$ for all (x,ε). Hence

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

where $F_{Y|X}$ is the distribution induced by $u, c, F_{\epsilon|X}$. Likewise, under *CI-1,2,3*, 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 Ψ is a $\mathbb{R}^k$-valued function with the i-th coordinate Ψ_i defined as

$$\begin{aligned}
\Psi_i(x;\pi) \quad \equiv \quad & \int q_i(x')\lambda(x';u,\pi_{-i})dG(x'|x) + \int (1-q_i(x'))\beta\pi_i(x')dG(x'|x) + \\
& \int q_i(x')\beta\pi_i(x') \int 1\{y' < \textstyle\sum_i u^{-1}(\beta\pi_i(x'))\}dF_{Y|X}(y'|x')dG(x'|x)
\end{aligned} \tag{41}$$

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 ND' implies that for each individual i and observable state x, there is positive probability that an agreement is reached when he is not the proposer. Hence for each player i, $\psi_i(x)$ is observed over the support Ω_X as the physical shares for player i when agreements occur with $\kappa \neq i$ and $X = x$.

Define $y_i^* \equiv y - \sum_{j\neq i}\psi_j(x)$ for all i. Alternatively, (41) can be written as

$$\beta\pi_i(x) = \beta\int q_i(x')p_1(x')\bar{\lambda}_i(x') + \left[q_i(x')p_0(x') + 1 - q_i(x')\right]\beta\pi_i(x')dG(x'|x) \tag{42}$$

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 (42) as a "quasi-fixed-point equation" for $\beta\pi_i(x)$. Compared with (41), (42) differs in that it explicitly expresses how the observed quantities $\{\psi_i\}_{i=1}^k$, 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 for the set of parameters that are restricted to be observationally equivalent (which is the case in the statement of our lemma).

We complete the proof by contradiction. Suppose there exists $u \neq \tilde{u}$ in Θ_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 Ω_X,

$$\psi_i(x; u) \equiv u^{-1}(\beta \pi_i(x; u)) = \tilde{u}^{-1}(\beta \tilde{\pi}_i(x; \tilde{u})) \equiv \psi_i(x; \tilde{u}) \tag{43}$$

$$p_1(x; u) \equiv \Pr(D = 1|x; u) = \Pr(D = 1|x; \tilde{u}) \equiv p_1(x; \tilde{u}) \tag{44}$$

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 $u \overset{o.e.}{\sim} \tilde{u}$,

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

where the inequalities all follow from concavity of g and applications of Jensen's Inequality as well as (43) and (44). In addition, the last inequality also uses $g(0) = 0$ (implied by $u(0) = 0$ for all $u \in \Theta_U$). This constitutes a contradiction. 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 Θ_u such that $u \overset{o.e.}{\sim} \tilde{u}$. $\square$

Proof of Proposition 8. Let $v \equiv (v_j)_{j=1}^K$ be the vector of individual SSPE payoffs that solve the fixed-point equation in Lemma 6. Note under *CI-1,2*,

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

Under *ND,* the discounted individual continuation payoff $\beta_i \pi_i(x) \equiv \beta_i E[v_i(S', \rho')|X = x]$ is observed as $\psi_i(x)$ (i.e. the division of the cake received by i when he does not propose) for all i and x. By definition,

$$\begin{aligned} \beta_i \pi_i(x) &\equiv \beta_i \int \begin{pmatrix} E[v_i(S', \rho')|\kappa' = i, x'] \Pr(\kappa' = i|x') + \\ E[v_i(S', \rho')|\kappa' = i, x'] \Pr(\kappa' \neq i|x') \end{pmatrix} dG(x'|x) \\ &= \beta_i \int \beta_i \pi_i(x') + \int 1(\kappa' = i) \max\{c(x', \varepsilon') - \textstyle\sum_j \beta_j \pi_j(x'), 0\} dF_{\kappa', \epsilon'|X'=x'} dG(x'|x) \end{aligned} \tag{45}$$

Under *MT*, Y is increasing in ϵ given X, and using a change of variables between Y and ϵ, (45) can be written as

$$\psi_i(x) = \beta_i \int \psi_i(x') + \int 1(\kappa' = i) \max\{y' - \textstyle\sum_j \psi_j(x'), 0\} dF_{\kappa', Y'|X'} dG(x'|x)$$

Hence β_i is identified for all i as

$$\beta_i = \left(\int \psi_i(x') + \tilde{\phi}_i(x') dG(x'|x) \right)^{-1} \psi_i(x)$$

where $\tilde{\phi}_i(x; \psi, F_{\kappa, Y|X}) \equiv \int 1(\kappa = i) \max\{y - \sum_{j=1}^{K} \psi_j(x), 0\} dF_{Y|X=x}$. $\square$

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 (27) is:

$$\begin{aligned} \tfrac{49567930}{6444299} p_1^2 + \tfrac{23508550}{6444299} p_2^2 + \tfrac{55809500}{6444299} p_3^2 + 10 p_1 - 5 &= \tfrac{717442573}{165078240} \\ \tfrac{55814500}{6444299} p_1^2 + \tfrac{22413980}{6444299} p_2^2 + \tfrac{50657500}{6444299} p_3^2 + 10 p_2 - 5 &= \tfrac{97368349}{132062592} \\ \tfrac{50090600}{6444299} p_1^2 + \tfrac{22567500}{6444299} p_2^2 + \tfrac{56227880}{6444299} p_3^2 + 10 p_3 - 5 &= \tfrac{330851369}{264125184} \end{aligned} \tag{46}$$

