Nash bargaining with endogenous outside options*

Eduard Talamàs†

Abstract

This paper describes a non-cooperative theory of bargaining in stationary markets that provides a generalization of the Nash bargaining solution to coalition formation problems. This generalization prevents outside options from being determined in a circular way and, remarkably, always yields a unique prediction. The equilibrium uncovers an endogenous vertical market structure such that the effects of changes in fundamentals propagate—via outside options—from the top down, but not vice versa. In markets that are vertically differentiated by skill, changes at the bottom do not affect those at the top. Moreover, if there is positive assortative matching, changes at the top affect those at the bottom only if they also affect everyone in between.

1 Introduction

In many markets, agents simultaneously bargain over both which coalitions to form (e.g., which firms employ which workers, which entrepreneurs become partners, which businesses form strategic alliances, etc.) and how to share the resulting gains from trade (e.g., wages, equity shares, etc.), but predictions that are general, sharp and transparent are elusive in these settings. In this paper, I describe a non-cooperative theory of coalition formation in stationary markets that maintains the transparency of the Nash bargaining solution and provides sharp predictions about how various changes in economic fundamentals propagate through these markets via agents' outside options.

**Date printed:* November 17, 2020.

†IESE Business School and Cambridge University. The guidance of Benjamin Golub throughout the process of conducting, presenting, and writing this work has been essential. I am grateful to Matthew Elliott, Jerry Green, Pau Milán, Debraj Ray, Rakesh Vohra, Asher Wolinsky and the anonymous referees for detailed comments, as well as to numerous conference and seminar participants for useful feedback. This work has been supported by the Warren Center for Network & Data Sciences, the Rockefeller Foundation (#2017PRE301) and Cambridge INET. All errors are my own.

This paper is related to Binmore, Rubinstein, and Wolinsky (1986), who describe a non-cooperative theory of bargaining in a fixed coalition to investigate how exogenous outside options enter the Nash bargaining solution.[1] They derive a version of the "outside option principle" (e.g., Sutton 1986), by which the surplus in the coalition of interest is shared according to the Nash bargaining solution, with the *Nash threat points* corresponding to the utilities that the agents get in autarky, and the *outside options* entering as lower bounds on payoffs. For example, consider a worker and an employer (both risk neutral) that can generate 1 unit of surplus by matching. The worker can sell her labor elsewhere at wage $w < 1$, the employer can hire an equally valuable worker at wage $w' > w$, and neither the employer nor the worker in autarky generate any value. The outside option principle suggests that the employer hires the worker at wage $1/2$ (as specified by the Nash bargaining solution with the Nash threat points corresponding to autarky), unless $w > 1/2$ or $w' < 1/2$ (in which case it suggests that the employer hires the worker at wage w and w', respectively). Intuitively, an agent's outside option is only relevant if it is better than what the Nash bargaining solution would otherwise give her.[2]

Crucially, however, the outside option principle is silent about how the relevant outside options in each coalition are determined. For instance, in the example just described, the wages w and w' at which the worker and the employer, respectively, can match elsewhere are taken as given. But, in many cases, these wages are themselves the result of bargaining with third parties. From this perspective, the contribution of this paper is to provide a non-cooperative framework that shows not only how outside options enter the Nash bargaining solution, but also how the Nash bargaining solution pins down the relevant outside options in each coalition.

The main innovation of the model relative to the literature on non-cooperative bargaining in stationary markets is that it allows agents to strategically choose which coalitions to propose—instead of assuming that the coalitions that can form are determined at random in

[1] In many applications, there are various sensible alternatives for both what the relevant outside options are and how they enter the Nash bargaining solution—and different alternatives have qualitatively different implications. For example, the extent to which unemployment is a relevant outside option in wage bargaining determines the effects of unemployment insurance on the labor market—e.g., Pissarides (2000), Krusell et al. (2010), Hagedorn et al. (2013) and Chodorow-Reich et al. (2018)—and the ability of macroeconomic models to generate realistic employment fluctuations—e.g., Shimer (2005), Hall and Milgrom (2008), Sorkin (2015), Chodorow-Reich and Karabarbounis (2016), Hall (2017) and Ljungqvist and Sargent (2017).

[2] Binmore et al. (1989) provide experimental evidence that is consistent with the outside option principle. More recently, Jäger et al. (2020) find that real-world wages are insensitive to sharp increases in unemployment insurance benefits, which suggests that unemployment is not a credible outside option in wage bargaining.

each period.[3] This seemingly-small departure has qualitative effects on the equilibrium predictions, and provides three main advantages with respect to the standard random matching framework.[4]

The first advantage of this framework is its generality: Not only it allows agents to be strategic about which coalitions to propose, but it also admits arbitrary heterogeneity in terms of coalitional surpluses, preferences and proposer probabilities. In particular, by imposing no restrictions on the coalitional surpluses, this framework is more general than the ones in Manea (2011), Polanski and Vega-Redondo (2018) and Talamàs (2019). By allowing heterogeneity in preferences, it is also more general than the one in Nguyen (2015).

The second advantage is that this framework allows a natural characterization of the equilibrium in terms of the familiar equilibria of canonical pure-sharing bargaining games (e.g., Binmore, Rubinstein, and Wolinsky 1986). In particular, the main result of this paper is that the equilibrium payoff profile is *the only one* that satisfies the following credibility property: Each agent's payoff is the maximum that she can justify as resulting from bargaining in some coalition in isolation while honoring the others' outside options—given by what they can themselves justify in this way. This credibility property prevents outside options from being determined in a circular way and, remarkably, it uniquely pins them down.

The third advantage is that a transparent algorithm constructs the equilibrium and provides sharp comparative statics. In particular, this algorithm uncovers an endogenous vertical market structure with the property that the effects of small changes in fundamentals propagate—via the endogenous outside options—from the top down, but not vice versa. For example, in multi-sided markets in which agents are vertically differentiated by skill, changes in the skills at the bottom of one side of the market do not affect those at the top. As in the canonical marriage market model of Becker (1973), positive assortative matching arises if and only if skills are complementary. But, in contrast to Becker's framework (and much of the subsequent literature), the present theory pins down wages uniquely, and hence provides sharp predictions about how positive assortative matching affects the way in which changes in economic fundamentals propagate via outside options. For example, positive assortative matching implies that changes in the skills at the top of one side of the market cannot affect those at the bottom unless they also affect everyone in between.

[3]The exception is Talamàs (2019), where I analyze a special case of this framework tailored to networked buyer-seller markets. As I discuss in detail in subsubsection 6.1.1, both the framework and the characterization of the present paper are substantially more general than the ones in Talamàs (2019).

[4]I illustrate via a simple example in subsubsection 6.1.1 how allowing agents to choose which coalitions to propose makes a substantial difference to the equilibrium outcome even in the limit as frictions vanish.

Roadmap. Section 2 presents a canonical non-cooperative model of bargaining in one coalition with exogenous outside options—similar to the one in Binmore, Rubinstein, and Wolinsky (1986). Section 3 builds on this classical framework to characterize the equilibrium of the coalition formation game of interest, and section 4 characterizes the limit of this equilibrium as bargaining frictions vanish. Section 5 describes comparative statics in vertically differentiated markets. Finally, section 6 discusses the main connections with the existing literature and how—under a standard interpretation of the stationary market assumption—the equilibrium outcome in the limit as bargaining frictions vanish is efficient.

2 Preliminaries

There is a finite set N of agents. *Coalitions* are subsets of N. *Profiles* are elements in $\mathbb{R}^N_{\geq 0}$. The i^{th} element of a profile α is denoted by α_i. For any two profiles α and β, $\alpha \geq \beta$ indicates that $\alpha_i \geq \beta_i$ for every i in N. Finally, α_{-i} denotes the profile α after setting its i^{th} entry to 0.

2.1 Auxiliary bargaining game $\mathcal{G}(C, \theta)$

The game $\mathcal{G}(C, \theta)$ involves bargaining in a given coalition C with outside options given by an exogenous profile θ. This game generalizes the model with exogenous risk of breakdown in Binmore, Rubinstein, and Wolinsky (1986) by allowing the relevant coalition C to be of arbitrary size.[5] In each period $t = 1, 2, \ldots$, one of the agents is chosen to be the proposer. Agent i is chosen with probability p_i. If this agent is not in C, everyone waits for the next period. Otherwise, she chooses a profile that splits the surplus $y(C) > 0$ among the agents in C (i.e., a profile x with $\sum_{i \in C} x_i = y(C)$). The agents in C then decide in a pre-specified order whether to accept or reject. If all of them accept, they leave the market with their agreed shares. Otherwise, no one matches in this period. Every agent i that rejects an offer can choose to either wait for the next period or to leave the market with payoff θ_i. The bargaining friction is that, at the end of each period, the game ends exogenously with probability $q > 0$, in which case every agent i obtains her autarky surplus $y(i) > 0$. The preferences of each agent i are represented by the von Neumann-Morgenstern utility function u_i, which is a concave, strictly increasing, and twice-continuously differentiable function of her payoff.

[5] As in Binmore et al. (1986), the results of the present paper have analogous counterparts when the relevant bargaining friction is impatience instead of risk aversion.

2.2 Equilibrium in $\mathcal{G}(C,\theta)$

Throughout this paper, I focus on subgame-perfect equilibria that are stationary, in the sense that no agent's strategy conditions on the history of the game except—in the case of a response—on the going proposal. Clearly, every stationary equilibrium is in cutoff strategies: Each agent i in C accepts (rejects) all offers that give her strictly more (less) than a cutoff α_i.

