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 11-035

"Identification and Estimation of Stochastic Bargaining Models
Fourth Version"

by

Antonio Merlo and Xun Tang

http://ssrn.com/abstract=1946367

Identification and Estimation of Stochastic Bargaining Models*

Antonio Merlo† Xun Tang‡

Revised, October 2011

Abstract

Stochastic sequential bargaining models (Merlo and Wilson (1995, 1998)) have found wide applications in different fields including political economy and macroeconomics due to their flexibility in explaining delays in reaching an agreement. This paper presents new results in nonparametric identification and estimation of such models under different data scenarios.

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

JEL codes: C14, C35, C73, C78.

*We thank co-editor Jean-Marc Robin and three anonymous referees for their helpful suggestions. Steven Berry, Herman Bierens, Hanming Fang, Philip Haile, Isabelle Perrigne, Katja Seim, Elie Tamer, Petra Todd, Quang Vuong, Ken Wolpin and seminar and conference participants at several institutions provided useful comments. We thank Hulya Eraslan for sharing her data with us and Chamna Yoon for providing excellent research assistance. The usual disclaimer applies.

†Department of Economics, University of Pennsylvania, CEPR, CESifo and NBER, <merloa@econ.upenn.edu>.

‡Department of Economics, University of Pennsylvania, <xuntang@sas.upenn.edu>.

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 (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 (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), Ghosal, Miller and Thampanishvong (2010)).

The existing literature on the structural estimation of noncooperative bargaining models is 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.[1] 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 and estimation of stochastic bargaining models.[2]

Empirical contexts of stochastic bargaining games may differ in what econometricians observe in the 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 bargain over when to liquidate a portfolio they jointly own and how to divide the proceeds. The size of the cake is the market value of the portfolio which is determined by state variables, such as market or macroeconomic conditions, that evolve over time according to a stochastic process. 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 are not observed by the econometricians (USV). In the complete data scenario, the econometricians observe the evolution of the market value of

[1] 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.

[2] One of the main objectives of our analysis 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 bargaining game. 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, 2007), Haile and Tamer (2003), Haile, Hong and Shum (2004), Hendricks, Pinkse and Porter (2003). Chiappori and Donni (2006) also address related questions in the context of a static, cooperative (or axiomatic) bargaining framework and derive 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.

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 of complete data, we derive conditions for a joint distribution of observed states, surplus, agreements and divisions of the cake to be rationalized by a stochastic sequential bargaining model. 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 show it can be recovered under an appropriate normalization. For the case of incomplete data with censored cakes, we show that 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. For both data scenarios we provide consistent estimators of the discount factor and the cake function. Also, we illustrate our approach for estimation with an empirical application.

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 stochastic bargaining under the assumption that the USV distribution is known to the econometricians. Our contribution in this scenario is to relax the assumption of a known USV distribution, and show that partial identification of counterfactual outcomes (i.e. the probability of reaching an agreement conditional on observed states) is possible under nonparametric shape restrictions on 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.[3] 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.

[3]For example, the expected market value of a portfolio of foreign assets must be monotone in exchange rates holding other state variables fixed.

The rest of the paper is organized as follows. Section 2 introduces the canonical model of stochastic sequential bargaining and characterizes its equilibrium. It also describes the three data scenarios we consider and relates them to the existing empirical literature. Sections 3, 4 and 5 present our results on identification and estimation of the canonical model in the complete data, incomplete data with censored cakes and incomplete data with unobservable cakes scenarios, respectively. Section 6 concludes.

2 The Canonical Model of Stochastic Bargaining

Many real bargaining situations involve negotiations among two or more players over the allocation of some surplus. In many negotiations, the terms of an agreement may depend on aspects of the environment which change during the negotiating period. In such cases, the surplus to be allocated may evolve stochastically over time, and naturally lead to the possibility that agreement is delayed whenever the players perceive that a better agreement may be achieved by waiting.

To analyze these situations, Merlo and Wilson (1995, 1998) propose a general class of sequential bargaining games with complete information in which both the surplus to be allocated and the identity of the proposer follow a stochastic process. Here, we describe the prototypical stochastic sequential bargaining model which is commonly used in empirical applications (which we refer to as the "canonical model").

Consider an infinite-horizon bargaining game with $K \geq 2$ players (denoted by $i = 1, ..., K$) who share the same discount factor $\beta \in (0,1)$. In each period $t = 0, 1, ...$, all players observe a vector of states (X_t, ϵ_t) with support $\Omega_{X,\epsilon}$. (Throughout the paper, 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. $R^t \equiv \{R_0, R_1, ., R_t\}$.) The realized state (x_t, ε_t) determines the set of feasible utility vectors to be allocated in period t, $\mathcal{C}\left(x_t, \varepsilon_t\right) \equiv \{u \in \mathbb{R}_+^K : \sum_{i=1}^K u_i \leq c(x_t, \varepsilon_t)\}$, where $c : \Omega_{X,\epsilon} \rightarrow \mathbb{R}_+^1$ determines the total surplus to be agreed upon in that period, or the size of the "cake".[4] In each period, player i is randomly selected to be the proposer with probability $L_i \in (0,1)$, $\sum_{i=1}^K L_i = 1$, $L \equiv \{L_1, ..., L_K\}$. We denote the (random) identity of the proposer in any given period t by κ_t and its realization by $\Bbbk_t$.[5]

[4] 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.

[5] The proposer identity κ_t may be allowed to depend on the current states X_t. As long as, once we condition on X_t, κ_t is independent of ϵ_t and of past and future states, all our results generalize. We present

We assume that unlike the players in the game, researchers can only observe the vector of states X_t, but not the scalar noise ϵ_t. For generic random vectors R_1, R_2, let $F(R_1|R_2)$ or $F_{R_1|R_2}$ denote the distribution of R_1 conditional on R_2. Throughout the paper, we maintain the following restriction on the transition of states.

CI *(Conditional independence) The transition between states satisfies:*

$$F(X_{t+1}, \epsilon_{t+1}|X^t, \epsilon^t) = F(\epsilon_{t+1}|X_{t+1})G(X_{t+1}|X_t). \tag{1}$$

For the rest of the paper, we use (R, R') to denote random vectors in the current and the next period, respectively. Assumption *CI* requires that the dynamics between current and subsequent states (X, ϵ) and (X', ϵ') be completely captured by the stochastic process governing the transition of states that are observable to the researchers. This is a condition commonly shared by a wide range of structural dynamic models in industrial organization and labor economics (e.g. Rust (1987)).

The bargaining game proceeds as follows. At the beginning of each period, players observe (x, ε) and know the identity of the proposer $\Bbbk$. The proposer then chooses to either propose an allocation in $\mathcal{C}(x, \varepsilon)$ or pass and let the game move to the next period. If he proposes an allocation, the other players respond by either accepting or rejecting the proposal. If no proposal is offered or the proposal is rejected by some player, the game moves to the next period where new states (x', ε') are realized and a new proposer $\Bbbk'$ is selected with probabilities L. The procedure is then repeated with total surplus given by $c(x', \varepsilon')$. This game continues until an allocation is proposed and unanimously accepted (if ever).

The structural parameters $(\beta, c, L, G, F_{\epsilon|X})$ are common knowledge among all players (although they are not known to the researchers). Let $S_t \equiv (X_t, \epsilon_t, \kappa_t)$ denote the information available to the players at time t. Given any initial realized information s_0, an outcome of the bargaining game (τ, η_τ) consists of a stopping time τ and a random K-vector η_τ (measurable with respect to S^τ) such that $\eta_\tau \in \mathcal{C}(X_\tau, \epsilon_\tau)$ is a feasible division of the cake in state (X_τ, ϵ_τ) if $\tau < +\infty$, and $\eta_\tau = 0$ if $\tau = +\infty$. Given a realized sequence of information $(s_0, s_1, ..)$, τ is the period in which a proposal is accepted by all players, and η_τ is the accepted proposal when the state is (X_τ, ϵ_τ) and the identity of the proposer is κ_τ. For a game starting with initial states (x, ε) and proposer $\Bbbk$, an outcome (τ, η_τ) implies a von Neumann-Morgenstern payoff to player i, $E[\beta^\tau \eta_{\tau,i}|S_0 = (x, \varepsilon, \Bbbk)] = E[\beta^\tau \eta_{\tau,i}|X_0 = x, \epsilon_0 = \varepsilon, \kappa_0 = \Bbbk]$, where $\eta_{\tau,i}$ is the i-th coordinate of η_τ.

the case where κ_t does not depend on X_t to simplify the exposition.

A stationary outcome is such that there exists a measurable subset ω_S of the support of S, and a measurable function $\mu : \omega_S \rightarrow \mathbb{R}^K_+$ such that (i) $S_t \notin \omega_S$ for all $t = 0, 1, ., \tau - 1$; (ii) $S_\tau \in \omega_S$; and (iii) $\eta_\tau = \mu(S_\tau)$. That is, no allocation is implemented until some state and proposer identity $s = (x, \varepsilon, \Bbbk) \in \omega_S$ is realized, in which case a feasible proposal $\mu(x, \varepsilon, \Bbbk)$ is accepted by all players. Given property (iii), we let $v^\mu(s) \equiv E[\beta^\tau \mu(S_\tau)|S_0 = s]$ denote the vector of individual von Neumann-Morgenstern payoffs given the initial state and proposer identity in s. It follows from the definition of stationary outcome that $v^\mu(s) = \mu(s)$ for all $s \in \omega_S$ and $v^\mu(s) = E[\beta^\tau \mu(S_\tau)|S_0 = s]$ for all $s \notin \omega_S$. Hence, a stationary outcome is characterized by the triplet (ω_S, μ, τ).

A history up to time t is a finite sequence of realized states, identities of proposers, and actions taken up to time 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 payoffs induced by a subgame perfect strategy profile as a SPE outcome and SPE payoffs respectively. A strategy profile is stationary if the actions prescribed at any history depend only on the current state, proposer identity, and offer. A stationary SPE (SSPE) outcome and payoffs are the outcome and payoffs generated by a subgame perfect strategy profile which is stationary.

The following lemma characterizes the players' actions and outcomes in the unique SSPE of the game. Let $v_i : \Omega_S \rightarrow \mathbb{R}^1_+$ denote the SSPE payoff for player i, $v = (v_1, ..., v_K)$ the SSPE payoff vector, and $w = \sum_{i=1}^K v_i$ the SSPE total payoff of all players in the bargaining game. Let F^K denote the set of bounded measurable functions mapping from Ω_S to $\mathbb{R}^K$.

Lemma 1 *(Characterization of SSPE) Suppose CI holds. Then: (a) $v \in F^K$ is the unique SSPE payoff vector if and only if $v = A(v)$, where for any s and $i, j = 1, ..., K$, the operator A is defined as:*

$$A_i(v(s)) \equiv \max\left\{c(x,\varepsilon) - \beta E\left[\textstyle\sum_{j\neq i} v_j(S')|x\right], \beta E\left[v_i(S')|x\right]\right\} \text{ for } i = \Bbbk, \tag{2}$$

$$A_j(v(s)) \equiv \beta E[v_j(S')|x] \text{ for } j \neq \Bbbk; \tag{3}$$

(b) the SSPE total payoff w does not depend on the identity of the proposer κ, and solves

$$w(x,\varepsilon) = \max\{c(x,\varepsilon), \beta E[w(X',\epsilon')|x]\}; \tag{4}$$

and (c) an agreement is reached in states (x,ε) if and only if $c(x,\varepsilon) \geq \beta E[w(X',\epsilon')|x]$.

The proof of Lemma 1 is based on results in Theorems 1-3 in Merlo and Wilson (1998). A striking feature of the equilibrium of the canonical model of stochastic bargaining is that the SSPE total payoff and the occurrence of agreement in equilibrium only depend on the current states (X, ϵ), but not on the identity of the proposer κ. This important property of SSPE, known as the "separation principle", is instrumental for some of our identification strategies below. In contrast, the individual SSPE payoffs $(v_i)_{i=1}^K$ do depend on the identity of the proposer. In particular, only $i = \Bbbk$ can claim the additional "gains to the proposer" $c(x,\varepsilon) - \beta E[\sum_{j=1}^K v_j(S')|x]$ in addition to his own continuation payoff $\beta E[v_i(S')|x]$, while all other players just get their individual continuation payoffs.

The operator A we use to characterize the SSPE payoffs is based on the fundamental observation that, if an agreement is reached in any period, the proposer may extract any surplus over what the players obtain by delaying agreement until the next period (i.e. their equilibrium continuation payoffs). The separation principle implies that the identity of the proposer affects how the cake is allocated, but not the states in which it is allocated. Furthermore, the gains from proposing are also independent of the identity of the proposer and the characterization of the SSPE total payoff is equivalent to the solution of the single agent problem of deciding when to consume a stochastic cake. It follows that the SSPE total payoff maximizes the expected discounted size of the cake (i.e. the expected discounted total surplus allocated among the players). The unique SSPE of the game (and hence any delay in agreement) is therefore Pareto efficient. The fact that a temporary delay in agreement is a possible equilibrium outcome follows from the possibility that the discounted size of the cake need not decline in every period. In particular, equilibrium delays occur in states where the cake is "too small": that is, the sum of the continuation payoffs of all players (including the proposer) exceeds the current size of the cake.

Econometricians are interested in recovering the underlying structure of the model, summarized by the parameters $(\beta, c, L, F_{\epsilon|X}, G)$ under assumption *CI*, using the distributions of states, actions, proposers and allocations observed in the data from a large sample of bargaining games. For each bargaining game in the data, we assume that researchers observe the time to agreement and the history of the states X, but not ϵ. Also, we posit that all the bargaining games observed in the data share the same transition of states and the same cake function $c : \Omega_{X,\epsilon} \to \mathbb{R}^1_+$, and all players follow SSPE strategies. In the next three sections of the paper, we discuss the (nonparametric) identification and estimation of the structure under different scenarios where the total surplus, the identity of the proposers and the agreed proposals may or may not be reported in the data.

We consider three scenarios with increasing data limitations. In the first scenario, researchers have access to "complete data" and observe the identity of the proposer and the size of the surplus in each period regardless of whether an agreement is reached. This represents a useful benchmark to establish the extent to which model structures can be nonparametrically identified and estimated under ideal circumstances. While not common, this scenario is empirically relevant given the increased availability of bargaining data. For example, in the context of negotiations over residential real estate transactions in England, Merlo, Ortalo-Magne and Rust (2009) observe the entire history of offers made by each potential buyer, including the sequence of rejected offers within individual negotiations.