(Innocuous location and scale normalizations) For example, suppose $a = 3, b = 2$. Then the nonlinear system in (28) is

$$\begin{aligned} \tfrac{29740758}{6444299} p_1^2 + \tfrac{14105130}{6444299} p_2^2 + \tfrac{33485700}{6444299} p_3^2 + 6 p_1 - 1 &= \tfrac{1267703373}{275130400} \\ \tfrac{33488700}{6444299} p_1^2 + \tfrac{13448388}{6444299} p_2^2 + \tfrac{30394500}{6444299} p_3^2 + 6 p_2 - 1 &= \tfrac{537576989}{220104320} \\ \tfrac{30054360}{6444299} p_1^2 + \tfrac{13540500}{6444299} p_2^2 + \tfrac{33736728}{6444299} p_3^2 + 6 p_3 - 1 &= \tfrac{1211268649}{440208640} \end{aligned} \tag{47}$$

which is the same system as (47). 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 (21)-(25) as :

$$M_I V > 0 \tag{48}$$
$$M_E V = d \tag{49}$$

where $V \equiv [Q^0, Q^1, \Phi^0, \Phi^1]$ is the vector of unknown distributional parameters from F_ϵ. Then substitute out a subvector of V in (48) using the equalities in (49). 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

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

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*.

8.3 Part C: More on the issue of "normalizations"

Suppose researchers choose some arbitrary distribution $\tilde{F}_{\epsilon|X=x}(\tilde{\varepsilon})$ for each x that is increasing in $\tilde{\varepsilon}$ in structural estimation, while the true underlying parameters are $\{c_0, F^0_{\epsilon|X}\}$. Then the cake function is recovered as

$$\tilde{c}(x, \tilde{\varepsilon}) = F^{*-1}_{Y|X=x}(\tilde{F}_{\epsilon|X=x}(\tilde{\varepsilon})) \tag{50}$$

It is straightforward to show that $\tilde{c}, c_0$ are related as

$$\tilde{c}(x, \tilde{\varepsilon}) = c_0(x, Q^0_{\epsilon|X=x}(\tilde{F}_{\epsilon|X=x}(\tilde{\varepsilon}))) \tag{51}$$

where $Q^0_{\epsilon|X}(\alpha)$ denotes the inverse of $F^0_{\epsilon|X}$ at α. Or alternatively,

$$\tilde{c}^{-1}(x, y) = \tilde{Q}_{\epsilon|X=x}(F^0_{\epsilon|X=x}(c_0^{-1}(x, y))) \tag{52}$$

for all x, y, where $c_0^{-1}, \tilde{c}^{-1}$ are inverses of $c_0, \tilde{c}$ at y for any given x , and $\tilde{Q}_{\epsilon|X=x}$ is the inverse of $\tilde{F}_{\epsilon|X=x}$. Suppose researchers are interested in knowing the distribution of cake sizes if the cake function is perturbed to $c_0^g(x,\varepsilon) = c_0(g(x),\varepsilon)$ for all x,ε. That is, for a given USV, the cake size under $X = x$ in the counterfactual environment would equal that in state $X = g(x)$ in the current data-generating process.

With normalization $\tilde{F}_{\epsilon|X}$ in place, the econometrician can first recover $\tilde{c}(x,\tilde{\varepsilon})$ from $F^*_{Y|X=x}$ as in (50), and then construct the counterfactual structural function of interest from $\tilde{c}$ as $\tilde{c}^g(x,\tilde{\varepsilon}) \equiv \tilde{c}(g(x),\tilde{\varepsilon})$. However, the true counterfactual distribution of cake sizes is $\Pr\{c_0(g(X),\epsilon) \leq y|\ X = x;\ F^0_{\epsilon|X}\} = F^0_{\epsilon|X=x}(c_0^{-1}(g(x),y))$, while the one predicted under the normalization is:

$$
\begin{aligned}
& \Pr\{\tilde{c}(g(X),\tilde{\epsilon}) \leq y|X = x; \tilde{F}_{\epsilon|X}\} = \tilde{F}_{\epsilon|X=x} \circ \tilde{c}^{-1}(g(x),y)) \\
= \ & \tilde{F}_{\epsilon|X=x} \circ \tilde{Q}_{\epsilon|X=g(x)} \circ F^0_{\epsilon|X=g(x)} \circ c_0^{-1}(g(x),y)
\end{aligned}
$$

where $f \circ g(.)$ is a shorthand for the composite function $f(g(.))$, and the second equality follows from (52). In general, $\tilde{F}_{\epsilon|X=x} \circ \tilde{Q}_{\epsilon|X=g(x)} \circ F^0_{\epsilon|X=g(x)} \neq F^0_{\epsilon|X=x}$, and hence the normalization $\tilde{F}_{\epsilon|X}$ may lead to errors in predicting the distribution of (X,Y) in the counterfactual context. In the special case where $F^0_{\epsilon|X}$ is known to be independent of X, choosing any $\tilde{F}_\epsilon$ (independent of X) indeed amounts to a normalization that is innocuous for the counterfactual exercise. This is obvious from the fact that with F^0_ϵ and $\tilde{F}_\epsilon$ both independent from X, $\tilde{F}_\epsilon(\tilde{Q}_\epsilon(F^0_\epsilon(\varepsilon))) = F^0_\epsilon(\varepsilon)$ holds trivially for all ε.

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

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

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

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

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

[16] 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

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

[18] Matzkin, R., "Nonparametric Estimation of Nonadditive Random Functions," Econometrica, September 2003, Vol. 71, No. 5, p. 1339-1375

[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] Rust, J., "Optimal Replacement of GMC Bus Engines: An Empirical Model of Harold Zurcher," Econometrica, September 1987, 55–5, 999-1033.

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

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

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

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

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