Lemma 2.1. *The game $\mathcal{G}(C,\theta)$ admits unique equilibrium cutoffs.*

Proof. If $(\alpha_i)_{i\in C}$ are equilibrium cutoffs in $\mathcal{G}(C,\theta)$, then, for each i in C,

$$(1)\quad u_i(\alpha_i) = \max\left[u_i(\theta_i), qu_i(y(i)) + (1-q)\left(p_i u_i(\max(\alpha_i, y(C) - \sum_{j\in C-i}\alpha_j)) + (1-p_i)u_i(\alpha_i)\right)\right].$$

In words, each agent i in C is indifferent between accepting and rejecting an offer that gives her α_i. When agent i rejects an offer, she can choose to obtain θ_i right away, or to incur the risk of breakdown (which materializes with probability q, in which case she gets $y(i)$). If this risk does not materialize, in the next period she is the proposer with probability p_i, in which case she can obtain $y(C) - \sum_{j\in C-i}\alpha_j$ by making an acceptable proposal.[6] With the remaining probability $1-p_i$, she either receives an offer that gives her α_i, or she does not receive any offer; in either case, her expected utility is $u_i(\alpha_i)$. Lemma 2.1 follows from the fact that, as shown in Appendix A.1, there is a unique $(\alpha_i)_{i\in C}$ that satisfies (1) for all i in C. □

While the agents in $N-C$ do not play any role in $\mathcal{G}(C,\theta)$, and hence have no equilibrium cutoffs in this game, it is useful to define artificial cutoffs for these agents so that the vector of equilibrium cutoffs $\alpha(C,\theta)$ in $\mathcal{G}(C,\theta)$ has the same dimension for all pairs (C,θ).

Notation 2.1. For every agent i in N, let $\alpha_i(C,\theta)$ be i's equilibrium cutoff in $\mathcal{G}(C,\theta)$ if i is in C, and 0 otherwise.

2.3 Five auxiliary results

The following five simple corollaries are useful for the arguments below. The first three follow from the fact that—as is evident from (1)—the outside option θ_i of each agent i in C enters only as a lower bound on her equilibrium cutoff $\alpha_i(C,\theta)$.

[6] If $\alpha_i \geq y(C) - \sum_{j\in C-i}\alpha_j$, agent i does not benefit from making an acceptable proposal. Otherwise, agent i obtains exactly $y(C) - \sum_{j\in C-i}\alpha_j$ when she is the proposer because, in this case, every agent j in $C-i$ accepts α_j (otherwise, i would not have a best response).

The first observation highlights that an agent's equilibrium cutoff is weakly decreasing in the others' outside options.

Corollary 2.2. *For any profile $\theta' \leq \theta$ with $\theta'_i = \theta_i$, $\alpha_i(C, \theta') \geq \alpha_i(C, \theta)$.*

Similarly, Corollary 2.3 highlights that an agent's equilibrium cutoff is weakly increasing in her own outside option.

Corollary 2.3. *For any profile $\theta' \leq \theta$ with $\theta'_j = \theta_j$ for all $j \neq i$, $\alpha_i(C, \theta') \leq \alpha_i(C, \theta)$.*

The following observation highlights that the equilibrium payoffs in $\mathcal{G}(C, \theta)$ do not change if we replace agents' outside options θ by their equilibrium cutoffs $\alpha(C, \theta)$.

Corollary 2.4. *For any profile θ' such that, for all i, θ'_i is either θ_i or $\alpha_i(C, \theta)$, $\alpha(C, \theta) = \alpha(C, \theta')$.*

The following property of the equilibrium sharing rule $\alpha(C, \theta)$ plays a key role in ensuring that the credibility notion uncovered by this paper uniquely pins down equilibrium outcomes in the coalition formation game of interest (described in the next section).

Corollary 2.5. *For any outside option profile θ and any two coalitions C and D containing any two agents i and j, if $\alpha_i(C, \theta) \geq \alpha_i(D, \theta)$, then $\alpha_j(C, \theta) \geq \alpha_j(D, \theta)$ as well.*

Proof. See Appendix A.2. □

Finally, the last observation is that, as long as an agent benefits from bargaining to form coalition C, her equilibrium cutoff is strictly increasing in the surplus $y(C)$ of this coalition.

Corollary 2.6. *If $\alpha_i(C, \theta) > \max(\theta_i, y(i))$, then $\alpha_i(C, \theta)$ strictly increases with $y(C)$.*

Proof. See Appendix A.3. □

3 Coalition formation in a stationary market

3.1 Bargaining game $\mathcal{G}$

The set N of agents, their utility functions $\{u_i\}_{i \in N}$ and their proposer probabilities $(p_i)_{i \in N}$ are as in $\mathcal{G}(C, \theta)$. The game $\mathcal{G}$ differs from $\mathcal{G}(C, \theta)$ in three respects. First, agents can choose which coalition to propose. Second, they do not have exogenous outside options. Third, dynamic entry of agents into the market keeps the market stationary.

In each period $t = 1, 2, \ldots$, one agent is selected to be the proposer as in $\mathcal{G}(C, \theta)$. The proposer chooses a coalition $C \subseteq N$ that contains her, and proposes a split of its surplus $y(C) > 0$ among its members. The agents in C then decide in a pre-specified order whether to accept or reject this proposal. If all of them accept, then they leave the market with the agreed shares, while the others wait for the next period. Otherwise, no one matches in this period, and everyone waits for the next period. In the transition between periods, every agent i waiting for the next period has a probability q of being forced to leave the market, in which case she obtains her autarky payoff $y(i)$. For simplicity, the market is kept stationary by assuming that every agent that leaves the market is immediately replaced by an identical agent. I discuss a natural interpretation of this stationarity assumption in section 6.2.

3.2 Credibility and equilibrium

As in $\mathcal{G}(C, \theta)$, every stationary equilibrium is in cutoff strategies: Each agent i accepts (rejects) all offers that give her strictly more (less) than a certain cutoff β_i. The main result of this paper is that the unique equilibrium cutoff profile in $\mathcal{G}$ is the only profile that is *credible*, in the following sense.

Definition 3.1. A profile β is *credible* if

(2) $$\beta_i = \max_{i \in C \subseteq N} \alpha_i(C, \beta_{-i}) \text{ for all } i \text{ in } N,$$

where recall that β_{-i} denotes β after setting its i^{th} entry to 0.

In words, a profile β is credible if each agent i's cutoff β_i is the maximum that she can justify by bargaining in some coalition C while honoring *the others'* outside options β_{-i}. Theorem 1 shows that there is a unique credible profile, and Theorem 2 shows that this is the unique equilibrium cutoff profile in $\mathcal{G}$. Algorithm $\mathcal{A}$ (Definition 3.3) constructs this profile. Informally, the first step of this algorithm is as follows. Start with the outside option profile 0. Have each agent i point to all the coalitions C where her share $\alpha_i(C, 0)$ is maximum. By Corollary 2.5, there is at least one coalition that is pointed at by all its members. I refer to such coalitions as the 0-*perfect coalitions*. For every 0-perfect coalition C, update the outside option of each agent i in C to $\alpha_i(C, 0)$. In the next steps, the algorithm similarly updates the remaining agents' outside options.

Definition 3.2. For any profile θ, a coalition C is θ-*perfect* if, for all i in C,

$$\alpha_i(C, \theta) = \max_{i \in D \subseteq N} \alpha_i(D, \theta).$$

Definition 3.3 (Algorithm $\mathcal{A}$). Let $\gamma^0 = 0 \in \mathbb{R}^N$. In step $s = 1, 2, \dots$, for each agent i in a γ^{s-1}-perfect coalition C, let $\gamma_i^s = \alpha_i(C, \gamma^{s-1})$. For all other agents j, let $\gamma_j^s = 0$. End in the first step S such that $\gamma^S = \gamma^{S-1}$, and denote γ^S by γ.

Note that γ^s is well defined in every step s, because $\alpha_i(C, \gamma^{s-1})$ is the same for all γ^{s-1}-perfect coalitions C that contain i. Moreover, $\gamma^{s-1} \leq \gamma^s$, because for any γ^{s-1}-perfect coalition C, Corollary 2.4 implies that $\alpha(C, \gamma^{s-1}) = \alpha(C, \gamma^s)$. Hence, by Corollary 2.2, every γ^{s-1}-perfect coalition is also a γ^s-perfect coalition, so algorithm $\mathcal{A}$ updates each agent's outside option (from 0 to a strictly positive number) at most once.[7] Finally, Corollary 2.5 implies that in the step S in which algorithm $\mathcal{A}$ ends, every agent is in a γ^S-perfect coalition, so algorithm $\mathcal{A}$ updates each agent's outside option exactly once.

Theorem 1. *The profile γ constructed by algorithm $\mathcal{A}$ is the unique credible profile.*

Proof. Let's first show that γ is credible. Consider an arbitrary agent i. Note that $\gamma_i = \max_{i \in C \subseteq N} \alpha_i(C, \gamma) \geq \max_{i \in C \subseteq N} \alpha_i(C, \gamma_{-i})$, where the equality holds by construction and the inequality follows from Corollary 2.3. To see that this inequality cannot be strict, note that letting s be the first step of algorithm $\mathcal{A}$ in which agent i is in a γ^{s-1}-perfect coalition C, we have that $\gamma_i = \alpha_i(C, \gamma^{s-1}) = \alpha_i(C, \gamma_{-i})$, where the second equality follows from Corollary 2.4.

To show that γ is in fact the only credible profile, fix an arbitrary credible profile β. It is enough to show that, for each $0 \leq s \leq S$ and each agent i, γ_i^s is either 0 or β_i. This holds trivially for $s = 0$. The induction hypothesis is that this holds for a given $s \geq 0$. Consider an arbitrary γ^s-perfect coalition C one of whose members i is such that $\gamma_i^s = 0$. It is enough to show that $\alpha_i(C, \gamma^s) = \beta_i$. By the induction hypothesis, for all j in C with $\gamma_j^s \neq 0$, we have that $\beta_j = \gamma_j^s = \alpha_j(C, \gamma^s)$, so $\beta_{-i} \geq \gamma^s$. Hence, using Corollary 2.2,

$$\alpha_i(C, \gamma^s) = \max_{i \in D \subseteq N} \alpha_i(D, \gamma^s) \geq \max_{i \in D \subseteq N} \alpha_i(D, \beta_{-i}) =: \beta_i.$$

Since i is chosen arbitrarily among the agents in $\{j \in C \mid \gamma_j^s = 0\}$, this implies that $\alpha_j(C, \gamma^s) \geq \beta_j$ for all j in C. Hence, $\alpha_i(C, \gamma^s) = \alpha_i(C, \alpha_{-i}(C, \gamma^s)) \leq \alpha_i(C, \beta_{-i}) = \beta_i$, where the first equality follows from Corollary 2.4, the inequality follows from Corollary 2.2, and the last equality follows from the fact that β is credible. □

Theorem 2. *The unique credible profile γ is the unique equilibrium cutoff profile in $\mathcal{G}$.*

[7] Note that for every agent i in a γ^{s-1}-perfect coalition C, we have that $\alpha_i(C, \gamma^{s-1}) \geq y(i) > 0$, so $\gamma_i^s = 0$ if and only if i is not in any γ^{s-1}-perfect coalition.