In the second scenario, researchers have access to "incomplete data with censored cakes" and observe the size of the cake only in the period when an agreement is reached. Also, the identity of the proposer in any given period may or may not be observable. This is the most common data scenario in the existing empirical literature. For example, in his empirical analysis of the process of government formation in Italy, Merlo (1997) specifies the size of the cake to be the expected duration of the government, which depends on the state of the economy (X) as well as the political climate (ϵ) while negotiations take place. In addition to observing the duration of the negotiation, the sequence of proposers (or *formateurs*), and the time-series of several macroeconomic variables during the negotiation for a sample of 47 bargaining episodes between 1947 and 1994, the data contain the durations of the coalition governments that actually formed, but not the expected duration of proposed governments that were rejected (see also Diermeier, Eraslan and Merlo (2003)). Similarly, Benjamin and Wright (2008) and Ghosal, Miller and Thampanishvong (2010) use a stochastic bargaining model to explain the length of delays in a sample of 90 sovereign debt restructurings during 1998-2005 and only observe the size of the "haircuts" that were agreed upon at the conclusion of each negotiation. This data scenario also applies to the empirical study of corporate bankruptcy reorganizations by Eraslan (2008). Since we use this application to illustrate our methodology, we discuss it in more details in Section 4.3 below.

In the third scenario, researchers have access to "incomplete data with unobservable cakes" and only observe the timing of agreement, but never observe the size of the cake or the identity of the proposer. Since this scenario only requires minimal data, there are several possible applications where large data sets exist that only contain information about the duration of negotiations. For example, in his empirical study of the time to the adoption of standards by the Internet Engineering Task Force, which he models as a stochastic bargaining problem, Simcoe (2008) uses a dataset containing the time between initial submission and final revision of 2,601 Internet Protocols between 1993 and 2003. Similarly, Diaz-Moreno and Galdon (2000) estimate a stochastic model of collective bargaining using data on the duration of 545 collective negotiations in Spain in the late 1980s.

In practice, data may contain 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 all $i = 1, ..., K$. Such profiles of individual characteristics may vary across bargaining episodes in the data, but remain the same throughout any given negotiation. Of course, primitives $(\beta, c, L, G, F_{\epsilon|X})$ may also depend on (K, Z^K). Since these individual characteristics are observable in the data and fixed over time, our identification arguments throughout the paper should be interpreted as conditional on (K, Z^K). We suppress dependence of the structural elements on the vector (K, Z^K) only for the sake of notational simplicity.

3 Complete Data

In this section, we discuss the empirical content of the canonical stochastic bargaining model when researchers observe the complete history of *(i)* X_t and cake sizes $Y_t = c(X_t, \epsilon_t)$, but not ϵ_t; *(ii)* whether an agreement is reached at time t (denoted by a dummy variable D_t); and *(iii)* the proposer identity κ_t. Researchers also observe the agreed allocations (i.e. $\eta_\tau \in \mathcal{C}(X_\tau, \epsilon_\tau)$, where τ denotes the period when agreement occurs), but may not observe rejected proposals (if any) in any other period. As we already pointed out, this observational environment is both theoretically interesting and empirically relevant. We first discuss testable restrictions under our conditional independence assumption (CI) as well as monotonicity of the cake function. Then, we show constructive identification of the discount factor and the cake function and propose consistent nonparametric estimators for both objects.

3.1 Preliminaries

Let F_{X_0} denote the distribution of the initial observable state X_0 at the beginning of the bargaining game. For the rest of the paper, we maintain a regularity condition that the support of the observable states is the same for all periods (i.e. for all x and t, $\Pr(X_{t+1} \in \omega_X \mid X_t = x) > 0$ for all $\omega_X \subseteq \Omega_X$ such that $\Pr(X_0 \in \omega_X) > 0$). Under assumption *CI*, only $\theta \equiv \{\beta, c, F_{\epsilon|X}\}$ remain to be identified, while the proposer-selection mechanism L, the transition of observed states G and the distribution of initial states F_{X_0} can be recovered from the data. Hence, for identification of θ, we can treat L and G as fixed and known.

We say a joint distribution of the time to agreement τ, the accepted allocations η_τ and the history $(X^\tau, Y^\tau, \kappa^\tau)$ is rationalized by some θ in a bargaining game (whose structure satisfies CI) if this distribution arises in a SSPE given θ. Define a feature $\Gamma(.)$ as a mapping from the space of parameters in θ into that of the features (e.g. $\Gamma(\theta)$ could be a subvector or some functional of θ, such as the location (median) or the scale (variance) of ϵ given X).

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

Any feature of the true parameters $\Gamma(\theta_0)$ that can be expressed in terms of observed distributions of $\{\tau, \eta_\tau, X^\tau, Y^\tau, \kappa^\tau\}$ is identified. Note that identification is defined under *CI* and any additional restrictions on θ as captured in Θ. The role of *CI* in defining identification is to reduce the parameter space of interest to that of $\theta \equiv \{\beta, c, F_{\epsilon|X}\}$. For any $\theta \in \Theta$, let

$$\pi_i(X;\theta) \equiv E[v_i(S';\theta)|X] \;\; \text{for } i = 1, ..., K \;\; \text{and} \;\; \pi_w(X;\theta) \equiv E[w(X', \epsilon'; \theta)|X] \tag{5}$$

denote respectively individual and total continuation payoffs under θ in a SSPE, where v_i and w are respectively the SSPE individual and total payoffs given by (2)-(4) in Lemma 1. We maintain two additional restrictions on the parameter space.

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

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

The monotonicity restriction *MT* ensures there exists a one-to-one mapping between cake sizes and unobserved states given any x.[6] The non-degeneracy condition *ND* requires that for any x, there is enough variation in unobserved states so that an agreement may or may not occur with positive probability. In other words, the discounted total continuation payoff is in the interior of the support of cake sizes for all observed states. This condition helps to rule out uninteresting cases where agreement or non-agreement become a certainty once X reaches some realized value x. Since each player has a positive probability of being the proposer in any state, this assumption also implies that for each player i and state x there is always a positive probability that an agreement is reached on someone else's proposal.[7]

[6] This assumption is fairly restrictive and whether it is likely to hold or not depends on the structural interpretation of ϵ in the specific context of each application.

[7] The restriction in *ND* can be tested using the distribution of states, cakes and decisions.

3.2 Testable restrictions

We derive necessary and sufficient conditions for the joint distribution of $\{\tau, \eta_\tau, X^\tau, Y^\tau, \kappa^\tau\}$ to be rationalized under *CI* by some θ satisfying *MT* and *ND*. Let Θ_{MN} denote the set of values of θ that satisfy *MT* and *ND*. Let $p(x) \equiv \Pr\{D = 1|X = x\}$ and $L_i \equiv \Pr\{\kappa = i\}$ denote observable probabilities. For all $\alpha \in (0,1)$, let $Q_{Y|x}(\alpha) \equiv F^{-1}_{Y|x}(\alpha)$ denote the conditional quantiles of cake sizes. Define $\lambda^*(x) \equiv Q_{Y|x}(1 - p(x))$.

Lemma 2 *(Testable Restrictions) A joint distribution of $\{\tau, \eta_\tau, X^\tau, Y^\tau, \kappa^\tau\}$ is rationalized under CI by some $\theta \in \Theta_{MN}$ if and only if: (i) for all $t \geq 0$ and $(x^t, y^t, D^t, \mathbb{k}^t)$,*

$$F\left(Y_{t+1}, D_{t+1}, \kappa_{t+1}, X_{t+1}|t < \tau, y^t, \mathbb{k}^t, x^t\right) = L_{\kappa_{t+1}} F(Y_{t+1}, D_{t+1}|X_{t+1}) G(X_{t+1}|x_t) \qquad (6)$$

and $F(Y_0, D_0, \kappa_0|x_0) = L_{\kappa_0} F(Y_0, D_0|x_0)$ for all x_0; (ii) for all x, $F_{Y|X=x}$ is strictly increasing and $p(x) \in (0,1)$; (iii) $Y_t < \lambda^*(X_t)$ for all $t < \tau$ and $Y_\tau \geq \lambda^*(X_\tau)$; (iv) there exists $\alpha \in (0,1)$ such that for all x,

$$\left(\int \max\{y', \lambda^*(x')\} dF_{Y',X'|x}\right)^{-1} \lambda^*(x) = \alpha; \qquad (7)$$

and (v) there exist K functions $(\lambda_i^*)_{i=1}^K$ such that $\eta_{\tau,i} = \lambda_i^*(X_\tau)$ for $i \neq \kappa_\tau$, $\eta_{\tau,i} = Y_\tau - \sum_{j \neq i} \lambda_j^*(X_\tau)$ for $i = \kappa_\tau$, and for all i, x,

$$\lambda_i^*(x) = \alpha \int \lambda_i^*(x') + L_i \int \max\{y' - \lambda^*(x'), 0\} dF_{Y'|X'} dG_{X'|x}. \qquad (8)$$

The conditions in Lemma 2 arise naturally from properties of the SSPE and from the assumptions *CI*, *MT* and *ND*. Condition (6) says that the distribution of cake sizes and the probability of reaching an agreement do not depend on history once we condition on contemporary observed states. They are also independent of the identity of the proposer in any given period. These are direct implications of our conditional independence assumption (given current observed states, the history of unobserved and past observed states are not informative about the future), and the fact that the proposer-selection mechanism is independent of the evolution of states. The conditional distribution of cake sizes is increasing, as stated in condition *(ii)*, because the cake function is monotone in the unobserved state, whose conditional distribution is also increasing.

The separation principle implies that the occurrence of agreement in any given period only depends on whether the cake size exceeds the discounted total continuation payoff in that period. Under *CI*, the latter is a function of current observed states alone. Under *ND*, it must also lie in the interior of the conditional support of cake sizes. Therefore, condition *(iii)* in Lemma 2 says that for any given observed state there is a single "threshold cake size" beyond which agreement occurs. Hence, the threshold for any state can be recovered as the appropriate conditional quantile of cake sizes. Condition *(iv)* also builds on a similar argument to link the common discount factor to the distribution of observables. We elaborate more on the intuition for this result following Proposition 1 below.

Finally, in any SSPE, a player who is not the proposer always accepts an allocation that gives him his discounted individual continuation payoff. Assumption *ND* implies that for each player i and state x there is a positive chance that an agreement occurs when i is not the proposer in state x. Condition *(v)* simply relates each individual's discounted continuation payoff to the allocation he accepts when someone else proposes in state x.[8]

3.3 Identification

We discuss identification of $\{\beta, c, F_{\epsilon|X}\}$ when the joint distribution of $\{\tau, \eta_\tau, X^\tau, Y^\tau, \kappa^\tau\}$ observed from the data is rationalizable (i.e. it satisfies the conditions in Lemma 2). Let $F_{\tau,\eta_\tau,X^\tau,Y^\tau,\kappa^\tau}$ denote the rationalizable joint distribution observed from the data. To clarify the exposition in the rest of this section, let $\theta_0 \equiv \{\beta_0, c_0, F^0_{\epsilon|X}\}$ denote the true parameters in the data-generating process and $\theta \equiv \{\beta, c, F_{\epsilon|X}\}$ a generic set of parameters. Let $c^{-1}(x, y)$ denote the inverse function of $c(x, \varepsilon)$ at y given x. Recall that the proposer-selection mechanism L is independent from other states and is directly identified from observables.

Proposition 1 *Suppose CI, MT and ND hold. (i) The discount factor is identified as:*

$$\beta_0 = \left(\int \int \max\{y', \lambda^*(x')\} dF_{Y'|X'} dG_{X'|x} \right)^{-1} \lambda^*(x). \tag{9}$$

(ii) A pair $(c, F_{\epsilon|X}) \overset{o.e.}{\sim} (c_0, F^0_{\epsilon|X})$ *if and only if* $F_{\epsilon|x}(c^{-1}(x, y)) = F_{Y|x}(y)$ *for all* x*. (iii) If in addition* ϵ *is independent of* X*, then* c_0 *is identified as*

$$c_0(x, \varepsilon) = Q_{Y|x}(F^0_\epsilon(\varepsilon)) \tag{10}$$

with F^0_ϵ *normalized to a known distribution.*

Identification of β_0 builds on two intuitive observations. First, under *CI*, ϵ is not informative about next period's total SSPE payoff given X. Hence, the true total continuation payoff

[8] Lemma 2 summarizes all the restrictions imposed by the model on the distribution of observables in the complete data scenario. These restrictions can therefore be used to construct a test for the null that the data is rationalized by the canonical stochastic sequential bargaining model. We leave the development of such a test for future research.

$\pi_w(x;\theta_0)$ does not depend on ε. Under *ND*, $\beta_0\pi_w(x;\theta_0)$ lies in the interior of the support of cake sizes, and *MT* implies it can be directly recovered as a conditional quantile $\lambda^*(x)$ as in Lemma 2. Second, changing variables between ϵ and Y under *MT* helps to relate the discounted total continuation payoff to observed distributions through a "quasi-structural" fixed-point equation:

$$\beta_0\pi_w(x;\theta_0) = \beta_0 \int\int \max\{y', \beta_0\pi_w(x';\theta_0)\} dF_{Y'|X'} dG_{X'|x} \tag{11}$$

where the prefix "quasi-" is due to the fact that $(c_0, F^0_{\epsilon|X})$ only enters through the observed distribution of cake sizes $F_{Y|X}$ it implies. Substituting $\lambda^*(x)$ in place of $\beta_0\pi_w(x;\theta_0)$ in the quasi-structural form gives (9).

Part *(ii)* of Proposition 1 states that a generic pair $(c, F_{\epsilon|X})$ is observationally equivalent to the true parameters if and only if it implies the same cake distribution $F_{Y|X}$ as observed in the data-generating process. This is fairly intuitive because the separation principle implies that agreements do not depend on the identity of the proposer, and proposer identities evolve independently of the other state variables. It follows from *(ii)* that the cake function and the distribution of the unobserved state are non-identified because they cannot be jointly identified from the observed cake distribution $F_{Y|X}$ alone.

When ϵ is independent of X, following Matzkin (2003) we could have normalized $c_0(\bar{x},\varepsilon) = \varepsilon$ for some $\bar{x}$ and recovered $F^0_\epsilon(t)$ and $c_0(x,\varepsilon)$ as $F_{Y|\bar{x}}(t)$ and $Q_{Y|x}(F_{Y|\bar{x}}(\varepsilon))$, respectively. Instead, we propose an alternative normalization in *(iii)* of Proposition 1 that sets F^0_ϵ to a known distribution (such as a uniform on $[0,1]$). This allows us to recover $c_0(x,\varepsilon)$ as $Q_{Y|x}(F^0_\epsilon(\varepsilon))$. On the other hand, if the distribution of the unobserved state depends on X, assuming a specific form for $F^0_{\epsilon|X}$ in order to recover $c_0(x,\varepsilon)$ would not be an innocuous "normalization" in general because the chosen form of $F^0_{\epsilon|X}$ can affect predictions in some counterfactual analyses (see Appendix B for details).