Proof. An equilibrium cutoff profile β of $\mathcal{G}$ is such that each agent i is indifferent between obtaining β_i right away and waiting for the next period; that is,

$$u_i(\beta_i) = qu_i(y(i)) + (1-q)\left[p_i u_i\left(\max_{i\in C\subseteq N}\left[y(C) - \sum_{j\in C-i}\beta_j\right]\right) + (1-p_i)u_i(\beta_i)\right] \text{ for all } i \text{ in } N, \tag{3}$$

where we have used that, conditional on the others' cutoffs β_{-i}, the maximum agent i can obtain when she is the proposer is $\max_{i\in C\subseteq N}\left[y(C) - \sum_{j\in C-i}\beta_j\right]$. Theorem 2 follows from the fact that, as shown in Appendix A.4, γ is the only profile that satisfies (3) for all i in N. □

Theorem 1 and Theorem 2 are useful mainly for two reasons. First, they show that the equilibrium predictions of the coalition formation game $\mathcal{G}$ are the necessary result of taking the equilibrium prediction in the canonical game $\mathcal{G}(C, \theta)$ as the relevant bargaining solution in each coalition C and imposing a simple credibility condition on the relevant outside options in each coalition. Second, algorithm $\mathcal{A}$ provides a transparent way to compute the equilibrium predictions which—as the analysis that follows illustrates—is very useful to understand these predictions and how they change with the economic fundamentals.

3.3 Essentially unique equilibrium

Theorem 2 implies that each agent i's equilibrium cutoff γ_i is the maximum that she can justify as resulting from bargaining in some coalition C in isolation, while honoring the others' outside options γ_{-i}, in the sense that $\gamma_i = \max_{i\in C\subseteq N}\alpha_i(C, \gamma_{-i})$. I refer to a coalition C in which i can justify her payoff γ_i in this way as *i's best coalition*.

Definition 3.4. Coalition C is *i's best coalition* if $\gamma_i = \alpha_i(C, \gamma_{-i})$.

Consider an agent i that strictly benefits from matching with others in equilibrium, in the sense that γ_i is strictly larger than her autarky payoff $y(i)$. It is clear from (1) that C is i's best coalition if and only if it solves $\max_{i\in C\subseteq N}\left(y(C) - \sum_{j\in C-i}\gamma_j\right)$. Hence, on the equilibrium path, i always proposes that one of her best coalitions forms, offers γ_j to each other member j of this coalition, and everyone accepts. In particular, a coalition that is no one's best coalition never forms in equilibrium. If i has more than one best coalition, the equilibrium does not pin down which among her best coalitions she proposes, but this is generically not the case.

Proposition 3.1. *Generically, each agent i with $\gamma_i > y(i)$ has a unique best coalition.*

Proof. Corollary 2.6 implies that in the first step s of algorithm $\mathcal{A}$ in which an agent i with $\gamma_i > y(i)$ is in a γ^{s-1}-perfect coalition, she is generically in only one such coalition C. In particular, $\gamma_i = \alpha_i(C, \gamma^{s-1}) > \alpha_i(D, \gamma^{s-1}) \geq \alpha_i(D, \gamma_{-i})$, for every coalition $D \neq C$ with i in D, where the last inequality follows from Corollary 2.2. $\square$

Hence, the game $\mathcal{G}$ has a generically unique stationary equilibrium, except for the following two possible sources of minor indeterminacies. First, each agent i is indifferent between accepting and rejecting an offer that gives her exactly her cutoff γ_i. While each agent accepts such offers with probability 1 on the equilibrium path, the responses to such offers are not pinned down off the equilibrium path. Second, there might be agents that cannot benefit from matching with others in equilibrium. The equilibrium cutoff of any such agent i is her autarky payoff $y(i)$ and her proposals are not pinned down in equilibrium , because she does not benefit from making acceptable offers.

3.4 Equilibrium tier structure

For simplicity, I focus on the generic case in which every agent i has a unique best coalition. The bargaining outcome in i's best coalition C generates her binding outside option in every other coalition D that she is part of, in the sense that, if D forms in equilibrium, agent i obtains γ_i in this coalition by appealing to her outside option $\alpha_i(C, \gamma_{-i}) > \alpha_i(D, \gamma_{-i})$.

In principle, there could exist two coalitions C and D such that the bargaining outcome in C generates a binding outside option in D, and vice versa. But Corollary 2.5 implies that this does not occur in equilibrium. In particular, for any two coalitions C and D, and any pair of agents i and j in $C \cap D$, if C is i's best coalition, then D is not j's best coalition. More generally, the coalitions that form in equilibrium can be organized into tiers in such a way that the binding outside options in every coalition are exclusively generated by bargaining in coalitions that are in higher tiers. The *first-tier coalitions* are those coalitions C with

$$\gamma_i = \alpha_i(C, 0) \text{ for each } i \text{ in } C.$$

Proceeding inductively, after identifying the k^{th}-tier coalitions for every $k \leq \ell - 1$, a coalition C is in the *ℓ^{th} tier* if it satisfies the following two properties. First, it contains at least one agent that is also in an $(\ell-1)^{\text{th}}$-tier coalition. Second, only the outside options of agents that are in a higher-tier coalition are binding; that is, for each i in C who is not a member of any coalition in the first, second, . . . , or $(\ell-1)^{\text{th}}$ tier, $\gamma_i = \alpha_i(C, \gamma_{-i})$.[8]

[8] The equilibrium tier structure can be directly derived from algorithm $\mathcal{A}$ as follows: The γ^0-perfect coalitions identified in step $s = 0$ are included in the first tier. Proceeding inductively in s, after having placed each γ^{s-1}-perfect coalition in a tier, if a γ^s-perfect coalition that is not a γ^{s-1}-perfect coalition contains an agent that

Since no one's best coalition changes after small-enough changes in the fundamentals (i.e., changes in the surplus function y, the utility functions $(u_i)_{i\in N}$ and the proposer probability profile p), the equilibrium tier structure remains intact after such changes, and uncovers sharp comparative statics. Indeed, given that the binding outside options in any given coalition are determined in higher tiers, its equilibrium sharing rule can be affected by changes in the surplus of the coalitions that are in higher tiers and the preferences or proposer probabilities of such coalitions' members—but not by changes in the productivity of other coalitions in the same or lower tiers, or the preferences or proposer probabilities of such coalitions' members. The following example (illustrated by Figure 1) describes the equilibrium tier structure and the resulting comparative statics in a simple market.

3.5 Example

Suppose that $N = \{1, 2, 3, 4\}$, with homogeneous preferences and proposer probabilities, and only three coalitions $\{1, 2\}$, $\{2, 3\}$ and $\{3, 4\}$ with non-negligible surpluses, satisfying $0 < y_{12} - y_{23} < y_{23} - y_{34}$.[9] Coalition $\{1, 2\}$ is both agent 1 and agent 2's best coalition. In particular, agents 1 and 2 always make offers to each other in equilibrium, and their equilibrium cutoffs γ_1 and γ_2 are exactly as in the bilateral bargaining game $\mathcal{G}(\{1, 2\}, 0)$. Agent 3's best coalition is $\{2, 3\}$, so her equilibrium cutoff is $\alpha_3(\{2, 3\}, \gamma_{-3})$, and she always makes offers to agent 2, obtaining $y_{23} - \gamma_2$ when she is the proposer. Finally, agent 4's best coalition is $\{3, 4\}$, so her equilibrium cutoff is $\alpha_4(\{3, 4\}, \gamma_{-4})$, and she always makes offers to agent 3, obtaining $y_{34} - \gamma_3$ when she is the proposer. Coalition $\{1, 2\}$ is in the first tier, since it is the best coalition of both its members. Agent 2's outside option binds in $\{2, 3\}$ (i.e., $\gamma_2 > \alpha_2(\{2, 3\}, \gamma_{-2})$), so this coalition belongs to the second tier. Finally, agent 3's outside option binds in $\{3, 4\}$, (i.e., $\gamma_3 > \alpha_3(\{2, 3\}, \gamma_{-3})$) so this coalition belongs to the third tier.

This example illustrates how the effects of changes in market fundamentals propagate from the top of the equilibrium tier structure down, but not vice versa. For instance, an increase in y_{12} not only benefits agents 1 and 2, but it also benefits agent 4, by harming agent 3. Indeed, an increase in y_{12} raises agent 2's outside option when bargaining with agent 3 which, in turn, reduces agent 3's outside option when bargaining with agent 4. Note also that the effects of changes in market fundamentals do not propagate from the bottom up. For example, the only effect of an increase in y_{34} is to increase agent 4's payoff γ_4.

is in a coalition that has been placed in tier k but contains no agent that is in a coalition that has been placed in any tier $k' > k$, then this coalition C is included in tier $k + 1$.

[9]For brevity, I denote $y(\{i, j\})$ by y_{ij}.

(a) An arrow from i to j indicates that agent i makes offers to agent j in equilibrium.

(b) The equilibrium tier structure.

Figure 1: Illustration of the example in section 3.5.

4 Small bargaining frictions

4.1 Auxiliary game $\mathcal{G}(C, \theta)$

Lemma 4.1 below highlights that, as the bargaining friction q goes to 0, the equilibrium cutoffs in $\mathcal{G}(C, \theta)$ converge to the Nash bargaining solution in C, with the outside options θ entering only as lower bounds on payoffs. This is a version of the *outside option principle* (e.g., Binmore, Rubinstein, and Wolinsky, 1986, Proposition 6).

Lemma 4.1. *As q goes to 0, the equilibrium cutoffs $(\alpha_i(C, \theta))_{i \in C}$ converge to the unique solution of*

$$\text{(4)} \qquad \underset{(s_i \geq y(i))_{i \in C}}{\text{argmax}} \prod_{j \in C} (u_j(s_j) - u_j(y_j))^{p_j} \quad \text{subject to } \; y(C) \geq \sum_{j \in C} s_j \; \text{ and } \; s_j \geq \theta_j \; \text{ for all } j \text{ in } C$$

if $y(C) \geq \sum_{i \in C} \max(\theta_i, y(i))$, and to $(\max(\theta_i, y(i)))_{i \in C}$ otherwise.

Proof. The proof is standard and is deferred to Appendix A.5. □

It is again useful to define artificial cutoffs for the agents in $N - C$ so that the vector of limit equilibrium cutoffs $\alpha^\star(C, \theta)$ in $\mathcal{G}(C, \theta)$ has the same dimension for all pairs (C, θ).

Notation 4.1. For each i in N, let $\alpha_i^\star(C, \theta)$ be i's limit equilibrium cutoff in $\mathcal{G}(C, \theta)$ if i is in C, and 0 otherwise.

The five auxiliary results in subsection 2.3 regarding the equilibrium cutoffs $\alpha(C, \theta)$ also hold for the limit equilibrium cutoffs $\alpha^\star(C, \theta)$. In particular, Corollary 2.5 follows from the fact that we can assign a number $\chi(C, \theta)$ to each coalition, so that each agent i's Nash bargaining share $\alpha_i^\star(C, \theta)$ in each coalition C that she is part of is increasing in $\chi(C, \theta)$.[10]

[10] The solution to (4) is such that, for all agents i in C with $\alpha_i^\star(C, \theta) > \max(\theta_i, y(i))$, $\frac{u_i(\alpha_i^\star(C,\theta)) - u_i(y(i))}{p_i u_i'(\alpha_i^\star(C,\theta))}$ is equal to a constant $\chi(C, \theta)$ and, since u is concave, $\alpha_i^\star(C, \theta)$ increases with this constant $\chi(C, \theta)$.

4.2 Coalitional bargaining game $\mathcal{G}$

This section shows that, as bargaining frictions vanish, the equilibrium cutoff profile γ converges to the unique profile that satisfies the limit version of the credibility property above, defined as follows.

Definition 4.1. A profile β in $\mathbb{R}^N$ is *Nash credible* if

$$\beta_i = \max_{i \in C \subseteq N} \alpha_i^\star(C, \beta_{-i}) \text{ for all } i \text{ in } N.$$

In other words, β is *Nash credible* if, for each agent i, β_i is the maximum that she can justify as the result of Nash bargaining in some coalition C subject to the constraint that each other agent j in C obtains at least β_j. Definition 4.3 describes the limit analog of algorithm $\mathcal{A}$, and Theorem 3 highlights that this algorithm constructs the unique Nash-credible profile.

Definition 4.2. For any profile θ, a coalition C is *θ-Nash-perfect* if, for every i in C,

$$\alpha_i^\star(C, \theta) = \max_{i \in D \subseteq N} \alpha_i^\star(D, \theta).$$

Definition 4.3 (Algorithm $\mathcal{A}^\star$)**.** Let $\gamma^0 = 0 \in \mathbb{R}^N$. In step $s = 1, 2, \ldots$, for each γ^{s-1}-Nash-perfect coalition C and each i in C, let $\gamma_i^s = \alpha_i^\star(C, \gamma^{s-1})$. For all other agents j, let $\gamma_j^s = 0$. End in the first step S such that $\gamma^S = \gamma^{S-1}$, and denote γ^S by $\gamma^\star$.

Theorem 3. *The profile $\gamma^\star$ constructed by algorithm $\mathcal{A}^\star$ is the unique Nash-credible payoff profile.*

Proof. The argument is analogous to the proof of Theorem 1, using $\alpha^\star(C, \theta)$ instead of $\alpha(C, \theta)$ as the relevant bargaining solution in each coalition C with outside option profile θ. □

Corollary 4.2. *As the bargaining friction q goes to 0, the unique equilibrium cutoff profile γ of the game $\mathcal{G}$ converges to the unique Nash-credible payoff profile.*

In other words, the limit equilibrium payoff profile $\gamma^\star$ is the only profile that is such that, for every agent i, $\gamma_i^\star$ is the maximum that can be justified as resulting from Nash bargaining in some coalition C while honoring the others' outside options $\gamma_{-i}^\star$. I refer to a coalition C in which i can justify her payoff γ_i in this way as *i's Nash-best coalition*.

Definition 4.4. Coalition C is *i's Nash-best coalition* if $\gamma_i^\star = \alpha_i(C, \gamma_{-i}^\star)$.

The argument analogous to the one in section 3.4 shows that each agent i has a generically unique Nash-best coalition C. In this case, for all sufficiently small bargaining frictions q, in equilibrium each agent makes offers to her Nash-best coalition. In particular, algorithm $\mathcal{A}^*$

identifies the coalitions that form in equilibrium for all small enough bargaining frictions, as well as the corresponding tier structure. The limit sharing rule in each coalition is given by the Nash bargaining solution, with the binding outside options determined by the Nash bargaining solution in higher tiers.

Example 4.1. In the example discussed in section 3.5, the equilibrium offers for all small enough bargaining frictions are as described there, with limit payoffs

$$\gamma_1^\star = \gamma_2^\star = \frac{y_{12}}{2} \quad \gamma_3^\star = y_{23} - \gamma_1^\star \quad \gamma_4^\star = y_{34} - \gamma_3^\star. \tag{5}$$

Since the only heterogeneity among agents comes from the surplus function y, the Nash bargaining solution in each coalition boils down to equal sharing. Hence, in this case $\gamma^\star$ is the only profile that gives each agent the maximum that she can justify using equal sharing in some coalition—subject to the constraint that her opponent j gets at least $\gamma_j^\star$. The limit sharing rule in the first-tier coalition $\{1, 2\}$ is equal sharing; the limit sharing rule in the second-tier coalition $\{2, 3\}$ is as close as possible to equal sharing subject to the constraint that agent 2 obtains at least her share in $\{1, 2\}$; and the limit sharing rule in the third-tier coalition $\{3, 4\}$ is as close as possible to equal sharing subject to the constraint that agent 3 obtains at least her share in $\{2, 3\}$.

5 Vertically Differentiated Markets

This section describes comparative statics in markets that are vertically differentiated by skill, risk aversion, or proposer probability. For simplicity, I restrict attention to limit equilibrium outcomes, the generic case in which every agent has a unique Nash-best coalition, and comparative statics that do not alter the ranking of agents in terms of their attributes.

5.1 Vertical differentiation by skill

Denote agents by $1, 2, \ldots, n$, ordered from most to least skilled, in the following sense. Suppose that, for every two agents i and j with $i < j$ and every coalition $C \subseteq N - i - j$, i's marginal contribution to C is larger than j's, or, equivalently, $y(C + i) > y(C + j)$.[11] Suppose also that skill is the only source of heterogeneity. Corollary 5.1 highlights that agents' payoffs are increasing in their skill.

Corollary 5.1. $\gamma_1^\star \geq \gamma_2^\star \geq \cdots \geq \gamma_n^\star$.

[11] I write $N - i - j$ instead of $N \backslash \{i, j\}$, $y(C + i)$ instead of $y(C \cup \{i\})$, etc.

Proof. Letting $i < j$, we want to show that $\gamma_i^\star \geq \gamma_j^\star$. Since $\gamma_i^\star \geq y(i) \geq y(j)$, we can assume without loss of generality that $\gamma_j^\star > y(j)$. Let s_i and s_j be the steps of algorithm $\mathcal{A}^\star$ in which i and j's outside options are updated from 0 to a strictly positive number, respectively. Since i is more skilled than j, the only way in which we can have $\gamma_i^\star < \gamma_j^\star$ is that $s_i > s_j$, so that outside options are higher in s_i than in s_j (i.e., $\gamma^{s_i - 1} \geq \gamma^{s_j - 1}$). Suppose for contradiction that $s_i > s_j$. Let C be a γ^{s_j-1}-Nash perfect coalition containing j. This coalition C cannot contain i and, since $\gamma_j^\star > y(j)$, C is not a singleton. Since $y(C + i - j) > y(C)$, the γ^{s_j-1}-Nash share of every agent in C other than j is strictly larger in $C + i - j$ than in C, which contradicts the fact that C is a γ^{s_j-1}-Nash perfect coalition. □

Corollary 5.2 below highlights that changes in the skills of low-skill agents can only have limited effects on the payoffs of high-skill agents. Naturally, an increase in the skill of an agent that increases the surplus $y(C)$ of a coalition C can have direct (positive) effects on the payoffs of the members of C with higher skills. Also, this increase can positively or negatively affect the payoffs of agents with lower skills that are not in C (via the propagation from top tiers down illustrated in section 3.5, for example). But this increase cannot have indirect effects on the agents with higher skills, in the following sense.