3.4 Consistent estimation

We construct a multi-step nonparametric estimator for the discount factor β_0 by plugging in sample analogs in the identification arguments presented above. We show consistency of the estimator when the support of X is finite.[9] The data contains N independent bargaining games, each indexed by g. For a game g, let T_g denote the number of bargaining periods (which are indexed by t) observed in the data. It is possible that, for some games, players may

[9] We can generalize our estimator to allow for mixed discrete and continuous covariates by using kernel smoothing. This generalization is however omitted to economize on space.

not reach any agreement over the time periods observed in the data. To simplify notation, let $\sum_{g,t}(.) \equiv \sum_{g=1}^{N}\sum_{t=1}^{T_g}(.)$. The estimation procedure is as follows. First, estimate the probabilities of agreement and the conditional distribution of cake sizes as:

$$\hat{p}(x) \equiv \frac{\sum_{g,t} 1\{D_{g,t} = 1 \;\wedge\; X_{g,t} = x\}}{\sum_{g,t} 1\{X_{g,t} = x\}} \tag{12}$$

$$\hat{F}_{Y|X}(r|x) \equiv \frac{\sum_{g,t} 1\{Y_{g,t} \leq r \;\wedge\; X_{g,t} = x\}}{\sum_{g,t} 1\{X_{g,t} = x\}}. \tag{13}$$

Next, estimate the discounted ex ante total continuation payoffs $\beta_0 \pi_w(x;\theta_0)$ as:

$$\hat{\lambda}(x) \equiv \arg\min_{r \in \mathbb{R}^1} \left[\hat{F}_{Y|X}(r|x) - \hat{p}(x)\right]^2. \tag{14}$$

Third, estimate the conditional expectation of total payoffs $m_0(x) \equiv E[\max\{Y \;,\; \beta_0 \pi_w(X;\theta_0)\} \mid X = x]$ as

$$\hat{m}(x) \equiv \frac{\sum_{g,t} \max\{Y_{g,t} \;,\; \hat{\lambda}(x)\} 1\{X_{g,t} = x\}}{\sum_{g,t} 1\{X_{g,t} = x\}}. \tag{15}$$

Fourth, estimate the transition of observed states $G_{X'|X}$ as $\hat{G}_{x'|x} \equiv \left(\Sigma_{\bar{x} \in \Omega_X} n_{\bar{x}|x}\right)^{-1} n_{x'|x}$, where $n_{x'|x}$ denotes the number of transitions from state x to x' observed in the data. Let $n_x \equiv \sum_{g,t} 1\{X_{g,t} = x\}$ and $w_x \equiv \left(\sum_{\bar{x} \in \Omega_X} n_{\bar{x}}\right)^{-1} n_x$. Then, our estimator for the discount factor is:

$$\hat{\beta} \equiv \sum_{x \in \Omega_X} w_x \left[\frac{\hat{\lambda}(x)}{\sum_{x' \in \Omega_X} \hat{G}_{x'|x} \hat{m}(x')}\right]. \tag{16}$$

Let $\xrightarrow{p}$ denote convergence in probability and for a generic vector z, let $\mathcal{N}_e(z)$ be an e-neighborhood around z.

Proposition 2 *Suppose CI, MT and ND hold; $E[c_0(X,\epsilon)|X = x] > 0$ for all x; and there exists $e > 0$ such that $\sup_{\zeta \in \mathcal{N}_e(\lambda^*(x;\theta_0))} |E[\max\{c_0(X,\epsilon),\zeta\}|X = x]| < \infty$. Then $\hat{\beta} \xrightarrow{p} \beta_0$.*

To estimate c_0, we can normalize F_ϵ^0 to $N(0,1)$ and estimate $c_0(x,\varepsilon)$ by $\hat{c}(x,\varepsilon) \equiv \hat{Q}_{Y|X=x}(\Phi(\varepsilon))$ where Φ is the standard normal distribution. The consistency of $\hat{c}$ holds under standard regularity conditions for nonparametric quantile estimators.

4 Incomplete Data with Censored Cakes

In this section, we discuss nonparametric identification and estimation of the canonical stochastic bargaining model when the cake size is only observed in the event of an agreement.

As we pointed out above this is the predominant data scenario in empirical applications. In this scenario, conditional on any observed state, the distribution of cake sizes is censored at the discounted total continuation payoff. In addition, researchers may or may not observe the identity of the proposer in any given period. We show that the common discount factor can still be recovered from the distribution of observables under our conditional independence assumption (*CI*). Furthermore, we show that if the total surplus is additively separable in observed and unobserved states, and if the distribution of the unobserved state satisfies some exclusion restrictions, then the cake function can be identified. We construct consistent nonparametric estimators for the discount factor and the cake function and illustrate our methodology by performing an actual estimation of a simple stochastic bargaining model of corporate bankruptcy reorganization. In what follows, we drop subscripts and superscripts 0 from the true parameters in the data-generating process to simplify notation.

4.1 Identification

For generic random vectors R_1, R_2, let $Q^{\alpha}_{R_2|R_1}$ denote the α-quantile of R_2 conditional on R_1. Throughout this section, we maintain the following restrictions on the model structure.

AS *(Additive separability)* $c(x,\varepsilon) = c(x) - \varepsilon$ *for all* x, ε, *where* $c(x)$ *is bounded over* Ω_X.

RS *(Rich support)* ϵ *has positive densities with respect to the Lebesgue measure over* $\mathbb{R}^1$.

MI *(Median independence)* $Q^{0.5}_{\epsilon|x} = 0$ *for all* $x \in \Omega_X$.

ER *(Exclusion restrictions)* $X = (X^a, X^b)$ *and* ϵ *is independent of* X^b *conditional on* X^a.

SV *(Sufficient variation)* $\Pr(X \in \Omega^+_X | x^a) > 0$ *for all* x^a, *where* $\Omega^+_X \equiv \{x : c(x) \geq \beta\pi_w(x)\}$.

Under additive separability of the cake function (*AS*), the equilibrium characterization in Lemma 1 implies that $c(X) - \beta\pi_w(X) - \epsilon$ is the difference between the proposer's contemporary payoff should he decide to offer the other players their continuation values and his own continuation value. Hereinafter, we refer to $c(X) - \beta\pi_w(X) - \epsilon$ as the "potential proposer gains" and $\max\{c(X) - \beta\pi_w(X) - \epsilon, 0\}$ as the "actual proposer gains" that accrue to the proposer when agreements occur. Under median independence of the unobserved state (*MI*), $c(x) - \beta\pi_w(x)$ is the median of the potential proposer gains given x. The rich support assumption (*RS*) requires the unobserved state ϵ to be distributed with a large support over $\mathbb{R}^1$. This requirement guarantees that there is a positive probability that the potential proposer gains are strictly greater than zero in any observed state. Hence, in spite of censoring, at least some high quantile of cake sizes is observable in each state.[10]

[10]The restriction that the support of ϵ is unbounded in $\mathbb{R}^1$ is stronger than necessary. Identification only

The exclusion restrictions (ER) imply that a subvector of observed states X^b is independent of the noise in the potential proposer gains ϵ conditional on the remaining observed states X^a. These restrictions may arise naturally if the noise unobserved by researchers affects not only the total surplus but also a subset of the observed state variables. The sufficient variation condition (SV) ensures that for any x^a there is always enough variation in the excluded states X^b so that (x^a, X^b) leads to strictly positive proposer gains, and hence agreements, with over 50% chance. We discuss the practical implications of assumptions ER and SV in the context of an empirical application in Section 4.3 below.

Proposition 3 establishes identification of the discount factor and the cake function.[11] For any bounded function b defined over Ω_X, let $H \circ b(X) \equiv \int \max\{c(X') - \epsilon', b(X')\} dF_{\epsilon'|X'} dG_{X'|X}$. For any x, let $\tilde{\lambda}(x)$ denote the lower bound of the support of cake sizes Y under agreement ($D = 1$). Note that even when the observed distribution of cake sizes is censored, the conditional median of cake sizes is identified for all observed states where the probability of agreement is greater than one half. In particular, for any such state x, if we define Y^c to be equal to the observed cake size whenever an agreement occurs and to $\tilde{\lambda}(x)$ when it does not (i.e. $Y^c \equiv DY + (1 - D)\,\tilde{\lambda}(x)$), then $Q^{0.5}_{Y^c|x} = Q^{0.5}_{Y|x}$.

Proposition 3 *Suppose CI, AS, RS, MI, ER and SV hold, and H is a self-map over the space of bounded functions on Ω_X. Then: (i) β is identified as*

$$\beta = \left(\int \int D'Y' + (1 - D')\tilde{\lambda}(X') dF_{Y',D'|X'} dG_{X'|x} \right)^{-1} \tilde{\lambda}(x) \tag{17}$$

for all $x \in \Omega_X$; (ii) $x \in \Omega_X^+$ if and only if $p(x) \geq \frac{1}{2}$; (iii) for any $x \in \Omega_X^+$, $c(x)$ is identified as:

$$c(x) = Q^{0.5}_{Y|x};$$

and (iv) for any $x = (x^a, x^b) \notin \Omega_X^+$, $c(x)$ is identified as:

$$c(x) = Q^{\alpha}_{Y|D=1,x} + Q^{0.5}_{Y|\tilde{x}} - Q^{\alpha'}_{Y|D=1,\tilde{x}}$$

for any $\alpha \in (0, 1)$ and $\tilde{x} \equiv (x^a, \tilde{x}^b) \in \Omega_X^+$, where $\alpha' = 1 - (1 - \alpha)\frac{p(x)}{p(\tilde{x})}$.

---

requires that conditional on all $x \in \Omega_X$ the support of ϵ is large enough to induce positive gains to the proposer (and therefore an agreement) with positive probability. However, imposing the stronger condition simplifies the exposition.

[11] Our identification arguments extend immediately to cases where condition ER holds after further conditioning on observable instruments X^z, which may or may not enter the cake function.

Proof of Proposition 3. Proof of (i) and (ii). Because H is a self-map, the total continuation payoffs and the median potential proposer gains are bounded. Then AS and RS imply $p(x) \equiv \Pr\{\epsilon \leq c(x) - \beta\pi_w(x)|x\} \in (0,1)$ for all x. The discounted total continuation payoffs $\beta\pi_w(x)$ are then identified as $\tilde{\lambda}(x)$ for any x. By Lemma 1,

$$\int \max\{Y, \beta\pi_w(X)\} dF_{Y|X} = \int DY + (1-D)\beta\pi_w(X) dF_{Y,D|X}$$

It then follows the discount factor is (over-)identified as in part *(i)* for all $x \in \Omega_X$. Because Lemma 1 suggests $\Pr\{D = 1|x\} = \Pr\{\epsilon \leq c(x) - \beta\pi_w(x)|x\}$, part *(ii)* follows from *RS*.

Proof of (iii). For any $x \in \Omega_X$ and $\alpha \in (0,1)$, define $\gamma_x(\alpha) \equiv 1 - p(x) + \alpha p(x)$. By the Law of Total Probability, $Q^{\alpha}_{Y|D=1,x} = Q^{\gamma_x(\alpha)}_{Y|x}$ for all (x, α). For any $x \in \Omega^+_X$, invert γ_x at $\frac{1}{2}$ to get $\gamma_x^{-1}(0.5) = 1 - 0.5/p(x)$. By construction, $0 < 1 - 0.5/p(x) < \frac{1}{2}$ and

$$Q^{1-0.5/p(x)}_{Y|D=1,x} = Q^{0.5}_{Y|x} = c(x) \tag{18}$$

for all x in Ω^+_X, where the second equality follows from AS and *MI*.

Proof of (iv). Recall that by construction, $\gamma_x(\alpha) > 1 - p(x)$ for all $\alpha \in (0,1)$ and

$$Q^{\alpha}_{Y|D=1,x} = Q^{\gamma_x(\alpha)}_{Y|x} = c(x) - Q^{1-\gamma_x(\alpha)}_{\epsilon|x} \tag{19}$$

for all $x \in \Omega_X$ (including $x \notin \Omega^+_X$). Consider any $x \notin \Omega^+_X$. Because $Q^{\alpha}_{Y|D=1,x}$ is directly identified, $c(x)$ will be identified if $Q^{1-\gamma_x(\alpha)}_{\epsilon|x}$ can be recovered from observed distributions. Suppose x consists of (x^a, x^b). Condition SV implies for all x^a that $\tilde{x} \equiv (x^a, \tilde{x}^b) \in \Omega^+_X$ must exist with positive probability. And ER implies $Q^{\alpha}_{\epsilon|\tilde{x}} = Q^{\alpha}_{\epsilon|x} = Q^{\alpha}_{\epsilon|x^a}$ for any $\alpha \in (0,1)$ and any such $\tilde{x}$. Recall α' is defined as

$$\alpha'(x, \tilde{x}; \alpha) \equiv 1 - (1-\alpha)\tfrac{p(x)}{p(\tilde{x})} \tag{20}$$

for any such $\tilde{x}$ and $\alpha \in (0,1)$. Note $\alpha < \alpha' < 1$, because $p(x) < \frac{1}{2} \leq p(\tilde{x})$ by definition of Ω^+_X. Also note $\gamma_x(\alpha) = \gamma_{\tilde{x}}(\alpha')$ by construction. Therefore $Q^{1-\gamma_x(\alpha)}_{\epsilon|x} = Q^{1-\gamma_{\tilde{x}}(\alpha')}_{\epsilon|x^a} = Q^{1-\gamma_{\tilde{x}}(\alpha')}_{\epsilon|\tilde{x}} = c(\tilde{x}) - Q^{\alpha'}_{Y|D=1,\tilde{x}}$ where the first and second equalities follow from ER and the third follows from (19). Finally, note $Q^{\alpha'}_{Y|D=1,\tilde{x}}$ is directly identified and $c(\tilde{x})$ is recovered as $Q^{0.5}_{Y|\tilde{x}}$ because $\tilde{x}$ is chosen from Ω^+_X. Therefore for any $x = (x^a, x^b) \notin \Omega^+_X$, the median cake function c is over-identified as

$$\begin{aligned} c(x) &= Q^{\alpha}_{Y|D=1,x} + c(\tilde{x}) - Q^{\alpha'}_{Y|D=1,\tilde{x}} \\ &= Q^{\alpha}_{Y|D=1,x} + Q^{0.5}_{Y|\tilde{x}} - Q^{\alpha'}_{Y|D=1,\tilde{x}} \end{aligned} \tag{21}$$

for any $\tilde{x} = (x^a, \tilde{x}^b) \in \Omega^+_X$ and $\alpha \in (0,1)$. $\quad Q.E.D.$

The results in Proposition 3 are intuitive. Part *(i)* identifies the discount factor using a similar argument to the one in Proposition 1. The only difference is that when the cakes are censored and are only observed when agreement occurs, the discounted total continuation payoff $\beta\pi_w$ in each state is identified as the lower bound of the support of observed cakes in that state. Parts *(ii)* and *(iii)* follow from the observation that for states where the probability of reaching an agreement is greater than one half ($x \in \Omega_X^+$), the conditional median of cake sizes is directly observable.[12] This is a direct implication of the conditional median independence (MI) assumption.