Corollary 5.2. *A change in the surplus $y(C)$ of a coalition C that contains agent j cannot affect the payoff of any agent $i < j$ that is not in C.*

Proof. Let $i < j$, let C be a coalition that contains j but not i. We show that γ_i^* is determined independently of $y(C)$ both before and after the change in $y(C)$. Consider the situation before the change in $y(C)$ (the argument after the change in $y(C)$ is similar). Let D be i's Nash best coalition. Note that, since skill is the only source of heterogeneity, we have that $\gamma_i^\star \leq \gamma_k^\star$ for all k in D, where the inequality is strict for those agents whose Nash best coalition is not D. Hence, $\gamma_i^\star$ is determined by $y(D)$ and the productivity of (a subset of) the Nash-best coalitions of the agents k with $\gamma_k^\star > \gamma_i^\star$. The corollary follows from the fact that C is not the Nash best coalition of any such agent k, because otherwise we would have $\gamma_j^\star \geq \gamma_k^\star > \gamma_i^\star$, a contradiction of Corollary 5.1. □

The combination of Corollary 5.1 and Corollary 5.2 implies that, when the source of heterogeneity among agents is their skill, bargaining outcomes are determined from the agents with the highest payoffs down—in the sense that changes in the skill of an agent do not have indirect effects on the agents with higher payoffs. In contrast, as I now illustrate, when the source of heterogeneity among agents is either their risk aversion or their bargaining power, the bargaining outcomes are determined from the agents with the lowest payoffs up.

5.2 Vertical differentiation by risk aversion

Denote agents by $1, 2, \ldots, n$, ordered from most to least risk averse, in the following sense. Suppose that, for any two agents $i < j$, there exists a strictly concave function g such that u_i is the composition $g \circ u_j$ of the functions g and u_j. Suppose also that risk aversion is the only source of heterogeneity. It is well known that an increase in the risk aversion of an agent—in the sense that her utility function changes from u to $g \circ u$, for some strictly concave utility function g—reduces her Nash bargaining share in any given coalition (e.g., Osborne and Rubinstein, 1990, section 2.4.1). Corollary 5.3 provides an analog of this result in the present coalition formation setting.

Corollary 5.3. $\gamma_1^\star \leq \gamma_2^\star \leq \cdots \leq \gamma_n^\star$.

Proof. The proof is similar to the one of Corollary 5.1, and is deferred to Appendix A.6. □

Corollary 5.4 below highlights that a change in the risk aversion of an agent can only have limited effects on the payoffs of more risk averse agents. An increase in the risk aversion of an agent j increases the payoff of a more risk averse agent i when the Nash-best coalition of both agents is the same. Also, this increase can positively or negatively affect the payoffs of less risk averse agents (via the propagation from top tiers down illustrated in section 3.5, for example). However, this increase cannot have indirect effects on more risk averse agents, in the following sense.

Corollary 5.4. *Consider two agents $i < j$ that do not have a common Nash-best coalition. A change in j's risk aversion cannot affect i's payoff.*

Proof. Let $i < j$. I argue that, in every step s of algorithm $\mathcal{A}^*$, $\gamma_j^s \neq 0$ only if $\gamma_i^s \neq 0$. This implies that, if i and j do not have a common Nash-best coalition, then i's payoff is determined independently of j's risk aversion. Suppose for contradiction that j is in a γ^s-Nash-perfect coalition C, but i is not in any γ^s-Nash-perfect coalition. Let s' be the first step s for which this is the case. If C is a singleton (that is, $C = \{j\}$), then $\max_{j \in C \subseteq N} \alpha_j^\star(C, \gamma^{s'}) = y(j)$, which implies that $\max_{i \in C \subseteq N} \alpha_i^\star(C, \gamma^{s'}) = y(i)$ as well, so $\{i\}$ is a $\gamma^{s'}$-perfect coalition, a contradiction. If C is not a singleton, then for every member $k \neq j$ in C, $\alpha_k^\star(C + i - j, \gamma^{s'}) > \alpha_k^\star(C, \gamma^{s'})$, which contradicts the fact that C is a γ^s-Nash perfect coalition. □

The case in which agents are vertically differentiated by proposer probability is analogous: Denote agents by $1, 2, \ldots, n$, ordered from the one with the lowest proposer probability to the one with the highest proposer probability (that is, $p_i < p_j$ if $i < j$). If proposer probabilities are the only source of heterogeneity, similar arguments show that Corollary 5.3 and the following analog of Corollary 5.4 hold in this case as well: Consider two agents $i < j$ that do not have a common Nash-best coalition. A change in j's proposer probability cannot affect i's payoff.

The contrast between markets that are vertically differentiated by skill and markets that are vertically differentiated by risk aversion or proposer probability suggests that the effects of government policies like price caps or taxes can have very different effects depending on the source of heterogeneities in the market. For example, a minimum wage of 40 in the example of a market vertically differentiated by skill of section 3.5 does not affect the sharing rule in the top-tier coalition, and harms the agents with the lowest payoffs because it prevents them from matching with others. In contrast, a similar policy in a market that is vertically differentiated by risk aversion or proposer probability can influence the sharing rule in the top-tier coalitions—hence increasing the payoffs of the agents with the lowest payoffs.

5.3 Assortative matching in multi-sided markets

In markets in which agents are vertically differentiated either by risk aversion or by proposer probability, there is no sense in which agents match assortatively. Indeed, in this case every non-singleton coalition that forms in equilibrium contains one of the most risk averse agents or one of the agents with the lowest proposer probability, respectively, because these are the agents with the lowest payoffs. I now investigate the conditions under which agents match assortatively when they are vertically differentiated by skill, and the implications of positive assortative matching for how the effects of changes in fundamentals propagate through the market via agents' outside options.

Consider multi-sided markets where, for each coalition C, $y(C)$ is negligible unless C contains exactly one agent on each side of the market. Label the agents on each side of the market by integers, ordered from most to least skilled, in the following sense. Suppose that, for every two agents $i < j$ on the same side of the market, and for every coalition C consisting of exactly one agent on every other side of the market, $y(C + i) > y(C + j)$.

In this case, Corollary 5.2 above implies that an increase in the skill of an agent has no effects on the agents on her side of the market with higher skills. Corollary 5.6 below highlights how, when there is positive assortative matching (Definition 5.1), this increase propagates to agents with lower skills in a highly structured way.

Definition 5.1. Say that agents i and j *match* if someone's Nash-best coalition contains both i and j.[12] There is *positive assortative matching* if, for any two agents $i < i'$ on the same side of the market, and any two agents $j < j'$ on another side of the market, if i matches with j', then i' does not match with j.

As in the canonical marriage market model of Becker (1973), positive assortative matching arises when skills are complementary. More precisely, Lemma 5.5 below highlights that positive assortative matching is guaranteed when the surplus function is supermodular (Definition 5.2).

Definition 5.2. The surplus function y is supermodular if, for every two agents $i < i'$, every two agents $j < j'$ on another side of the market, and every coalition C that is such that $C + i + j$ contains exactly one agent on each side of the market,

$$y(C + i + j) + y(C + i' + j') > y(C + i' + j) + y(C + i + j').$$

Lemma 5.5. *If the surplus function y is supermodular, then there is positive assortative matching.*

Proof. The proof is standard and is deferred to Appendix A.7. □

In contrast to Becker's cooperative framework and much of the subsequent matching literature, the present theory pins down payoffs uniquely, and hence provides testable predictions about how positive assortative matching affects the way in which economic shocks propagate via outside options.

Corollary 5.6. *Let $i < j < k$ be in the same side of the market, and suppose that the surplus function y is supermodular. If an increase in the surplus $y(C)$ of a coalition C containing agent i does not affect j, then it does not affect k either.*

Proof. Consider two agents $i < j$ in the same side of the market, and let D be j's Nash best coalition. Say that an agent is strictly above (below) D if she strictly more (less) skilled than the agent in her side of the market that is in D. By Lemma 5.5, the fact that y is supermodular implies that there is positive assortative matching. In particular, an agent that is strictly below D is not in any γ^s-perfect coalition that contains an agent strictly above D. Hence, if an increase in the surplus $y(C)$ of a coalition that contains i does not change j's Nash-best coalition D nor the equilibrium sharing rule in D, then it does not affect the Nash-best coalition of any agent k that is strictly below D nor the equilibrium sharing rule in this coalition. □

[12] Note that if i and j match, then for all small enough bargaining frictions q, there is a coalition that forms in equilibrium that contains both of them.

6 Discussion

6.1 Related literature

6.1.1 Coalition formation in stationary markets

This paper contributes to the literature on bargaining in stationary markets (e.g., Rubinstein and Wolinsky 1985; Gale 1987; Binmore and Herrero 1988; Wolinsky 1990; de Fraja and Sákovics 2001; Manea 2011; Lauermann 2013; Nguyen 2015). Example 6.1 illustrates how the main departure of this paper from this literature—allowing agents to choose which coalition to propose—substantially changes the equilibrium outcome.

Example 6.1. A simple modification embeds the example discussed in section 3.5 into the standard random-matching framework: After the proposer is selected, instead of allowing her to choose which coalition to propose, one of the coalitions that she is part of is selected uniformly at random, and she can only propose how to split the surplus of this randomly-chosen coalition.[13] In this case, the limit payoff profile $\nu^\star$ satisfies[14]

$$\nu_1^\star = \frac{1}{4}\left(3y_{12} - 2y_{23} + y_{34}\right), \quad \nu_2^\star = y_{12} - \nu_1^\star, \quad \nu_3^\star = y_{23} - \nu_2^\star, \quad \nu_4^\star = y_{34} - \nu_3^\star. \tag{6}$$

Note that $\nu^\star$ differs from the Nash-credible payoff profile (5) and, in contrast to this profile, it does not have a natural interpretation in terms of the Nash bargaining solution. In particular, when bargaining to form coalition $\{1, 2\}$, agent 2 can effectively leverage the fact that she can also form a coalition with agent 3 to obtain more than half of the surplus y_{12}.[15] Agent 2's payoff cannot be justified as resulting from Nash bargaining in either $\{1, 2\}$ or $\{2, 3\}$. Instead, her payoff is a combination of the surpluses of all the relevant coalitions, which is at odds with the outside option principle and implies that the equilibrium lacks the tier structure described in section 3.4. □

It might seem surprising that allowing agents to choose which coalition to propose leads to qualitatively different predictions. Indeed, in the traditional random-matching framework just illustrated by Example 6.1, agents can always wait to be able to make an offer to any coalition of their choosing, and the expected cost of such waiting goes to zero as frictions vanish. Recall, however, that relative proposer probabilities are important drivers of

[13] This is equivalent to assuming that one coalition is selected uniformly at random and an agent of this coalition is selected uniformly at random to be the proposer.