The exclusion restriction (ER) and the assumption of sufficient variation in excluded states (SV) are instrumental for identifying median cake sizes conditional on states where the probability of reaching an agreement is less than one half ($x \notin \Omega_X^+$) in part *(iv)*. The reasoning behind the proof is as follows. Even for any $x \equiv (x^a, x^b) \notin \Omega_X^+$, additive separability of the cake function (AS) and rich support of unobserved states (RS) imply the conditional ρ-quantile of uncensored cake sizes $Q_{Y|x}^{\rho} = c(x) - Q_{\epsilon|x}^{1-\rho}$ can be recovered for some ρ close enough to 1. Hence, identification of $c(x)$ only hinges on identification of $Q_{\epsilon|x}^{1-\rho}$. Under *ER,* $Q_{\epsilon|x}^{1-\rho} = Q_{\epsilon|\tilde{x}}^{1-\rho}$ for any $\tilde{x} = (x^a, \tilde{x}^b)$. The role of SV is to ensure there exists such a $\tilde{x} \in \Omega_X^+$, so that the latter quantile can be recovered as $Q_{\epsilon|\tilde{x}}^{1-\rho} = c(\tilde{x}) - Q_{Y|\tilde{x}}^{\rho} = Q_{Y|\tilde{x}}^{0.5} - Q_{Y|\tilde{x}}^{\rho}$.[13] The regularity condition in Proposition 3 that H is a self-map ensures that the total continuation payoffs are bounded. This requirement is natural and sensible in most applications.

4.2 Consistent estimation

As in Section 3.4, our multi-step estimation procedure consists of plugging in sample analogs in the identification arguments presented above. Also, we only construct nonparametric estimators for the case with a finite Ω_X for expositional simplicity. The first step is to estimate the total continuation payoff as well as the discount factor β. For each $x \in \Omega_X$, estimate $\beta\pi_w(x)$ by:

$$\hat{\lambda}_{\mathcal{I}}(x) \equiv \inf\{Y_{g,t} : X_{g,t} = x, D_{g,t} = 1\} \tag{22}$$

where subscripts $\mathcal{I}$ are used here to distinguish estimators in the case of incomplete data with censored cakes from the ones in the previous section. Under appropriate regularity

[12] Our argument can be extended to show identification under any general conditional quantile independence (i.e. $Q_{\epsilon|x}^{\alpha} = Q^*$ for all $x \in \Omega_X$, where $\alpha \in (0, 1)$ and $Q^* \in \mathbb{R}^1$ are known constants).

[13] Chen, Dahl and Khan (2005) use a similar argument to nonparametrically identify a censored regression with an independent error term that has multiplicative heterogeneity. Also, note that the argument for identification of the cake function in part *(iv)* also applies for all $x \in \Omega_X^+$ such that there exists $\tilde{x} \in \Omega_X^+$ with $p(\tilde{x}) > p(x)$. Hence, the cake function is overidentified on Ω_X^+.

conditions, $\hat{\lambda}_{\mathcal{I}}(x) \xrightarrow{p} \tilde{\lambda}(x)$ for all x. Then define:

$$\hat{m}_{\mathcal{I}}(x) \equiv \frac{\sum_{g,t} 1\{X_{g,t}=x\}\left[D_{g,t}Y_{g,t} + (1-D_{g,t})\,\hat{\lambda}_{\mathcal{I}}(x)\right]}{\sum_{g,t} 1\{X_{g,t}=x\}}. \tag{23}$$

The estimator for the discount factor $\hat{\beta}_{\mathcal{I}}$ is defined similarly to (16), but with $\hat{\lambda}$ and $\hat{m}$ replaced by $\hat{\lambda}_{\mathcal{I}}$ and $\hat{m}_{\mathcal{I}}$. Arguments similar to those in Section 3.4 apply to show that $\hat{\beta}_{\mathcal{I}} \xrightarrow{p} \beta$ under appropriate conditions.

The proof of identification of the cake function in Proposition 3 holds for any $\alpha \in (0,1)$ and appropriately chosen $\tilde{x}$. Therefore the cake function c is over-identified. We propose an estimator of $c(x)$ for all $x \equiv (x^a, x^b) \in \Omega_X$ that exploits such an over-identification by averaging over multiple pairs of $(\alpha, \tilde{x})$'s. First, estimate Ω_X^+ by the set of x's for which the null hypothesis $H_0 : p(x) \geq 1/2$ cannot be rejected at some significance level δ. That is,

$$\hat{\Omega}_X^+ \equiv \{x \in \Omega_X : \hat{p}(x) \geq \frac{1}{2}(1 - z_{1-\delta} n_x^{-1/2})\}$$

where $z_{1-\delta}$ denotes the $(1-\delta)$-quantile of the standard normal distribution and n_x is defined as in Section 3.4.[14] Then, for any $x \in \hat{\Omega}_X^+$, estimate $c(x) = Q^{0.5}_{Y|x}$ by the median of the empirical analog of Y^c defined in Section 4.1, $Y^c_{g,t} \equiv D_{g,t}Y_{g,t} + (1-D_{g,t})\sum_{x \in \Omega_X} \hat{\lambda}_{\mathcal{I}}(x) 1\{X_{g,t}=x\}$. That is, $\hat{c}(x) \equiv \hat{Q}^{0.5}_{Y^c|x}$.

For any $x \notin \hat{\Omega}_X^+$, estimate $c(x)$ as follows. First, specify an arbitrary $\alpha \in (0,1)$ and estimate $Q^{\alpha}_{Y|D=1,x}$ by the conditional empirical α-quantile of Y under agreement in state x (denoted by $\hat{Q}^{\alpha}_{Y|D=1,x}$). This is done by inverting the empirical distribution of Y conditional on $D=1$ and x at α. Second, select some $\tilde{x} \equiv (\tilde{x}^a, \tilde{x}^b)$ such that $\tilde{x}^a = x^a$ and $\tilde{x} \in \hat{\Omega}_X^+$. For any such pair $(\alpha, \tilde{x})$, estimate α' by $\hat{\alpha}' = 1 - (1-\alpha)\hat{p}(x)/\hat{p}(\tilde{x})$, and then estimate $Q^{\alpha'}_{Y|D=1,\tilde{x}}$ by $\hat{Q}^{\hat{\alpha}'}_{Y|D=1,\tilde{x}}$. Define an estimator for $c(x)$ associated with this pair of $(\alpha, \tilde{x})$ as

$$\hat{c}_{\alpha}(x;\tilde{x}) \equiv \hat{Q}^{\alpha}_{Y|D=1,x} + \hat{Q}^{0.5}_{Y^c|\tilde{x}} - \hat{Q}^{\hat{\alpha}'}_{Y|D=1,\tilde{x}}. \tag{24}$$

Finally, repeat the previous two steps to construct estimators for other α's and $\tilde{x}$'s, and take the average of these estimators. Specifically, for any $x \notin \hat{\Omega}_X^+$, define

$$\hat{c}(x) \equiv (\#\mathcal{A})^{-1} \sum_{\alpha \in \mathcal{A}} \left(\frac{\sum_{\tilde{x} \in \Omega_X} \hat{c}_{\alpha}(x;\tilde{x}) 1\{\tilde{x}^a = x^a \wedge \tilde{x} \in \hat{\Omega}_X^+\}}{\sum_{\tilde{x} \in \Omega_X} 1\{\tilde{x}^a = x^a \wedge \tilde{x} \in \hat{\Omega}_X^+\}} \right), \tag{25}$$

where $\mathcal{A}$ is the set of arbitrarily chosen α (with its cardinality denoted by $\#(\mathcal{A})$). Proposition 4 proves point-wise consistency of $\hat{c}(x)$ for all $x \in \Omega_X$ under the following additional regularity conditions:

[14] Alternatively, we could use $\hat{\Omega}_X^+ \equiv \{x \in \Omega_X : \hat{p}(x) \geq \frac{1}{2}\}$ for consistent estimation of $c(x)$.

RG *(Regularity) (i) The distribution* $F_{Y|D=1,x}$ *is absolutely continuous and has positive density* $f_{Y|D=1,x}$ *with respect to the Lebesgue measure over its conditional support; (ii) for all* x *and* $\alpha \in (0,1)$*, there exists* $e > 0$ *such that* $f_{Y|D=1,x}$ *is bounded below by some positive constant* h *over the e-neighborhood around the quantile* $Q^{\alpha}_{Y|D=1,x}$, $\mathcal{N}_e(Q^{\alpha}_{Y|D=1,x})$*, where both* e and h *may depend on* α and x*; and (iii)* $\Pr\{p(X) = \frac{1}{2}\} = 0$.

Proposition 4 *Suppose CI, AS, RS, MI, ER, SV and RG hold. Then* $\hat{c}(x) \xrightarrow{p} c(x)$ *for all* $x \in \Omega_X$*.*

4.3 An application: corporate bankruptcy reorganization

In this section, we illustrate our methodology by performing nonparametric estimation of a simple model of corporate bankruptcy reorganization in the United States. When a company files for Chapter 11 bankruptcy, it keeps operating while its claimants (which include both debt-holders and equity-holders) negotiate over a plan to reorganize the firm as an alternative to its liquidation. In particular, the claimants bargain over the allocation of the company's reorganization value after the company emerges from bankruptcy. For this event to occur, all classes of creditors (which include secured and unsecured creditors as well as equity-holders) have to agree on a plan to restructure the bankrupt firm. Clearly, the potential value from reorganizing the company can vary during the negotiating period due to fluctuations in aggregate conditions (such as, for example, fluctuations in interest rates), or changes in industry- and firm-specific conditions that are reflected, for example, in the movement of stock prices. The reorganization value may also depend on private information revealed to the claimants while negotiations take place, but not observable by others. Hence, this situation represents a natural application of the stochastic bargaining framework.

Our goal here is not to provide a comprehensive empirical study of corporate bankruptcy reorganizations in the U.S., for which we refer the reader to Eraslan (2008).[15] More simply, our aim is to illustrate the feasibility of our nonparametric identification and estimation strategies in the context of a real application. The U.S. Corporate Bankruptcy Data (UCBD) collected by Eraslan (2008) contains information on 128 large, publicly held firms that filed for Chapter 11 bankruptcy between 1990 and 1997, and had a confirmed reorganization plan by the end of 2000. In addition to the beginning and end dates of the negotiations, and the

[15] Eraslan (2008) specifies a bargaining model of corporate bankruptcy reorganization that explicitely incorporates the role of the court in Chapter 11 bankruptcies and the possibility that a case is converted to Chapter 7. She estimates the model parametrically using a novel data set she collected. She then uses the estimated structural model to conduct counterfactual experiments to quantify the liquidation values of bankrupt firms that successfully reorganize and to assess the consequences of a mandatory liquidation policy.

dollar values of the total claims of the creditors at the time of filing, for 77 firms the data also contain their reorganization values and the agreed upon allocations among their claimants. However, since when a firm files for Chapter 11 its stock is suspended from trading until the firm emerges from bankruptcy, the potential reorganization value of a firm during the negotiation over its reorganization is not publicly observable. Hence, this situation fits the incomplete data scenario with censored cakes. Table I reports descriptive statistics for the 77 bargaining episodes we use in our analysis.[16]

Table I: Descriptive statistics

	mean	std. dev.	median
Duration (days)	489.33	505.37	360
Firm value (million $)	658.13	971.87	394.74
Total claims (million $)	956.24	1353.10	466.36
Firm value/Total claims	0.78	0.45	0.72

To capture the evolution of state variables that may affect the value of a firm while the claimants bargain over its reorganization, we supplement the UCBD with time-series data on interest rates of 3-month Treasury bills from the Federal Reserve Economic Data (FRED) and industry stock price indices from COMPUSTAT over the relevant time periods. For each firm in the data, we use its industry classification defined by the Global Industry Classification Standard (GICS) to collect the time-series of the stock price index pertaining to the firm's industry between the time it files from Chapter 11 and the time its reorganization plan is confirmed.[17]

Given the relatively small size of the sample, it would not be feasible to estimate nonparametrically a bargaining model of corporate bankruptcy reorganizations with heterogeneous firms and a large state space. For this reason, we abstract from firm- and industry-level heterogeneity across bargaining episodes and estimate our model using a simple structure for the state space. Using the notation of the canonical model of stochastic sequential bargaining presented in Section 2, we take a period t to be six months.[18] We normalize the

[16] For a detailed description of the dataset see Eraslan (2008).

[17] To express stock prices in real terms, we divide them by the consumer price index obtained from the Bureau of Labor Statistics. To match each firm in the UCBD with its industry classification code in GICS we use standard SIC codes for industry classifications which are provided in the UCBD. For each of the six observations in the UCBD that do not have an SIC code, we assign a GICS classification based on the description of the business scope contained in the company's website.

[18] As pointed out by Eraslan (2008), proposals to restructure a bankrupt firm are complex objects that take

firms' reorganization values by dividing them by the total claims of all equity-holders and debt-holders (secured and unsecured) at the beginning of the negotiations and define the cake $c(x)-\varepsilon$ to be the normalized value, where ε is median-independent of x.[19] The observable states, $X=\left(X^a,X^b\right)$, are the fluctuations in industry stock price indices (X^a) and in the interest rate (X^b). To simplify the analysis, we focus on *changes* in the stock prices and the interest rate, and define states in terms of changes relative to the previous period. In particular, the state space has four possible values: $x_1 \equiv (\textit{Stock} \uparrow, \textit{Interest} \uparrow)$, $x_2 \equiv (\textit{Stock} \uparrow, \textit{Interest} \downarrow)$, $x_3 \equiv (\textit{Stock} \downarrow, \textit{Interest} \uparrow)$ and $x_4 \equiv (\textit{Stock} \downarrow, \textit{Interest} \downarrow)$.[20] The state unobservable to the econometricians, ϵ, contains idiosyncratic information about a firm's value that is only observed by the claimants. We apply the estimators described in Section 4.2 to recover the discount factor β and the cake function $c(x)$. In particular, we aim at inferring the conditional medians of the potential normalized reorganization values given fluctuations in the observable states.

Given our definition of states, we can justify the two main identifying restrictions we introduced in Section 4.1: exclusion restrictions (ER) and sufficient variation (SV). Specifically, unobserved noises which affect a firm's reorganization value are more likely to pertain to firm-level rather than aggregate information. Once movements in stock prices are controlled for, it is plausible to assume that such information is orthogonal to macroeconomic conditions such as the interest rate. On the other hand, fluctuations in interest rates affect the cost of capital on the market, and can have a substantial impact on how claimants forecast the reorganization values. Table II reports the sample probability of reaching an agreement in each of the four possible states and results from one-sided tests for the simple null hypothesis H_0: $\textit{Pr(agreement}|x) \geq 1/2$. Our test in Table II shows some evidence that changes in the interest rate can have a significant effect on median normalized reorganization values, regardless of movements in stock prices. The last column in Table II also suggests that we could not reject the null that conditional agreement probabilities are above 50% for states x_1 and x_3. This observation is consistent with our identifying condition in SV, which

time to formulate and implement. For example, each class of creditors has to separately vote on a proposed plan before it can be confirmed. When a period is equal to six months, our data contains observations on 204 bargaining periods across all bargaining episodes.

[19] As Table I shows, the firms in the sample differ considerably in their realized reorganization values. Such differences may be ascribed to industry- or firm-level factors that are not captured by our definition of states. The normalization we adopt is meant to partially account for such heterogeneities across bargaining episodes.

[20] A $\uparrow$ indicates that the value of the variable in the current period is strictly larger than its value in the previous period. A possible justification for this simplification is that claimants are often more concerned about the momentum of these variables, which tend to be more predictive about future changes.

requires that changes in interest rates be conditionally independent from other unobserved firm-specific factors given the changes in stock prices.

Table II: Summary statistics for states and agreements

States (x)	$\hat{p}(x)$	*p-value*
$x_1 \equiv$ (*Stock* $\uparrow$, *Interest* $\uparrow$)	0.*49*	*0.44*
$x_2 \equiv$ (*Stock* $\uparrow$, *Interest* $\downarrow$)	0.*28*	*0.00*
$x_3 \equiv$ (*Stock* $\downarrow$, *Interest* $\uparrow$)	0.*56*	*0.76*
$x_4 \equiv$ (*Stock* $\downarrow$, *Interest* $\downarrow$)	0.*32*	*0.01*

Notes: $\hat{p}(x)$ *are sample probabilities for reaching an agreement in state* x *(e.g.* $\hat{p}(x^4) = 0.32$ *means that among all bargaining periods in the data when the state is* x^4*, in* 32% *of the cases claimants are observed to reach an agreement. The p-value is based on a one-sided test for the null* H_0*:*$p(x) \geq 1/2$*.)*

In our estimation, we use $\hat{\Omega}_X^+ = \{x_1, x_3\}$. Table III reports point estimates and 90% confidence intervals for $c(x)$ (i.e. conditional medians) for each possible state x. The point estimate of β is 0.53, with a 90% confidence interval of [0.52, 0.70].[21] The confidence intervals are constructed by using a bootstrap resampling method as follows. We construct 500 bootstrap samples by drawing from our estimation data with replacement. We then calculate our estimates as described in Section 4.2 for each one of the bootstrap samples. The 90% confidence intervals for β and $c(x)$ are then constructed by pairing the 5th and the 95th percentile of all point estimates from the 500 bootstrap samples.

Table III: Estimates and confidence intervals for median cakes

States (x)	$\hat{c}(x)$	*90%-C.I. for* $c(x)$
$x_1 \equiv$ (stock $\uparrow$, interest $\uparrow$)	0.38	[0.31 , 0.53]
$x_2 \equiv$ (stock $\uparrow$, interest $\downarrow$)	0.26	[0.14 , 0.29]
$x_3 \equiv$ (stock $\downarrow$, interest $\uparrow$)	0.33	[0.26 , 0.65]
$x_4 \equiv$ (stock $\downarrow$, interest $\downarrow$)	0.17	[0.03 , 0.43]

The point estimates in Table III suggest that, for a given direction of changes in the interest rate, the median of the normalized reorganization values is higher if stock prices

[21] While this estimate may seem low, the six-month discount factor in this setting also incorporates the risk of liquidation, which we do not model explicitly here.

have gone up. This is consistent with the intuition that an upward trend in stock prices lead to more optimistic expectations for reorganization values by the claimants. On the other hand, for a fixed direction of changes in stock prices, the median reorganization value is higher as the interest rate increases. This may be attributed to the fact that a higher interest rate translates into higher costs of borrowing on capital markets. Thus, if claimants are short of cash, they may find it appealing to realize the reorganization values earlier so as to avoid the higher interests.

Table IV reports bootstrap confidence intervals for differences in median cake sizes under various states. At the 90% level, an increase in stock prices does not have a significant effect on the size of the cake, regardless of the fluctuations in the interest rate. On the other hand, increasing interest rates can lead to a significantly higher chance for claimants to reach an agreement if stock prices have gone up from the previous period. Taken together, these patterns seem to suggest that movements in the interest rate have a more pronounced effect on the probability of reaching agreements than changes in stock prices.

Table IV: Confidence intervals for difference in cakes

Δc	*90%-C.I. for* Δc
$c(x^1) - c(x^3)$	[-0.31 , 0.20]
$c(x^2) - c(x^4)$	[-0.21 , 0.23]
$c(x^1) - c(x^2)$	[0.05 , 0.38]
$c(x^3) - c(x^4)$	[-0.13 , 0.59]

5 Incomplete Data with Unobserved Cakes

We now discuss nonparametric identification of the canonical stochastic bargaining model in the scenario with the least data. In particular, we consider the case where researchers only observe the duration of the negotiation and the evolution of states X in each bargaining episode in the data, but never observe the size of the cake or the identity of the proposer. As we pointed out in Section 2, there exist several large data sets of this sort.[22]

The definition of observational equivalence of parameters needs to be modified to fit this scenario. Let $F_{D|X}$ denote the probability of agreement in state X. A vector of parameters is

[22] In addition to the data used by Diaz-Moreno and Galdon (2000) and Simcoe (2008) which we already mentioned, there are many instances of legal disputes that are settled out of court where the beginning and end dates of the disputes are recorded, but the details of the negotiations or the terms of the settlements are not disclosed.

observationally equivalent to the truth in the data-generating process if it implies the same $F_{D|x}$ as that observed in the data for all $x \in \Omega_X$. The definition of identification of features of the true parameters remains instead the same as in Section 3.

Our starting point for analyzing identification is that the model is correctly specified under our conditional independence assumption (CI) for some $(\beta, c, F_{\epsilon|X})$. This means that the distribution of agreements observed necessarily satisfies restrictions implied by the model assumptions (i.e. $F_{D',X'|D,X} = F_{D'|X'}G_{X'|X}$). As in the case of incomplete data with censored cakes analyzed in Section 4, throughout this section we maintain the assumption that the cake function is additively separable (AS): $c(x, \varepsilon) = c(x) - \varepsilon$.

For the case where the discount factor β and the distribution $F_{\epsilon|X}$ are known to researchers, it is fairly easy to show that $c(x)$ is identified for all x.[23] The separation principle in Lemma 1 states that an agreement occurs if and only if the total surplus exceeds the total continuation payoff in a SSPE, which has the same characterization as the continuation value in a single-agent optimal stopping problem. Hence, the arguments in Berry and Tamer (2006) on the identification of optimal stopping models can be applied here to show identification of c.[24]

However, in practice, it is often the case that researchers do not know a priori the distribution of unobserved states. In this case, misspecifying $F_{\epsilon|X}$ can lead to incorrect predictions of outcomes (i.e. probabilities of agreements conditional on observable states) under hypothetical changes in the transition of states or in the cake function. On the other hand, economic theory often implies that the cake function or the unobserved state distribution must satisfy certain shape or stochastic restrictions (such as monotonicity of c or independence of ϵ from X). Such conditions help restrict counterfactual outcomes to a subset that is consistent with model restrictions and outcomes observed in the data-generating process (DGP). We refer to this set as the *identified set of counterfactual outcomes* (ISCO).

For the rest of this section, we restrict ϵ to be independent of X, and the support of X to be finite (i.e. $\Omega_X \equiv \{x_1, ..., x_M\}$).[25] It is therefore convenient to introduce vector notation.

[23] The assumption that the discount factor β is known to researchers can be justified in situations where it can be directly recovered from the data. This is typically the case in macroeconomic applications where the discount factor is specified as $\beta \equiv 1/(1+r)$ where r is the interest rate.

[24] If β is not known, their arguments can also be extended to show that β can be identified as long as $F_{\epsilon|X}$ is known for all X and the cake function is known for some value of X. The proofs are relatively straightforward and are therefore omitted.

[25] The case of discretized state spaces is particularly important in the empirical literature on structural estimation. This assumption also simplifies our exposition of the characterization of the ISCO. We conjecture

Let the cake function be denoted by a M-vector C with coordinates $C_m \equiv c(x_m)$, and the transition of observable states be denoted by a $M \times M$ transition matrix G with entries $G_{mn} \equiv \Pr(X' = x_n | X = x_m)$, $m, n = 1, ..., M$. The observed outcome in a SSPE is then summarized by the M-vector of conditional agreement probabilities $p \equiv (p(x_1), ..., p(x_M))$. Also, we focus on shape restrictions on the cake function that can be represented as a system of linear inequalities on C, $BC > 0$, where B is a known matrix with as many rows as the number of restrictions. For example, partial or complete rankings of the sizes of the cake in different states, as well as monotonicity, additive separability, or super-modularity of the cake function in a subvector of x can all be expressed as linear restrictions on C.

We propose a simple algorithm for recovering the ISCO under two types of hypothetical changes in the model structure: the transition between states is changed from G to $\tilde{G}$ while the discount factor β, the cake function C and the distribution of unobserved states F_ϵ remain unchanged; or the cake function is changed from C to $\tilde{C} \equiv \gamma C$, while β, G, F_ϵ remain unchanged. (Here, γ is a known diagonal matrix with positive diagonal elements denoting percentage changes in the cake sizes in different states.) By definition, the task of recovering the ISCO under a given set of model restrictions amounts to finding all counterfactual outcomes $\tilde{p} \in (0,1)^M$ such that both $\tilde{p}$ and the outcomes observed from the DGP p are jointly implied in a SSPE by the same structure C, F_ϵ satisfying these restrictions.

To illustrate our algorithm consider the counterfactual environment where the transition between states is changed from G to $\tilde{G}$ while β, C, F_ϵ are kept the same. The algorithm recovers the ISCO by exploiting two simple observations about the structural link between model primitives and implied outcomes (i.e. the vector of agreement probabilities). First, in the DGP, the characterization of SSPE consists of a system of M equalities relating the vector of outcomes p to the cake function C and certain nuisance parameters of F_ϵ (which include M quantiles and M truncated expectations). Analogously, the characterization of SSPE for the counterfactual environment also consists of a system of M equalities relating the vector of outcomes $\tilde{p}$ to the same cake function C and a different set of M quantiles and M truncated expectations of the same distribution of unobserved states F_ϵ. Second, the shape restrictions on C and the independence between ϵ and X can be formulated as inequalities restricting C and all these nuisance parameters ($2M$ parameters for the DGP and $2M$ for the counterfactual environment), respectively. This system of inequalities provides the structural link between the DGP and the counterfactual environment. For example, all of the $2M$ unknown quantiles $(F_\epsilon^{-1}(p(x_m)), F_\epsilon^{-1}(\tilde{p}(x_m))_{m=1}^M$ must be ranked in the same order as the corresponding agreement probabilities $(p_m, \tilde{p}_m)_{m=1}^M$ (e.g. $p(x_m) < \tilde{p}(x_k)$ if and only if

that there is a continuous analog for our partial identification arguments below.

$F_\epsilon^{-1}(p(x_m)) < F_\epsilon^{-1}(\tilde{p}(x_k)))$. The combined system consists of M equalities for the DGP, M equalities for the counterfactual environment and all the inequalities linking the $4M$ nuisance parameters of F_ϵ to the implied outcomes p and $\tilde{p}$. This system is linear in the unknown structural elements (i.e. C and nuisance parameters of F_ϵ), and the two vectors of implied outcomes p and $\tilde{p}$ enter the system through the matrix of coefficients. As a result, a vector $\tilde{p}$ belongs to the ISCO if and only if it is such that the linear system admits solutions in C and nuisance parameters of F_ϵ given the outcomes observed in the DGP p. The algorithm for the second counterfactual environment is similar except that the cake function in the counterfactual environment is γC.

We formalize these ideas in Proposition 5 below. For a generic vector R, let R_j and $R_{(j)}$ denote its j-th coordinate and its j-th smallest coordinate, respectively. Let $\phi > 0$ be an arbitrary positive constant chosen to normalize the scale of C and F_ϵ.

Proposition 5 *Suppose CI and AS hold; C satisfies the shape restrictions $BC > 0$; ϵ is independent from X; β is known; and $p \in (0,1)^M$ is the vector of outcomes observed form the data-generating process. Then: (i) a vector $\tilde{p}$ is in the ISCO under the counterfactual transition of states $\tilde{G}$ if and only if, for all $j, l = 1, ..., 2M+1$ and $m = 1, ..., 2M$, there exist $R^Q, \tilde{R}^Q \in \mathbb{R}^M$ and $R^\Phi, \tilde{R}^\Phi \in \mathbb{R}^M_{++}$ that satisfy:*

$$R^Q + \beta(I - \beta G)^{-1} G R^\Phi = \tilde{R}^Q + \beta(I - \beta \tilde{G})^{-1} \tilde{G} \tilde{R}^\Phi \tag{26}$$

$$B\left[R^Q + \beta(I - \beta G)^{-1} G R^\Phi\right] > 0 \tag{27}$$

$$\bar{Q}_j \leq \bar{Q}_l \text{ if and only if } \bar{p}_j \leq \bar{p}_l \tag{28}$$

$$\bar{p}_{(m)}(\bar{R}^Q_{(m+1)} - \bar{R}^Q_{(m)}) \leq \bar{R}^\Phi_{(m+1)} - \bar{R}^\Phi_{(m)} \leq \bar{p}_{(m+1)}(\bar{R}^Q_{(m+1)} - \bar{R}^Q_{(m)}) \tag{29}$$

where $\bar{p} \equiv \left(p, \tilde{p}, \frac{1}{2}\right)$, $\bar{R}^Q \equiv \left(R^Q, \tilde{R}^Q, 0\right)$ and $\bar{R}^\Phi \equiv \left(R^\Phi, \tilde{R}^\Phi, \phi\right)$; and (ii) the characterization of the ISCO under the counterfactual cake function $\tilde{C}$ is the same as in part (i), except that (26) is replaced by:

$$\tilde{R}^Q + \beta(I - \beta G)^{-1} G \tilde{R}^\Phi = \gamma\left(R^Q + \beta(I - \beta G)^{-1} G R^\Phi\right). \tag{30}$$

Proposition 5 implies that recovering the ISCO's under both types of counterfactual changes is equivalent to finding all $\tilde{p}$'s such that the linear systems (26)-(30) admit solutions. Standard linear programming algorithms can then be used to find such $\tilde{p}$'s.[26]

[26] In an earlier version of the paper, we presented a simple numerical example with a low-dimension state space and showed the ISCO recovered is informative. It is also practical to conduct Bayesian inference of

Conditions (26)-(29) all have intuitive interpretations. Equation (26) derives from the requirement that the cake function remains unchanged as the transition of states shifts from G to $\tilde{G}$ in the first type of counterfactual exercises. To see this, let Q, Π and Φ denote M-vectors with coordinates $Q_m \equiv F_\epsilon^{-1}(p_m)$, $\Pi_m \equiv \pi_w(x_m)$ and $\Phi_m \equiv \int \max\{Q_m - \varepsilon, 0\} dF_\epsilon$, respectively.[27] The SSPE outcome from the data-generating process is characterized by $Q = C - \beta\Pi$, where $\Pi = G(\Phi + \beta\Pi)$. Thus, the cake function must be related to p as follows:

$$C = Q + \beta(I - \beta G)^{-1} G\Phi. \tag{31}$$

With C, F_ϵ unchanged but G replaced by $\tilde{G}$, we can derive a similar structural equation $C = \tilde{Q} + \beta(I - \beta\tilde{G})^{-1}\tilde{G}\tilde{\Phi}$, where $\tilde{Q}, \tilde{\Phi}$ are the same as Q, Φ except that $\tilde{p}$ (the implied SSPE outcomes under $\tilde{G}$) replaces p (the outcomes in the DGP). This means (26) must hold with $(R^Q, \tilde{R}^Q, R^\Phi, \tilde{R}^\Phi) = (Q, \tilde{Q}, \Phi, \tilde{\Phi})$, because the cake function remains the same both in the data-generating process and the counterfactual context under $\tilde{G}$. Equation (27) follows from the shape restrictions $BC > 0$.