[14] This follows from the equilibrium characterization in Nguyen (2015), which implies that the limit payoff profile minimizes $\sum_{1 \le i \le 4} \nu_i^2$ subject to $\nu_1 + \nu_2 \ge y_{12}$, $\nu_2 + \nu_3 \ge y_{23}$, and $\nu_3 + \nu_4 \ge y_{34}$.

[15] Recall that we have $y_{12} - y_{23} < y_{23} - y_{34}$, so $\nu_1^\star < \frac{y_{12}}{2}$, and hence $\nu_2^\star > \frac{y_{12}}{2}$.

bargaining power. Allowing proposers to choose which coalition to propose does not affect the proposer probabilities, but it does affect the probability that an agent can obtain positive surplus when she is selected to propose—which effectively amounts to the same thing. For instance, in Example 6.1, agent 2 makes acceptable offers twice as often as agent 1 (because agent 2 makes acceptable offers both when selected to propose $\{1, 2\}$ and $\{2, 3\}$, whereas agent 1 only makes an acceptable offer when selected to propose $\{1, 2\}$), and this translates into a higher payoff for agent 2 than for agent 1.

In the special case of networked buyer-seller markets for a homogeneous good, Talamàs (2019) already illustrated how allowing proposers to strategically choose whom to make offers to can make a substantial difference to the equilibrium outcome. Both the framework and the characterization of the present paper are substantially more general. In particular, the characterization in Talamàs (2019) (i) is only valid near the frictionless limit, (ii) does not identify exactly who trades with whom, and (iii) relies on the highly-structured matching surpluses that arise in buyer-seller markets for a homogeneous good (where the surplus that each buyer-seller pair can generate equals the buyer's value for the good minus the seller's cost of producing the good).[16] In contrast, the present paper provides a different characterization based on a novel notion of credibility that (i) works equally well for arbitrary bargaining frictions and in the limit as frictions vanish, (ii) works for arbitrary coalitional surpluses, (iii) identifies exactly which coalitions form in equilibrium and how they can be naturally organized into tiers, and (iv) characterizes the limit equilibrium payoff profile in terms of a natural generalization of the Nash bargaining solution.

6.1.2 Coalition formation in non-stationary markets

This paper is also closely connected to the literature on bargaining in non-stationary markets. The three most closely related papers in this literature are Elliott and Nava (2019), Chatterjee, Dutta, Ray, and Sengupta (1993) and Okada (2011).

Elliott and Nava (2019) consider non-cooperative bargaining in an assignment game (where all productive matches are bilateral and there is no dynamic entry, but which is otherwise similar to the one in the present paper). Focusing on the generic case in which the assignment game has a unique efficient match, one of their main results is that there exists an efficient Markov perfect equilibrium when agents are sufficiently patient if and only if the

[16]The characterization in Talamàs (2019) leverages this structure to decompose the market into different submarkets (subgroups of buyers and sellers) such that—in the limit as frictions vanish—the good is transacted at the same price within each submarket, but at different prices across submarkets.

profile $\sigma^\star$ that gives each agent her Nash bargaining payoff in her efficient match is in the core of the assignment game. In the language of the present paper, this condition is that the Nash-credible profile $\gamma^\star$ is in the core of the assignment game. Moreover, the limit payoff profile in the efficient equilibrium that they characterize is the Nash-credible profile.[17] This connection suggests that the central credibility notion of the present paper has the potential to prove useful in future analyses of coalition formation beyond the stationary markets considered here.

Both the system (3) of equations that determines the unique stationary equilibrium payoffs and the algorithm that computes them are similar to the ones that determine the *no-delay* stationary equilibria in Chatterjee, Dutta, Ray, and Sengupta (1993).[18] This connection might seem surprising because this earlier seminal work focuses on a *rejector-proposes protocol* (in which the rejector of a proposal becomes the proposer in the next period) instead of the *random-proposer protocol* of the present paper. Ray (2007, section 7.7) and Compte and Jehiel (2010) discuss how the former protocol gives considerably more bargaining power to the receiver of the offer than the latter, and how this explains the contrasting predictions often obtained under these two protocols. Intuitively, however, the dynamic entry of agents considered in the present paper implies that agents do not have to consider how the market might evolve after they reject an offer, which implies that the difference between these two protocols is much less pronounced.

Even in the special case of homogeneous preferences considered in Chatterjee, Dutta, Ray, and Sengupta (1993), the results of this paper differ in various dimensions. First, this paper characterizes the equilibrium payoff profile as being the only one that satisfies a natural credibility property. Second, this equilibrium always exists, and it is the unique stationary equilibrium instead of being one of the many possible stationary perfect equilibria and existing only under certain conditions.[19] Finally, not surprisingly given the qualitative differences between the settings, the predictions of the resulting theories are qualitatively distinct. For

[17] The bargaining friction in Elliott and Nava (2019) is delay instead of breakdown risk, which implies that the relevant Nash threat points in their setting are 0 instead of the autarky payoffs (see Binmore et al. 1986).

[18] See Ray (2007, section 7) for an explicit description of the algorithm that computes the *no-delay* stationary equilibria in Chatterjee, Dutta, Ray, and Sengupta (1993), and Ray and Vohra (1999) for an extension of this algorithm to games with externalities.

[19] Informally, a sufficient condition for a no-delay equilibrium to exist in the setting of Chatterjee, Dutta, Ray, and Sengupta (1993) is that no one's expected equilibrium payoff in such equilibrium increases when the set of active players shrinks. This condition is not relevant in the present setting because the dynamic entry of agents ensures that the relevant matching opportunities are constant over time. As a result, in contrast to this previous work, no conditions need to be put on the surplus function to guarantee equilibrium existence.

example, the equilibrium in the present setting is stationary (instead of evolving as different coalitions form), and the limit equilibrium payoffs do not depend on the order realization of the proposers. Moreover, the endogenous evolution of the market in Chatterjee, Dutta, Ray, and Sengupta (1993) implies that—unlike in the present setting—an agent does not necessarily benefit when she becomes more productive (because her improved outside option can incentivize others to avoid making offers to her, which can, in turn, make it more likely that the market will evolve against her).

Finally, Okada (2011) considers the random-proposer version of the model in Chatterjee, Dutta, Ray, and Sengupta (1993) and shows that, when the surplus function is superadditive, there is an efficient equilibrium (in which the grand coalition N forms with probability 1) for all sufficiently small bargaining frictions if and only if the Nash bargaining solution in the grand coalition N is in the core. When this condition is satisfied, the Nash-credible profile $\gamma^\star$ coincides with the Nash bargaining solution in the grand coalition N, so it can be seen as a generalization of Okada's prediction in stationary settings in which the surplus function need not be superadditive.[20]

6.1.3 Relation to other Nash bargaining approaches to coalition formation

The idea of building a theory of coalition formation from the Nash bargaining solution goes back at least to Rochford (1984), who defines a *symmetrically-pairwise-bargained* payoff profile of an assignment game with transferable utility as one that satisfies the following property: Each matched pair shares output according to the Nash bargaining solution—with each agent's disagreement point being the maximum that she could achieve in any other match taking the others' payoffs as given. While there are usually multiple symmetrically-pairwise-bargained payoff profiles, Burguet and Caminal (2020) show that a similar idea uniquely pins down agents' payoffs in a general coalitional setting in which only one coalition can form, and provide strategic foundations for the resulting coalition formation solution concept. The spirit of these solution concepts is close to the one described in the present paper, but the non-cooperative approach described here suggests that, as in Binmore, Rubinstein, and Wolinsky (1986), outside options do not enter as disagreement points, but act instead as bounds on the range of validity of the Nash bargaining solution.

[20] Focusing on environments in which the grand coalition generates the highest surplus and in which at most one coalition can form, Compte and Jehiel (2010) show that if an asymptotically efficient stationary equilibrium exists, the corresponding profile of limit payoffs is the profile that maximizes the product of agents' payoffs among those in the core—a property that the Nash-credible profile does not necessarily satisfy.

Collard-Wexler, Gowrisankaran, and Lee (2019) provide strategic foundations for the *Nash equilibrium in Nash bargains* (Horn and Wolinsky 1988), which is a widely used bargaining solution concept for bilateral oligopoly settings. In contrast to the coalition formation approach of the present paper—in which each agent can be part of at most one coalition—the *Nash-in-Nash* solution assumes that all the parties trade with each other (i.e., that all possible coalitions form) and derives prices for each bilateral contract as a function of the fundamentals. In particular, being a surplus division rule *for a given network*, the Nash-in-Nash solution does not necessarily provide sharp predictions about which network will form.

Corollary 2.5 is the key property of the equilibrium sharing rule in each coalition that guarantees that there exists a unique credible profile. Pycia (2012) considers a setting in which agents have preferences over coalitions, and shows that this property (pairwise aligned preference over coalitions) is key for the existence of stable outcomes in his setting.[21] Pycia (2012) also illustrates how the Kalai and Smorodinsky (1975) bargaining solution does not satisfy this property and, as a result, if the sharing rule in each coalition is given by this solution, then the existence of a stable coalition structure is not guaranteed. Similarly, the following example illustrates that the credibility notion of the present paper does not necessarily pin down payoffs uniquely when applied to a solution concept that does not satisfy this property.