Conditions in (28)-(29) result from two considerations. First, F_ϵ remains unchanged both in the data-generating process and the counterfactual context. Second, the independence between ϵ and X can be formulated in terms of inequality restrictions which are linear in $Q, \Phi, \tilde{Q}, \tilde{\Phi}$ and have $p, \tilde{p}$ enter in the coefficient matrix. For example, such independence implies the coordinates in Q must be ordered in the same way as in p; and $\Phi_m - \Phi_n = \int_{Q_n}^{Q_m} F_\epsilon(s) ds$ must lie between $p_n(Q_m - Q_n)$ and $p_m(Q_m - Q_n)$ for any pair (m, n) such that $p_m \geq p_n$. By the same reasoning, a similar set of linear restrictions involving $\tilde{p}, \tilde{Q}, \tilde{\Phi}$ can be derived for the counterfactual context. Because the unobserved state distribution remains unchanged both in the data-generating process and the counterfactual context, these two sets of restrictions can be combined into a single system as in (28)-(29).

It follows that a $\tilde{p}$ can be rationalized in a SSPE by certain C, F_ϵ satisfying the shape and independence restrictions if and only if the linear system in Proposition 5 admits solutions $(R^Q, \tilde{R}^Q, R^\Phi, \tilde{R}^\Phi) = (Q, \tilde{Q}, \Phi, \tilde{\Phi})$. Finally, note that the choice of ϕ has no impact on the feasibility of the linear system in Proposition 5.[28]

the ISCO in our model. For a dynamic model where a single agent chooses between binary actions each period, Norets and Tang (2010) propose a Markov Chain Monte Carlo (MCMC) algorithm for simulating the posterior of counterfactual choice probabilities. A similar approach is possible for our model here. However, we leave the inference of ISCO in stochastic bargaining games for future research.

[27] Norets and Tang (2010) use a similar argument to characterize the identified set of counterfactual choice probabilities in a model of dynamic binary choice processes.

[28] If a solution exists for a given ϕ, changing the constant to ϕ' would simply require a rescaling of the solution. This is not surprising because the scale of C and F_ϵ cannot be jointly identified with model

We conclude this section by noting that the conditions (i) and (ii) in Proposition 5 are not only *necessary but also sufficient* for a $\tilde{p}$ to be implied in a SSPE by some C, F_ϵ satisfying the shape restrictions $BC > 0$ and the independence of ϵ from X for each of the two counterfactual environments, respectively. Hence, the ISCO characterized above reveals the limit of what can be learned about counterfactual outcomes under these restrictions.

6 Conclusions

In this paper, we have presented positive results in the identification of structural elements and counterfactual outcomes in the canonical model of stochastic sequential bargaining under various data scenarios. A unifying theme of our analysis is that the model structure and the implied counterfactual outcomes can be point- or partially-identified under weak nonparametric restrictions (such as shape restrictions on the cake function or independence of the distribution of unobserved state variables) given different data availability. We have also proposed consistent estimators for the discount factor and the cake function.

The canonical model of stochastic sequential bargaining assumes that utilities are directly transferable among players and through time at a constant rate (i.e. the players have linear utilities and share a common discount factor). This feature of the model implies that the game has a unique SSPE which satisfies the separation principle. Also, in the canonical model, the current unobserved state is assumed to be independent of past states conditional on current observed states and the proposer-selection process does not depend on the history of states. The identification strategy we have used in this paper relies on the separation principle and on these independence assumptions. Depending on the application, the assumptions of the canonical model may or may not apply. Merlo and Wilson (1995) provide equilibrium characterization results for a large class of bargaining games, including games with non-transferable utility and environments with more general stochastic structures governing the cake and proposer processes. In such environments, the bargaining game typically admits multiple SSPE and the occurrence of agreement is no longer independent of the bargaining protocol. These aspects pose additional challenges for estimation. We intend to pursue the nonparametric identification and consistent estimation of general stochastic bargaining models in future work.

structures and counterfactual outcomes.

Appendix A: Proofs

Proof of Lemma 1. It follows from Theorem 1 in Merlo and Wilson (1998) that the individual SSPE payoff is characterized for all $s \equiv (x, \varepsilon, \Bbbk)$ as

$$\begin{aligned} A_i(v(s)) &\equiv \max\{c(x,\varepsilon) - \beta E[\textstyle\sum_{j\neq i} v_j(S')|s] \ , \ \beta E[v_i(S')|s]\} \ \textit{for } i = \Bbbk \\ A_j(v(s)) &\equiv \beta E[v_j(S')|s] \ \textit{for } j \neq \Bbbk \end{aligned}$$

From Theorem 2 in Merlo and Wilson (1998), the total payoff in SSPE must satisfy the fixed point equation $w(s) = \max\{c(x,\varepsilon) \ , \ \beta E[w(S')|s]\}$ for all s, and that agreements occur for s if and only if $c(x,\varepsilon) \geq \beta E[w(S')|s]$. Under *CI*, for any function $f(S)$,

$$\begin{aligned} E[f(S')|s] &= \int E[f(S')|x',\varepsilon',s]dF(x',\varepsilon'|s) \\ &= \int E[f(S')|x',\varepsilon']dF(x',\varepsilon'|x,\varepsilon) = E[f(S')|x] \end{aligned}$$

where the equalities follow from the independence of (X', ϵ', κ') from (ϵ, κ) given X, and independence of κ' from past states and κ. Then (a), (b) and (c) in the lemma follow. The uniqueness of SSPE payoffs follows from Theorem 3 in Merlo and Wilson (1998). *Q.E.D.*

Proof of Lemma 2. *(Necessity)* Suppose there exists $\{\beta, c, F_{\epsilon|X}\}$ that satisfies MT, ND and rationalizes the distribution of $\{\tau, \eta_\tau, Y^\tau, X^\tau, \kappa^\tau\}$ under CI. Recall $Y_t = c(X_t, \epsilon_t)$ and by Lemma 1, $D_t = 1$ if and only if $Y_t \geq \beta E(w(X', \epsilon')|x, \varepsilon)$ in any SSPE. Hence under MT, the equality in (6) is implied by $F(\epsilon_{t+1}, X_{t+1}, \kappa_{t+1}|\varepsilon^t, x^t, \Bbbk^t) = L_{\kappa_{t+1}} F_{\epsilon_{t+1}, X_{t+1}|X_{t+1}} G_{X_{t+1}|x_t}$, which follows from CI. The time-homogeneity of $F_{Y_t, D_t|X_t}$ follows from time-homogeneity of $F_{\epsilon|X}$ and G. Under *CI*, $E(w(X', \epsilon')|x, \varepsilon) = E(w(X', \epsilon')|x)$ and hence $p(x) = \Pr\{Y \geq \beta E(w(X', \epsilon')|x) \mid x\} = 1- \ F_{Y|x}(\beta E(w(X', \epsilon')|x))$. Under MT, $F_{Y|x}(t) = \Pr\{c(X, \epsilon) \leq t \mid X = x\} = F_{\epsilon|x}(c^{-1}(x, t))$, and is strictly increasing in t on the support of Y given x (where $c^{-1}(x, .)$ is the inverse of $c(x, .)$ given x). Under ND, $p(x) \in (0, 1)$ for all x, and $\beta E(w(X', \epsilon')|x)$ must be in the interior of the support of $F_{Y|x}$ and can be recovered as $\lambda^*(x)$. Then by Lemma 1, *(ii)* and *(iii)* must hold in SSPE. By definition,

$$\begin{aligned} \beta E\left[w(X', \epsilon')|x\right] &= \beta \int\int \max\{c(x',\varepsilon'), \beta E\left[w(X'',\epsilon'')|x'\right]\} dF_{\epsilon'|X'} dG_{X'|x} \\ \Leftrightarrow \lambda^*(x) &= \beta \int\int \max\{c(x',\varepsilon'), \lambda^*(x')\} dF_{\epsilon'|X'} dG_{X'|x} \qquad (32) \\ &= \beta \int\int \max\{y', \lambda^*(x')\} dF_{Y'|X'} dG_{X'|x} \qquad (33) \end{aligned}$$

where the equality in (33) follows from changing variables between y and ε given x under MT. Condition *(iv)* must hold because the discount factor $\beta \in (0, 1)$. Lemma 2 of Merlo and

Wilson (1998) implies that the individual continuation payoffs in SSPE ($\pi_i(s) \equiv \beta E[v_i(S')|s]$) are defined by the unique solution to the fixed point equation:

$$\pi_i(s) = \beta \int \pi_i(s') + 1\{\kappa' = i\} \max\{c(x', \varepsilon') - \beta \textstyle\sum_i \pi_i(s'), 0\} dF_{S'|s}. \tag{34}$$

Under *CI* and the independence of κ' from past states and κ, we have $S' = (X', \epsilon', \kappa')$ independent of (ϵ, κ) given X. Hence π_i, as a solution to (34), depends on x only. Thus, (34) can be written as

$$\pi_i(x) = \beta \int \pi_i(x') + L_i \int \max\{y' - \beta \textstyle\sum_i \pi_i(x'), 0\} dF_{Y'|X'} dG_{X'|x} \tag{35}$$

using change-of-variables between Y' and ϵ' conditional on X'. In any agreement in a *SSPE*, a non-proposer is offered his discounted continuation payoff $\pi_i(x)$. Hence condition *(v)* follows from (35) with λ_i^* playing the role of π_i and $E[w(X', \epsilon')|x] \equiv \sum_i \pi_i(x) = \lambda^*(x)$ for all x.

(Sufficiency) Suppose a joint distribution of $\{\tau, \eta_\tau, Y^\tau, X^\tau, \kappa^\tau\}$ is observed and satisfies conditions for rationalizability *(i)*-*(v)*. We need to find certain $\{\beta, c, F_{\epsilon|X}, L\}$ that satisfy MT and ND, and could rationalize this joint distribution under *CI*. First, choose *any* strictly increasing distribution $F_{\epsilon|x}$ and define $c(x, \varepsilon) \equiv F_{Y|x}^{-1}(F_{\epsilon|x}(\varepsilon))$. By condition *(ii)*, $F_{Y|x}(t)$ is increasing in t given x. Hence $c(x, \varepsilon)$ is increasing in ε given x, and MT is satisfied. Furthermore,

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

over the support of Y given x. The equalities follow from *(ii)*, and that $F_{\epsilon|x}(\epsilon)$ is uniform on $[0, 1]$ given any x. Next, define $\beta \equiv \left(\int \max\{y', \lambda^*(x')\} dF_{Y', X'|X=x}\right)^{-1} \lambda^*(x)$ for any x. Under condition *(iv)*, β is between $(0, 1)$. Finally, L is defined with $L_i \equiv \Pr\{\kappa = i\}$. By construction, $\lambda^*(x)$ is the *unique* solution for the following fixed-point equation:

$$\lambda^*(x) = \beta \int \int \max\{y', \lambda^*(x')\} dF_{Y'|X'} dG_{X'|x} \tag{36}$$

where the r.h.s. of (36) is a contraction mapping in f. Using change-of-variables between Y' and ϵ' conditional on X', (36) can be written as

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

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

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

where the second equality follows from condition *(iv)*. The condition ND is also satisfied, for *(ii)* implies $p(x) \in (0,1)$. Under MT, equation (8) in condition *(v)* can be written as

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

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

$$F_t(S_{t+1}|s^t) \equiv \; L_{\kappa_{t+1}} F_{\epsilon|X}(\epsilon_{t+1}|X_{t+1}) G(X_{t+1}|x_t).$$

Since (6) holds with $L, F_{Y,D|X}$ and $G_{X'|X}$ being time-homogenous in *(i)*, inductive arguments show $\{\beta, c, F_{\epsilon|X}, L\}$ rationalizes the joint distribution of $\{\tau, \eta_\tau, X^\tau, Y^\tau, \kappa^\tau\}$ as long as $\{\beta, c, L, F_{\epsilon|X}\}$ can rationalize $F_{Y,D|X}$ and $F_{\eta_\tau | \kappa_\tau, D_\tau = 1, X_\tau}$. *Q.E.D.*

Proof of Proposition 1. The proof uses Lemma A1 below, which reveals the degree of under-identification in $(c, F_{\epsilon|X})$.

Lemma A1 Suppose the true discount factor β_0 is known. Then a pair $(c, F_{\epsilon|X})$ that satisfies MT and ND can rationalize the $F_{\tau, \eta_\tau, X^\tau, Y^\tau, \kappa^\tau}$ observed if and only if it implies the distribution of cakes $F_{Y|X}$ observed for all $x \in \Omega_X$.

Proof of Lemma A1. Necessity follows immediately from the definition of rationalization. To prove sufficiency, first note under *CI* and *MT*, changing variables between ϵ and Y shows the discounted total continuation payoff can be expressed as a unique solution for the following fixed-point equation for all x:

$$\beta_0 \pi_w(x) = \beta_0 \int \max\{y', \beta_0 \pi_w(x')\} dF_{Y', X'|X=x}. \tag{37}$$

Likewise, discounted continuation payoffs for individual i can also be expressed as the unique solution in f_i for:

$$\beta_0 \pi_i(x) = \beta_0 \int \beta_0 \pi_i(x') + L_i \int \max\{y' - \beta_0 \pi_w(x'), 0\} dF_{Y'|X'} dG_{X'|X=x} \tag{38}$$

with f given in (37). Both π_i and π_w, as unique solutions of (37) and (38), are determined by β_0, L, G and $F_{Y|X}$, where L, G are directly identifiable and β_0 is fixed and known by our supposition in the lemma.

Suppose a generic pair $(c, F_{\epsilon|X})$ is such that

$$F_{\epsilon|x}(c^{-1}(x,y)) = F_{Y|x}(y) = F^0_{\epsilon|x}(c_0^{-1}(x,y)) \tag{39}$$

for all x, y. Then it necessarily implies the same total and individual continuation payoffs as the true parameters $(c_0, F^0_{\epsilon|X})$ does. As a result, $(c, F_{\epsilon|X})$ must imply the joint distribution of cakes and decisions $F_{Y,D|X}$ and the distribution of accepted allocations $F_{\eta|D=1,X,\kappa}$ as $(c_0, F^0_{\epsilon|X})$ does.

We then complete the proof using inductive arguments. When $\tau = 0$, the observed distribution is

$$\Pr\{\tau = 0, \eta_0, \kappa_0 | X_0 = x\} = L_{\kappa_0} \Pr\{Y_0 \leq \beta_0 \pi_w(X_0), \eta_0 | X_0 = x\}$$

for all x, which is determined by $F_{Y|X}$ and L only. Hence any $(c, F_{\epsilon|X})$ that implies $F_{Y|x}$ for all x can also rationalize this joint distribution. It follows immediately that $\Pr\{\tau > 0, Y_0, \kappa_0 | X_0 = x\}$ is also rationalized by such a $(c, F_{\epsilon|X})$ and L. When $\tau = 1$, the observed distribution is:

$$\begin{aligned}
& \Pr(\tau = 1, \kappa_0, \kappa_1, Y_0, X_1, \eta_1 | x_0) \\
= \ & \Pr(D_1 = 1, \kappa_1, \eta_1 | X_1, D_0 = 0, x_0, Y_0, \kappa_0) \Pr(x_1 | D_0 = 0, Y_0, \kappa_0, x_0) \Pr(D_0 = 0, Y_0, \kappa_0 | x_0) \\
= \ & \Pr(\eta_1 | D_1 = 1, \kappa_1, X_1) L_{\kappa_1} \Pr(D_1 = 1 | X_1) G(X_1 | x_0) L_{\kappa_0} \Pr(D_0 = 0, Y_0 | x_0)
\end{aligned}$$

where the equalities follow from *CI* and that an observed, rationalizable distribution $\Pr(\tau = 1, \eta_\tau, X^\tau, Y^{\tau-1}, \kappa^\tau)$ necessarily satisfies the condition *(i)* in Lemma 2. Recall L, G are directly identified from data. As shown above, any $(c, F_{\epsilon|X})$ that satisfies (39) implies the same $F_{Y|x}$ (and therefore $F_{\eta|D=1,x,\kappa}$ and $F_{Y,D|x}$) for all x as the truth $(c_0, F^0_{\epsilon|X})$. Hence such a pair $(c, F_{\epsilon|X})$ can generate the same $\Pr(\tau = 1, \eta_\tau, X^\tau, Y^{\tau-1}, \kappa^\tau)$ as $(c^0, F^0_{\epsilon|X})$. It follows immediately that $\Pr(\tau > 1, X^1, Y^1, \kappa^1)$ is also rationalized by such a $(c, F_{\epsilon|X})$ and L.

Now suppose for some $t \geq 1$, a pair $(c, F_{\epsilon|X})$ rationalizes the observable distribution $\Pr(\tau = l, \eta_\tau, X^\tau, Y^{\tau-1}, \kappa^\tau)$ for all $l \leq t$ as well as $\Pr(\tau > t, X^t, Y^t, \kappa^t)$. Then consider the case with $\tau = t+1$. For any x_0,

$$\begin{aligned}
& \Pr(\tau = t+1, \eta_{t+1}, \kappa^{t+1}, X^{t+1}, Y^t | x_0) \\
= \ & \Pr(D_{t+1} = 1, \kappa_{t+1}, \eta_{t+1} | X_{t+1}, Y^t, X^t, \tau > t) \\
& \cdot \Pr(X_{t+1} | Y^t, X^t, \tau > t) \Pr(Y^t, X^t, \tau > t | x_0) \Pi_{l=0}^t L_{\kappa_l} \\
= \ & \Pr(\eta_{t+1} | D_{t+1} = 1, \kappa_{t+1}, X_{t+1}) \Pr(D_{t+1} = 1 | X_{t+1}) G(X_{t+1} | X_t) \\
& \cdot \Pr(\tau > t, Y^t, X^t | x_0) \Pi_{l=0}^{t+1} L_{\kappa_l}
\end{aligned}$$

where the equalities again follow from necessary conditions for rationalizability in Lemma 2. By supposition at the beginning of this induction step, $(c, F_{\epsilon|X})$ rationalizes $\Pr(\eta|D = 1, \kappa, X)$, $\Pr(D = 1|X)$ and $\Pr(\tau > t, Y^t, X^t|x_0)$. It also follows that $\Pr(\tau > t+1,\ X^{t+1}, Y^{t+1},.\kappa^{t+1})$ is rationalized by such a $(c, F_{\epsilon|X})$ and L. $\square$

That β_0 is identified follows from the proof of the necessity of *(iv)* in Lemma 2. Part *(ii)* of Proposition 1 follows immediately from Lemma A1 above. When $\epsilon \perp X$ and the distribution of ϵ is normalized to any F_ϵ, the cake function is recovered as $c_0(x,\varepsilon) = F^{-1}_{Y|x}(F_\epsilon(\varepsilon))$. This is because $F_{Y|x}(t), c_0(x,\varepsilon)$ are strictly increasing in t and ε respectively given any x and $F_{Y|x}(t) = F_{Y|x}(c_0(x,\varepsilon)) = F_\epsilon(\varepsilon)$. *Q.E.D.*

Proof of Proposition 2. It suffices to show $\hat{\lambda}(x) \xrightarrow{p} \lambda^*(x)$ and $\sum_{x'\in\Omega_X} \hat{G}_{x'|x}\hat{m}(x') \xrightarrow{p} \lambda^*(x)/\beta_0$ for all x. Then Slutsky's Theorem applies to show that $\hat{\beta} \xrightarrow{p} \left(\sum_{x\in\Omega_X} w_x\right)\beta_0 = \beta_0$. For any fixed x,the weak law of large numbers implies $\hat{p}(x) \xrightarrow{p} p(x)$. Also for any fixed x, the Glivenko-Cantelli Lemma suggests $\sup_{r\in\mathbb{R}} |\hat{F}_{Y|X}(r|x) - F_{Y|X}(r|x)| \to 0$ *almost surely.* It then follows from standard arguments that $\hat{\lambda}(x) \xrightarrow{p} \lambda^*(x)$ for all x. By the Uniform Law of Large Numbers, for any x,

$$\sup_{\zeta\in N_e(\lambda^*(x))} \left| \frac{\sum_{g,t} \max\{Y_{g,t},\zeta\} 1\{X_{g,t} = x\}}{\sum_{g,t} 1\{X_{g,t} = x\}} - E\left[\max\{Y,\zeta\}|X = x\right] \right| \xrightarrow{p} 0$$

Since $E[\max\{Y,t\}|X = x]$ is continuous at $t = \lambda^*(x)$ conditional on any x, it then follows from Theorem 4.1.5 in Amemiya (1985) that for all x,

$$\hat{m}(x) \xrightarrow{p} m_0(x)$$

Also note $\hat{G}_{x'|x} \xrightarrow{p} G_{x'|x}$, and hence Slutsky's Theorem implies

$$\textstyle\sum_{x'\in\Omega_X} \hat{G}_{x'|x}\hat{m}(x') \xrightarrow{p} \sum_{x'\in\Omega_X} G_{x'|x} m_0(x') = E[\max\{Y', \lambda^*(X')\}|X = x]$$

and recall from (11) that $E[\max\{Y', \lambda^*(X')\}|X = x] = \pi_w(x;\theta_0)$. Since $\pi_w(x;\theta_0) \neq 0$ by the regularity conditions stated in the proposition, it follows from Slutsky's Theorem and $\lambda^*(x) = \beta_0\pi_w(x;\theta_0)$ that:

$$\left[\textstyle\sum_{x'\in\Omega_X} \hat{G}_{x'|x}\hat{m}(x')\right]^{-1} \hat{\lambda}(x) \xrightarrow{p} \lambda^*(x)/\pi_w(x;\theta_0) = \beta_0$$

for all $x \in X$. Then $\hat{\beta} \xrightarrow{p} \beta_0$ follows from the fact $\sum_{x\in\Omega_X} w_x = 1$. *Q.E.D.*

Proof of Proposition 4. The proof of consistency builds on results in *Lemma A2-4* below. For any $x \equiv (x^a, x^b) \in \Omega_X$, define $S(x) \equiv \{\tilde{x} \in \Omega_X : \tilde{x}^a = x^a \wedge p(\tilde{x}) \geq \frac{1}{2}\}$ and

$\hat{S}_n(x) \equiv \{\tilde{x} \in \Omega_X : \tilde{x}^a = x^a \wedge \tilde{x} \in \hat{\Omega}_X^+\}$. Note $\hat{S}_n(x)$ implicitly depends on the significance level δ used for testing the null $H_0 : p(x) \geq 1/2$.

Lemma A2 For any $\delta \in (0,1)$, $\Pr\{\hat{S}_n(x) = S(x)\} \to 1$ *as* $n \to +\infty$ *for all* $x \in \Omega_X$.

Proof of Lemma A2. By definition, $\hat{S}_n(x) \neq S(x)$ if and only if $\exists \hat{x}$ such that $\hat{x}^a = x^a$ and either "$\hat{x} \notin \hat{\Omega}_X^+$ but $\hat{x} \in \Omega_X^+$" or "$\hat{x} \in \Omega_X^+$ but $\hat{x} \notin \hat{\Omega}_X^+$". Hence:

$$\begin{aligned} & \Pr\{\hat{S}_n(x) \neq S(x)\} \\ \leq \ & \sum_{\hat{x} \notin S(x)} \Pr\{n_{\hat{x}}^{1/2} [\hat{p}(\hat{x}) - 0.5] \geq -0.5 z_{1-\delta}\} + \sum_{\hat{x} \in S(x)} \Pr\{n_{\hat{x}}^{1/2} [\hat{p}(\hat{x}) - 0.5] < -0.5 z_{1-\delta}\} \end{aligned} \tag{40}$$

By definition, $p(\hat{x}) < \frac{1}{2}$ for any $\hat{x} \notin S(x)$. Hence as $n \to +\infty$,

$$\Pr\{n_{\hat{x}}^{1/2}[\hat{p}(\hat{x}) - p(\hat{x})] + n_{\hat{x}}^{1/2}[p(\hat{x}) - 0.5] \geq -0.5 z_{1-\delta}\} \to 0$$

for all $\hat{x} \notin S(x)$. This follows from the fact that for all $\hat{x}$, $n_{\hat{x}}^{1/2} [\hat{p}(\hat{x}) - p(\hat{x})]$ converges in distribution to a normal distribution as $n \to +\infty$. By a similar argument and the condition (iii) in *RG*, we have $\Pr\{n_{\hat{x}}^{1/2}[\hat{p}(\hat{x}) - 0.5] < -0.5 z_{1-\delta}\} \to 0$ for all $\hat{x} \in S(x)$. Hence the r.h.s. of (40) converges to 0 as $n \to +\infty$. $\square$

Lemma A3 For any $x \in \Omega_X$ *and* $\alpha \in (0,1)$, $\hat{Q}_{Y|D=1,x}^{\hat{\alpha}} \xrightarrow{p} Q_{Y|D=1,x}^{\alpha}$ *whenever* $\hat{\alpha} \xrightarrow{p} \alpha$.

Proof of Lemma A3. Let $\hat{F}_{Y|D=1,x}$ denote the empirical distribution of Y given $D = 1$ and x. Suppose $\exists \varepsilon > 0$ such that $|\hat{Q}_{Y|D=1,x}^{\hat{\alpha}} - Q_{Y|D=1,x}^{\alpha}| > \varepsilon$. Then at least one of the following two conditions must hold: (a) $\exists \varepsilon_1 > 0$ such that $|\hat{\alpha} - \alpha| > \varepsilon_1$; or (b) $\exists \varepsilon_2 > 0$ such that $\sup_{t \in \mathbb{R}} |\hat{F}_{Y|D=1,x}(t) - F_{Y|D=1,x}(t)| > \varepsilon_2$. To see this, suppose $\hat{\alpha} = \alpha$ and $|\hat{t}_{\alpha,x} - t_{\alpha,x}| > \varepsilon$ where $\hat{t}_{\alpha,x}$, $t_{\alpha,x}$ are shorthands for $\hat{Q}_{Y|D=1,x}^{\alpha}$, $Q_{Y|D=1,x}^{\alpha}$ respectively. By *(ii)* in the regularity conditions *RG*, $\exists \delta > 0$ such that $|F_{Y|D=1,x}(\hat{t}_{\alpha,x}) - F_{Y|D=1,x}(t_{\alpha,x})| > \delta$. By construction $\hat{F}_{Y|D=1,x}(\hat{t}_{\alpha,x}) = F_{Y|D=1,x}(t_{\alpha,x}) = \alpha$, and this suggests $|F_{Y|D=1,x}(\hat{t}_{\alpha,x}) - \hat{F}_{Y|D=1,x}(\hat{t}_{\alpha,x})| > \delta$. Then (b) must hold with $\varepsilon_2 = \delta$. Hence for all $\varepsilon > 0$ and $x \in \Omega_X$,

$$\begin{aligned} & \Pr\left\{|\hat{Q}_{Y|D=1,x}^{\hat{\alpha}} - Q_{Y|D=1,x}^{\alpha}| > \varepsilon\right\} \\ \leq \ & \Pr\{|\hat{\alpha} - \alpha| > \varepsilon_1\} + \Pr\left\{\sup\nolimits_{t \in \mathbb{R}} |\hat{F}_{Y|D=1,x}(t) - F_{Y|D=1,x}(t)| > \varepsilon_2\right\} \end{aligned} \tag{41}$$

Under the condition that $\hat{\alpha} \xrightarrow{p} \alpha$, the first term on the r.h.s. converges to zero as $n \to +\infty$. Conditional on $D_{g,t} = 1$ and $X_{g,t} = x$, the observations $Y_{g,t}$ are i.i.d. and the Glivenko-Cantelli Theorem applies and $\sup_{t \in \mathbb{R}} |\hat{F}_{Y|D=1|x}(t) - F_{Y|D=1,x}(t)| \xrightarrow{a.s.} 0$ for all x. Hence the second term on the r.h.s. in (41) vanishes as $n \to +\infty$. $\square$

For each choice of α in $\mathcal{A}$ and all $x \in \Omega_X$, define

$$\hat{c}_\alpha^*(x) \equiv \frac{\sum_{\tilde{x}\in\Omega_X} \hat{c}_\alpha(x;\tilde{x}) 1\{\tilde{x} \in \hat{S}_n(x)\}}{\sum_{\tilde{x}\in\Omega_X} 1\{\tilde{x} \in \hat{S}_n(x)\}} \tag{42}$$

Lemma A4 For any $x \in \Omega_X$ and $\alpha \in \mathcal{A}$, $|\hat{c}_\alpha^*(x) - c(x)| \xrightarrow{p} 0$.