Example 6.2. Let $N = \{1, 2, 3, 4\}$, and suppose that there are three coalitions with non-negligible surplus: $\{1, 2\}$, $\{3, 4\}$ and $\{1, 2, 3, 4\}$, with

$$y(1, 2) = 180, \quad y(3, 4) = 60, \quad y(1, 2, 3, 4) = 240.$$

Let $u_1(s) = u_2(s) = u_4(s) = s$ and $u_3(s) = s^{\frac{1}{3}}$. In the spirit of the Kalai-Smorodinsky bargaining solution, suppose that, for every pair (C, θ) with $y(C) \geq \sum_{i \in C} \theta_i$, the sharing rule $\sigma(C, \theta)$ is such that there is a constant $a_{C,\theta}$ such that, for every i in C,

$$u_i(\sigma_i(C, \theta)) = \max\left[u_i(\theta_i), u_i(y(C))a_{C,\theta}\right] \text{ and } \sum_i \sigma_i(C, \theta) = y(C).$$

In this case, the first step of the version of algorithm $\mathcal{A}$ that uses the bargaining solution σ instead of α identifies $\{1, 2\}$ as a perfect coalition, and hence, $\gamma^1 = (90, 90, 0, 0)$. In the second step, we have

$$\sigma_3(\{3, 4\}, \gamma^1) \approx 19 \text{ and } \sigma_4(\{3, 4\}, \gamma^1) \approx 41$$

[21] Pycia (2012) highlights that among efficient and monotonic sharing rules, the ones that generate pairwise-aligned preferences are precisely those which are *consistent* in the sense of Harsanyi (1959) and Thomson and Lensberg (1989).

and

$$\sigma_3(\{1,2,3,4\},\gamma^1) \approx 7 \text{ and } \sigma_4(\{1,2,3,4\},\gamma^1) \approx 53.$$

In particular, agent 3's γ^1-best coalition is $\{3,4\}$ while agent 4's γ^1-best coalition is $\{1,2,3,4\}$, so the algorithm ends with $\gamma = (90, 90, 0, 0)$. Hence, for all $\sigma_3(\{1,2,3,4\},\gamma) < \nu_3 < \sigma_3(\{3,4\},\gamma)$, the profile $x = (90, 90, \nu_3, 60 - \nu_3)$ satisfies $x_i = \max_{i \in C \subseteq N} \sigma_i(C, x_{-i})$ for every i in N.

6.2 Optimal entry, stationarity and efficiency

In order to focus attention on how bargaining shapes market outcomes—and to stay as close as possible to the literature on bargaining in stationary markets—this paper has kept the market stationary by assuming that every agent that leaves the market is replaced by a replica. I now describe how the equilibrium characterization extends to a more realistic setting where the stationarity of the market is consistent with optimal entry decisions.[22] I then conclude by briefly discussing the efficiency of the equilibrium outcome in this case.

Interpret the set N as containing types of agents. To stay as close as possible to the benchmark game $\mathcal{G}$, assume that the surplus that each set of agents can generate by matching only depends on the types that it contains; that is, letting C be the set of types that are represented in a given set of agents, the surplus that these agents generate when they match is $y(C)$.[23] I start by modifying the bargaining protocol to allow for more than one proposer in each period so that, generically, the equilibrium outflow of agents is identical in every period. I then consider a population of potential entrants arriving on the scene in each period—as in Gale (1987) and Wolinsky (1987), for example—and I discuss the conditions under which the stationarity of the market is consistent with their optimal entry decisions.

Step 1 (Modification of the bargaining protocol). Start by assuming that there are always $n_i \geq n$ active agents of each type in the market, where n denotes the cardinality of N. The modified protocol is as follows: In each period, one agent of each type is selected uniformly at random to propose. Each proposer of type i chooses a coalition C (with $i \in C \subseteq N$), and proposes a split of the surplus $y(C)$. For each such proposal, one active agent of each type in $C - i$ is selected to respond to this proposal; assume that no proposer is called to respond to an offer and that, in any given period, no agent is called to respond to more than one proposal. The selected agents respond in (a pre-specified) sequence. If all of them

[22]See Manea (2017) for an investigation of how a stationary market can emerge via endogenous optimal entry in a pairwise-matching bargaining framework featuring random matching.

[23]For expositional clarity, here I reserve the term "coalition" to refer to a set of *types*, while I use the term "match" to mean a set of *agents* (that match).

accept, they match and leave the market with their agreed shares. Otherwise, they stay in the market for the following period.

Step 2 (Endogenous entry decisions). For any given set of primitives y, δ, $(u_i)_{i\in N}$, and $(p_i = \frac{1}{n_i})_{i\in N}$, the unique stationary equilibrium payoff profile under the modified protocol satisfies (3), exactly as in the original protocol, and is hence the unique credible profile γ. The advantage of this modified protocol is that, in the generic case in which each agent has exactly one best coalition, the equilibrium flow of agents out of the market is the same in every period. Focusing for simplicity on this generic case, let's now come back to the assumption that there are always n_i active agents of each type i. Letting m_i denote the number of agents of type i that flow out of the market in each period in equilibrium, such stationarity requires that m_i agents of each type i enter the market in each period. Suppose that, at the beginning of each period, there is an identical population of potential entrants of each type i. For simplicity, let's focus on the case in which the stock n_i of active agents of each type is large, so that each potential entrant's effect on the proposer probabilities is negligible. The potential entrant of type i with the m^{th} lowest opportunity cost of entry, c^i_m has incentives to enter the market if and only if the expected equilibrium payoff γ_i of doing so is larger than the cost c^i_m. Hence, the stationarity of the market is consistent with optimal entry decisions if $c^i_{m_i} \leq \gamma_i \leq c^i_{m_i+1}$ for each type i.

Note that when $c^i_{m_i} < \gamma_i < c^i_{m_i+1}$ for each type i, small changes in the primitives y, δ, $(u_i)_{i\in N}$, and $(p_i)_{i\in N}$ do not affect the equilibrium flow of agents into the market nor the equilibrium flow of agents out of the market, so such comparative statics are well defined—in the sense that these changes are consistent with the stationarity of the market even if we take the entry costs as given. Naturally however, as the following example illustrates, taking the entry costs as given narrows down the set of primitives under which a stationary market is consistent with optimal entry decisions.

Example 6.3. As in the example in section 3.5, $N = \{1, 2, 3, 4\}$, preferences are homogeneous, and there are only three coalitions $\{1, 2\}$, $\{2, 3\}$ and $\{3, 4\}$ with non-negligible surpluses y_{12}, y_{23} and y_{34}. Suppose that all types have the same entry costs (i.e., $c^i_m = c_m$ for all types i and all $m \geq 1$). Consider an equilibrium in which, in each period, the proposer of type 1 makes an offer to type 2 and vice versa, the proposer of type 3 makes an offer to type 2, the proposer of type 4 makes an offer to type 3, and all these offers are accepted. Letting p_i

denote i's proposer probability, the corresponding limit payoffs are[24]

$$\beta_1^* = \frac{p_1}{p_1+p_2} y_{12}, \quad \beta_2^* = \frac{p_2}{p_1+p_2} y_{12}, \quad \beta_3^* = y_{23} - \beta_2^* < \beta_1^*, \quad \beta_4^* = y_{34} - \beta_3^*$$

and the flow of agents out of the market is given by $m_1 = 2$, $m_2 = 3$, $m_3 = 2$ and $m_4 = 1$. This equilibrium is consistent with agents' entry decisions if $c_1 < \beta_4^* < c_2 < \beta_3^*$ and $\beta_1^* < c_3 < \beta_2^* < c_4$. In particular, the stationary of the market requires that the inflow of agents of type 2 is larger than the required inflow of agents of type 1. Since entry costs are the same across types, this requires that $\beta_2^* > \beta_1^*$ (which requires, in turn, that the stock n_1 of agents of type 1 is larger than the stock n_2 of agents of type 2, so that $p_2 > p_1$).

Finally, I conclude by discussing the efficiency properties of the equilibrium under this interpretation of the stationary market assumption. On the one hand, bargaining frictions can naturally lead to an inefficient equilibrium outcome. For example, suppose that $N = \{1,2,3,4\}$, and that there are only three coalitions with non-negligible surplus: $\{1,2\}$, $\{3,4\}$ and $\{1,2,3,4\}$, with $y(1,2) = y(3,4) = 2$ and $y(1,2,3,4) = 3$. When the bargaining frictions are high, the equilibrium cutoffs are close to 0, so everyone proposes the grand coalition $\{1,2,3,4\}$. Hence, the agents that have matched in equilibrium until any given period can be rearranged into bilateral matches to generate strictly higher surplus.

On the other hand, the equilibrium outcome in the limit as bargaining frictions vanish is efficient, in the following sense. Define the *net surplus* generated in a given period as the sum of the surplus of the coalitions that form in this period minus the sum of the entry costs of the members of these coalitions. I argue that the net surplus generated in equilibrium in the limit as bargaining frictions vanish is equal to the maximum surplus that could be generated by a social planner that can freely decide who enters as well as which matches form in each period. To see this, suppose for contradiction that such a social planner can generate a strictly higher net surplus in a given period than the one generated in equilibrium. This implies that there exists a group of agents that can generate a strictly higher net surplus than the sum of their limit equilibrium payoffs net of their entry costs. Note that $\gamma_i^\star - c_m^i$ is a lower bound on the limit net equilibrium payoff of the agent of type i with the m^{th} smallest entry cost.[25] Hence, there is a coalition $C \subseteq N$ that is such that $y(C) > \sum_{i \in C} \gamma_i^\star$, which contradicts the fact that $\gamma^\star$ is Nash credible.

[24]Such an equilibrium exists for all small enough bargaining frictions if $\frac{p_2}{p_1+p_2} y_{12} > \frac{p_2}{p_2+p_3} y_{23}$, $y_{23} - \frac{p_2}{p_1+p_2} y_{12} > \frac{p_3}{p_3+p_4} y_{34}$ and $y_{34} > y_{23} - \frac{p_2}{p_1+p_2} y_{12}$, which ensure that, for all small enough bargaining frictions, types 2, 3, and 4 indeed have incentives to make offers to types 1, 2, and 3, respectively.

[25]Consider the agent of type i with the m^{th}-lowest cost. If $\gamma_i^* - c_m^i < 0$, then for all small enough bargaining frictions this agent does not enter the market (and her limit net payoff is 0). Otherwise, this agent's limit net equilibrium payoff is $\gamma_i^\star - c_m^i$.