Proof of Lemma A4. Since $\hat{p}(x) \xrightarrow{p} p(x)$ for all $x \in \Omega_X$, $\hat{\alpha}' \xrightarrow{p} \alpha'$ for all $x \in \Omega_X$, $\alpha \in \mathcal{A}$ and the corresponding α'. Then by Lemma A3, we have, for any $\tilde{x} \in S(x)$,

$$\begin{aligned} \hat{c}_\alpha(x;\tilde{x}) &= \hat{Q}^{\alpha}_{Y|D=1,x} + \hat{Q}^{0.5}_{Y^c|\tilde{x}} - \hat{Q}^{\hat{\alpha}'}_{Y|D=1,\tilde{x}} \\ &\xrightarrow{p} Q^{\alpha}_{Y|D=1,x} + Q^{0.5}_{Y|\tilde{x}} - Q^{\alpha'}_{Y|D=1,\tilde{x}} = c(x) \end{aligned} \tag{43}$$

for *any* $x \in \Omega_X$, $\alpha \in \mathcal{A}$ and the corresponding α'. This follows from the Slutsky's Theorem and the fact that $\hat{Q}^{0.5}_{Y^c|\tilde{x}} \xrightarrow{p} Q^{0.5}_{Y|\tilde{x}}$ for any $\tilde{x} \in \Omega_X^+$. Note (43) holds for all x in Ω_X, including $x \in \Omega_X^+$. This is because the identification in part (iv) of Proposition 3 applies immediately for any $x \in \Omega_X^+$. Let ω_n denote the event "$\hat{S}_n(x) \subseteq S(x)$" and ω_n^c denote its complement. Then for any $\varepsilon > 0$,

$$\begin{aligned} & \Pr\{|\hat{c}_\alpha^*(x) - c(x)| > \varepsilon\} \\ = \; & \Pr\{|\hat{c}_\alpha^*(x) - c(x)| > \varepsilon \wedge \omega_n\} + \Pr\{|\hat{c}_\alpha^*(x) - c(x)| > \varepsilon \wedge \omega_n^c\} \end{aligned} \tag{44}$$

By Lemma A2, $\Pr\{\omega_n^c\} \to 0$ (and therefore the second term on the r.h.s. of (44) vanishes) as $n \to +\infty$. Note $\hat{c}_\alpha^*$ takes the form of an average and therefore the event "$|\hat{c}_\alpha^*(x) - c(x)| > \varepsilon$ and ω_n happens" implies $\max_{\tilde{x}\in S(x)} |\hat{c}_\alpha(x;\tilde{x}) - c(x)|$ must be larger than $\varepsilon^* \equiv \varepsilon/\#\{S(x)\} > 0$, where $\#\{S(x)\}$ denotes the cardinality of $S(x)$. It then follows:

$$\Pr\{|\hat{c}_\alpha^*(x) - c(x)| > \varepsilon \wedge \omega_n\} \leq \sum_{\tilde{x}\in S(x)} \Pr\{|\hat{c}_\alpha(x;\tilde{x}) - c(x)| > \varepsilon^*\} \tag{45}$$

with the r.h.s. of (45) vanishes as $n \to +\infty$ because of (43) and $\#\{S(x)\} < +\infty$. $\square$

For any $x \in \Omega_X$ and $\varepsilon > 0$, by construction,

$$\begin{aligned} & \Pr\{|\hat{c}(x) - c(x)| > \varepsilon\} \\ = \; & \Pr\left\{\left|\frac{1}{\#\mathcal{A}} \sum_{\alpha\in\mathcal{A}} \hat{c}_\alpha^*(x) - c(x)\right| > \varepsilon \wedge x \notin \hat{\Omega}_X^+\right\} + \Pr\left\{\left|\hat{Q}^{0.5}_{Y^c|x} - c(x)\right| > \varepsilon \wedge x \in \hat{\Omega}_X^+\right\} \end{aligned} \tag{46}$$

If $x \notin \Omega_X^+$, the second term on the r.h.s. of (46) vanishes as $n \to +\infty$ by Lemma A2 while the first term converges to 0 because $\hat{c}_\alpha^*(x) \xrightarrow{p} c(x)$ by Lemma A4 and $(\#\mathcal{A})^{-1} \sum_{\alpha\in\mathcal{A}} \hat{c}_\alpha^*(x) \xrightarrow{p} c(x)$ by the Slutsky's Theorem. If $x \in \Omega_X^+$, then the first term vanishes by Lemma A2 while the second term converges to 0 as $\hat{Q}^{0.5}_{Y^c|x} \xrightarrow{p} Q^{0.5}_{Y|x} = c(x)$ for all $x \in \Omega_X^+$. *Q.E.D.*

Proof of Proposition 5. A pair (c, F_ϵ) generates the outcome p in a SSPE if and only if a linear system of M equations holds:

$$Q = C - \beta\Pi \tag{47}$$

where Q is a M-vector with the m-th coordinate $Q_m \equiv F_\epsilon^{-1}(p_m)$ with $p_m \equiv p(x_m)$; and Π solves

$$\Pi = G(\beta\Pi + \Phi) \tag{48}$$

where the m-th coordinate of Φ is $\Phi_m \equiv \phi(p_m; F_\epsilon) = \int \max\{Q_m - \varepsilon, 0\}dF_\epsilon$. The outcome vector p enters (47) and (48) through Q and Φ. We adopt a pair of location and scale normalizations with $F_\epsilon^{-1}(\frac{1}{2}) = 0$ and $\phi(\frac{1}{2}; F_\epsilon) = \phi$ for some constant $\phi > 0$. Thus c is conditional median of cake sizes under this location normalization. As in the text, we use $R_{(j)}$ to denote the j-th smallest coordinate in a generic vector R. Proof of Proposition 5 uses Lemma A5 below.

Lemma A5 Suppose CI, AS hold and ϵ is independent from X. A vector p can be rationalized in a bargaining game with discount factor β, state transition G, and (c, F_ϵ) such that $BC > 0$ and F_ϵ is increasing on $\mathbb{R}^1$ with $F_\epsilon^{-1}(0.5) = 0$ if and only if there exists $R^Q \in \mathbb{R}^M$, $R^\Phi \in \mathbb{R}^M_{++}$ such that:

$$A\left[R^Q + \beta(I - \beta G)^{-1}GR^\Phi\right] > 0 \tag{49}$$

$$\hat{R}^Q_m \leq \bar{R}^Q_n \Leftrightarrow p^*_m \leq p^*_n,\ \forall m, n \in \{1, ., M\} \tag{50}$$

$$p^*_{(m)}(\hat{R}^Q_{(m+1)} - \hat{R}^Q_{(m)}) \leq \hat{R}^\Phi_{(m+1)} - \hat{R}^\Phi_{(m)} \leq p^*_{(m+1)}(\hat{R}^Q_{(m+1)} - \hat{R}^Q_{(m)}),\ \forall m \in \{1, ., M\} \tag{51}$$

with $p^* \equiv [p, \frac{1}{2}]$, $\hat{R}^Q \equiv [R^Q, 0]$ and $\hat{R}^\Phi \equiv [R^\Phi, \phi]$.

Proof of Lemma A5. (*Necessity*) Suppose p is rationalized by some (c, F_ϵ) with $BC > 0$ and ϵ is independent of X with median 0. Then let $Q_m = F_\epsilon^{-1}(p_m) \equiv Q_\epsilon(p_m)$ and $\Phi_m = \phi(p_m; F_\epsilon)$. It follows from the substitution of (48) into (47), $\epsilon \perp X$ and monotonicity of F_ϵ on $\mathbb{R}^1$ that (49) and (50) must hold for Q, Φ. The definition of ϕ and an application of the Leibniz rule for differentiating integrals suggest for any m, n,

$$\phi(p_m) - \phi(p_n) = \int_{Q_n}^{Q_m} F_\epsilon(\varepsilon)d\varepsilon$$

which is bounded between $p_n(Q_m - Q_n)$ and $p_m(Q_m - Q_n)$. Hence (51) holds for Q, Φ with $\phi \equiv \phi(1/2; F_\epsilon)$. More generally, if the normalization uses certain constant ϕ such that $\phi \neq \phi(1/2; F_\epsilon)$, the system linear restrictions (49)-(51) still hold for the scale multiples $Q \equiv \frac{\phi}{\phi(0.5;F_\epsilon)}[Q_\epsilon(p_1), ., Q_\epsilon(p_M)]$ and $\Phi \equiv \frac{\phi}{\phi(0.5;F_\epsilon)}[\phi(p_1; F_\epsilon), ., \phi(p_M; F_\epsilon)]$ as $\phi/\phi(0.5; F_\epsilon) > 0$. Hence (49)-(51) hold with $(R^Q, R^\Phi) = (Q, \Phi)$.

(Sufficiency) We need to show that if (49)-(51) holds for some R^Q, R^Φ then there must be a pair (c, F_ϵ) such that $\epsilon \perp X$, F_ϵ is increasing on $\mathbb{R}^1$, the shape restrictions $BC > 0$ are satisfied, and (c, F_ϵ) rationalizes p as the probability for agreements in SSPE. We can construct such a F_ϵ by first setting its p_m-percentile $Q_\epsilon(p_m)$ equal to Q_m, $Q_\epsilon(0.5) = 0$ and $\phi(1/2; F_\epsilon)$ equal to the positive constant ϕ, and then interpolating between $Q_\epsilon(p_m)$ so that $\phi(p(x_m))$ is equal to Φ_m. This is possible because inequality restrictions (51) are satisfied. An unobserved state distribution constructed this way is independent from X and increasing over $\mathbb{R}^1$ with $Q_\epsilon(0.5) = 0$. Then define $C = Q + \beta(I - \beta G)^{-1}G\Phi$ and the pair (c, F_ϵ) satisfies the restrictions on c, F_ϵ by construction. Furthermore (c, F_ϵ) also rationalizes p by construction because there exists Π such that (47) and (48) hold jointly. $\square$

In both types of counterfactual analyses considered, the distribution F_ϵ is fixed in the data-generating process and the counterfactual context. When C is fixed while G changed to $\tilde{G}$, we have $C = Q + \beta(I - \beta G)^{-1}G\Phi = \tilde{Q} + \beta(I - \beta\tilde{G})^{-1}\tilde{G}\tilde{\Phi}$. When G is fixed while C changed to $\tilde{C} = \gamma C$, we have $C = Q + \beta(I - \beta G)^{-1}G\Phi$ from the data-generating process and $\gamma C = \tilde{Q} + \beta(I - \beta\tilde{G})^{-1}\tilde{G}\tilde{\Phi}$ in the counterfactual context. The rest of the proof follows from arguments in Lemma A5. *Q.E.D.*

Appendix B: Choice of $F_{\epsilon|X}$ and Counterfactuals

Part *(ii)* of Proposition 1 suggests c_0 and $F^0_{\epsilon|X}$ cannot be jointly identified even with β_0 identified and considered known. Thus it is tempting to think setting $F^0_{\epsilon|x}$ to some known distribution (say, uniform on $[0, 1]$ for all x) in estimation is a necessary normalization for structural estimations. Nonetheless, such an arbitrary choice of the unobserved state distribution can lead to errors in predicting counterfactual distributions of (X, Y) if the cake function (mapping from states to surplus) is changed. Below we show the only special case where such a choice does not preclude correct counterfactual analyses is when $\epsilon \perp X$.

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

$$\tilde{c}(x, \tilde{\varepsilon}) = Q_{Y|x}(\tilde{F}_{\epsilon|x}(\tilde{\varepsilon})) \tag{52}$$

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

$$\tilde{c}(x, \tilde{\varepsilon}) = c_0(x, Q^0_{\epsilon|x}(\tilde{F}_{\epsilon|X=x}(\tilde{\varepsilon}))) \tag{53}$$

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

$$\tilde{c}^{-1}(x,y) = \tilde{Q}_{\epsilon|x}(F^0_{\epsilon|x}(c_0^{-1}(x,y))) \tag{54}$$

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 $\tilde{Q}_{\epsilon|x}$ is the inverse of $\tilde{F}_{\epsilon|x}$. 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 $F_{Y|x}$ as in (52), 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 $\Pr\{c_0(g(X),\epsilon) \leq y|\ X = x;\ F^0_{\epsilon|X}\} = F^0_{\epsilon|x}(c_0^{-1}(g(x),y))$, while the one predicted under the normalization is:

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

where $f \circ g(.)$ is a shorthand for the composite function $f(g(.))$, and the second equality follows from (54). In general, $\tilde{F}_{\epsilon|x} \circ \tilde{Q}_{\epsilon|X=g(x)} \circ F^0_{\epsilon|X=g(x)} \neq F^0_{\epsilon|x}$, 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] Amemiya, T. (2005): Advanced Econometrics, Cambridge: Harvard University Press.

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

[3] Athey, S. and P. Haile (2007): "Nonparametric Approaches to Auctions," in J. Heckman and E. Leamer (eds.), Handbook of Econometrics, Vol. 6A, ch. 60, pp. 3847–3965, Amsterdam: Elsevier.

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

[5] Berry, S. and E. Tamer (2006): "Identification in Models of Oligopoly Entry," in R. Blundell, W. Newey and T. Persson (eds.), Advances in Economics and Econometrics: Ninth World Congress of the Econometric Society, Vol. 2, pp. 46-85, Cambridge: Cambridge University Press.

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

[7] Bierens, H. (1987): "Kernel Estimators of Regression Functions," in T. F. Bewley (ed.), Advances in Econometrics: Fifth World Congress of the Econometric Society, Vol. I, pp. 99-144, New York: Cambridge University Press.

[8] Binmore, K., Osborne, M. and A. Rubinstein (1992): "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.

[9] Chen, S., Dahl, G., and S. Khan (2005): "Nonparametric Identification and Estimation of a Censored Location-Scale Regression Model," Journal of the American Statistical Association, 100, 212-221.

[10] Chiappori, P. and O. Donni (2006): "Learning from a Piece of Pie: The Empirical Content of Nash Bargaining," Working Paper 2128, IZA.

[11] Diaz-Moreno, C. and J. Galdon-Sanchez (2000): "Collective Bargaining Under Complete Information," Working Paper TE/2000/401, LSE STICERD.

[12] Diermeier, D., Eraslan, H. and A. Merlo (2003): "A Structural Model of Government Formation," Econometrica, 71, 27-70.

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

[14] Ghosal, S., Miller, M. and K. Thampanishvong (2010): "Delay and Haircuts in Sovereign Debt: Recovery and Sustainability," Working Paper, University of Warwick.

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

[16] Haile, P. and E. Tamer (2003): "Inference with an Incomplete Model of English Auctions," Journal of Political Economy, 111, 1-51.

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

[18] 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

[19] Laffont, J.J. and Q. Vuong (1996): "Structural Analysis of Auction Data," American Economic Review, 36, 414-420.

[20] Matzkin, R. (2003): "Nonparametric Estimation of Nonadditive Random Functions," Econometrica, 71, 1339-1375.

[21] Merlo, A. (1997), "Bargaining over Governments in a Stochastic Environment," Journal of Political Economy, 105, 101-131.

[22] Merlo, A., Ortalo-Magne, F. and J. Rust (2009), "The Home Selling Problem: Theory and Evidence," Working Paper, University of Maryland.

[23] Merlo, A. and C. Wilson (1995): "A Stochastic Model of Sequential Bargaining with Complete Information," Econometrica, 63, 371-399.

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

[25] Muthoo, A. (1999): Bargaining Theory with Applications, Cambridge: Cambridge University Press.

[26] Norets, A. and X. Tang (2010), "Semiparametric Inference in Dynamic Binary Choice Models," Working Paper, Princeton University.

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

[28] Rubinstein, A. (1982): "Perfect Equilibrium in a Bargaining Model," Econometrica, 50, 97-109.

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

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

[31] Simcoe, T. (2008): "Standard Setting Committees," Working Paper, University of Toronto.

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

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