A Appendix

A.1 Proof of Lemma 2.1

This proof generalizes standard arguments in the literature (e.g., Ray, 2007, Chapter 7) to the case in which utilities $(u_i)_{i\in N}$ are not necessarily linear. We can rewrite (1) as

$$u_i(\alpha_i) = \max\left[u_i(\theta_i), u_i(y(i)), \lambda_i u_i(y(i)) + (1-\lambda_i)u_i\left(y(C) - \sum_{j\in C-i}\alpha_j\right)\right] \text{ for all } i \text{ in } C, \tag{7}$$

where $\lambda_i := \frac{q}{1-(1-q)(1-p_i)}$. Definition A.1 is useful to compactly rewrite (7) as (8) below.

Definition A.1. For each type i, let the function $f_i : [y(i), \infty] \to \mathbb{R}$ be implicitly defined by

$$u_i(x) = \lambda_i u_i(y(i)) + (1-\lambda_i)u_i(f_i(x)).$$

In words, $f_i(x)$ is the amount that agent i must be able to obtain when she is the proposer in a stationary equilibrium if she is to be indifferent between accepting and rejecting the amount x (see Figure 2). Since u_i is strictly increasing, f_i is well defined and, since $\lambda_i \in (0,1)$, $f_i(x) > x$ for all $x > y(i)$. Moreover, as illustrated by Figure 2, the concavity of the utility function u_i implies that the difference between $f_i(x)$ and x is increasing in x. System (7) can be compactly written as follows.

$$f_i(\alpha_i) = \max\left(f_i(\theta_i), f_i(y(i)), y(C) - \sum_{j\in C-i}\alpha_j\right) \text{ for all } i \text{ in } C. \tag{8}$$

Existence of a solution to system (8) follows from Brouwer's fixed point theorem. To prove uniqueness, suppose for contradiction that there are two profiles α and α' that solve system (8). Define S to be the set $\{i \in C \mid \alpha_i \neq \alpha'_i\}$ of all agents for which these two solutions differ. Let i be one of the elements of the set S for which $f_i(\alpha_i) - \alpha_i$ is highest, and suppose without loss of generality that $f_i(\alpha_i) - \alpha_i$ is an upper bound on $\left\{f_j(\alpha'_j) - \alpha'_j\right\}_{j\in S}$. Since $f_i(\alpha_i) - \alpha_i$ is increasing in α_i, we also have that $\alpha_i > \alpha'_i$. Moreover we have that

$$f_i(\alpha_i) = y(C) - \sum_{j\in C-i}\alpha_j, \tag{9}$$

since, otherwise, $\alpha_i = \max(\theta_i, y(i))$, which contradicts the fact that $\alpha_i > \alpha'_i \geq \max(\theta_i, y(i))$. In particular,

$$f_j(\alpha_j) - \alpha_j \geq y(C) - \sum_{j\in C}\alpha_j = f_i(\alpha_i) - \alpha_i \geq f_j(\alpha'_j) - \alpha'_j \text{ for all } j \text{ in } S,$$

or, using again that $f_j(\alpha_j) - \alpha_j$ is increasing in α_j, $\alpha_j \geq \alpha'_j$ for all j in S. But then (9) combined with the fact that the function f_i is increasing and, by definition, $f_i(\alpha'_i) \geq y(C) - \sum_{j\in C-i}\alpha'_j$ implies that $\alpha'_i \geq \alpha_i$, a contradiction.

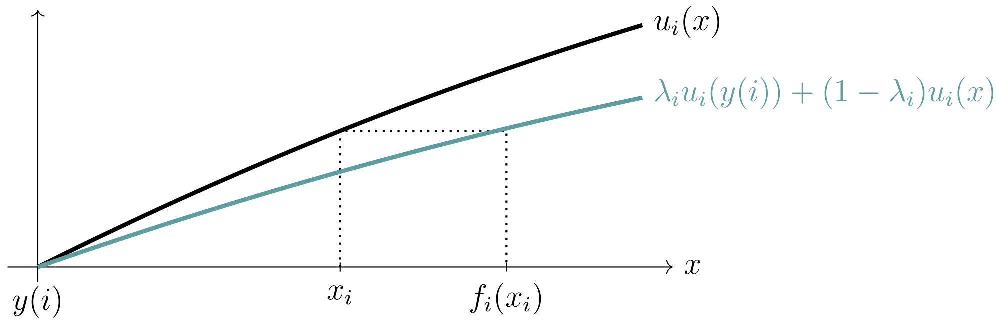

Figure 2: Illustration of the definition of the function f_i (Definition A.1).

A.2 Proof of Corollary 2.5

It follows from (8) that if $y(C) - \sum_{i\in C} \alpha_i(C,\theta) \geq y(D) - \sum_{i\in D} \alpha_i(D,\theta)$ then $f_i(\alpha_i(C,\theta)) - \alpha_i(C,\theta) \geq f_i(\alpha_i(D,\theta)) - \alpha_i(D,\theta)$ and hence $\alpha_i(C,\theta) \geq \alpha_i(D,\theta)$.

A.3 Proof of Corollary 2.6

Letting $S = \{i \in C \mid y(C) - \sum_{j\in C-i} \alpha_j > \max(\theta_i, y(i))\}$, we can rewrite (8) as

$$f_i(\alpha_i) - \alpha_i = \left(y(C) - \sum_{j\in C-S} \max(\theta_j, y(j))\right) - \sum_{j\in S} \alpha_j \quad \text{for all } i \text{ in } S. \tag{10}$$

Since $f_i(\alpha_i) - \alpha_i$ is the same for all i in S, and $f_i(\alpha_i) - \alpha_i$ is strictly increasing in α_i for all i in S, it follows from (10) that α_i is strictly increasing in $y(C)$.

A.4 Proof of Theorem 2

We can rewrite (3) as

$$u_i(\beta_i) = \max_{i\in C\subseteq N} \left[\lambda_i u_i(y(i)) + (1-\lambda_i)u_i\left(y(C) - \sum_{j\in C-i} \beta_j\right)\right] \text{ for all } i \text{ in } N. \tag{11}$$

We need to show that β solves (2) if and only if it solves (11). Using Definition A.1, (2) can be rewritten as

$$f_i(\beta_i) = \max_{i\in C\subseteq N} \left(y(C) - \sum_{j\in C-i} \alpha_j(C, \beta_{-i})\right) \text{ for all } i \text{ in } N \tag{12}$$

and (11) can be rewritten as

$$f_i(\beta_i) = \max_{i\in C\subseteq N} \left(y(C) - \sum_{j\in C-i} \beta_j\right) \text{ for all } i \text{ in } N. \tag{13}$$

An argument analogous to the one in section A.1 shows that (13) admits a unique solution. Hence, given that there is a unique credible payoff profile, it is enough to show that (12) implies (13). Suppose that β solves (2) (and hence its equivalent (12)). By Corollary 2.2, for all $i \neq j \in C$, we have $\alpha_j(C, \beta_{-i}) \geq \beta_j$ so

$$f_i(\beta_i) \leq \max_{i \in C \subseteq N} \left(y(C) - \sum_{j \in C-i} \beta_j \right) \text{ for all } i \text{ in } N.$$

Suppose for contradiction that this inequality is strict for some agent i, and let D achieve the maximum in the right-hand side of this strict inequality. Since β satisfies (2), $\alpha_j(D, \beta) = \beta_j$ for all j in D, so $f_i(\alpha_i(D, \beta)) < y(D) - \sum_{j \in D-i} \alpha_j(D, \beta)$, which contradicts the fact that, by definition, $\alpha(D, \beta)$ solves

$$f_i(\alpha_i(D, \beta)) = \max \left[f_i(\beta_i), f_i(y(i)), y(D) - \sum_{j \in D-i} \alpha_j(D, \beta) \right] \text{ for all } i \text{ in } D.$$

A.5 Proof of Lemma 4.1

This proof follows the lines of Proposition 4.2 in Osborne and Rubinstein (1990). Consider a pair (C, θ) with $y(C) \geq \sum_{i \in C} \max(y(i), \theta_i)$.[26] For every level of bargaining frictions $q > 0$, let $(\alpha_i(q))_{i \in C}$ denote the solution of (7). While the function f_i and the parameter λ_i also depend on q, for notational simplicity I do not specify this dependence. Letting $(\alpha_i^*)_{i \in C}$ be the limit of $(\alpha_i(q))_{i \in C}$ as $q \to 0$, we need to show that $(\alpha_i^*)_{i \in C}$ solves (4). For $i \in \{a, b\}$, let

$$u_i(\alpha_i) = \lambda_i u_i(y(i)) + (1 - \lambda_i) u_i \left(y(C) - \sum_{j \in C-i} \alpha_j \right). \tag{14}$$

It is enough to show that (α_a, α_b) converges to the pair (s_a, s_b) that maximizes

$$[u_a(s_a) - u_a(y(a))]^{p_a} [u_b(s_b) - u_b(y(b))]^{p_b} \tag{15}$$

subject to the constraint that $s_a + s_b \leq \alpha_a^* + \alpha_b^*$. Rewriting (14) as

$$u_i(f_i(\alpha_i)) - u_i(y(i)) = \frac{1}{1 - \lambda_i} [u_i(\alpha_i)) - u_i(y(i))].$$

and using that the first-order approximation of $\left(\frac{1}{1-\lambda_i} \right)^{p_i}$ around $q = 0$ is $1 + q$, we get that

$$[u_a(f_a(\alpha_a)) - u_a(y(a))]^{p_a} [u_b(\alpha_b) - u_b(y(b))]^{p_b}$$

[26] It follows from (1) that if $\alpha_i = y(i)$ then $\alpha_i \geq y(C) - \sum_{j \in C-i} \alpha_j$, so $\alpha_k \geq y(C) - \sum_{j \in C-k} \alpha_j$ for every k in C. In particular, $\alpha_i = \max(\theta_i, y(i))$ for all $i \in C$. Hence, when $y(C) \leq \sum_{i \in C} \max(\theta_i, y(i))$, the solution of (1) is $\alpha_i = \max(\theta_i, y(i))$ for all i in C.

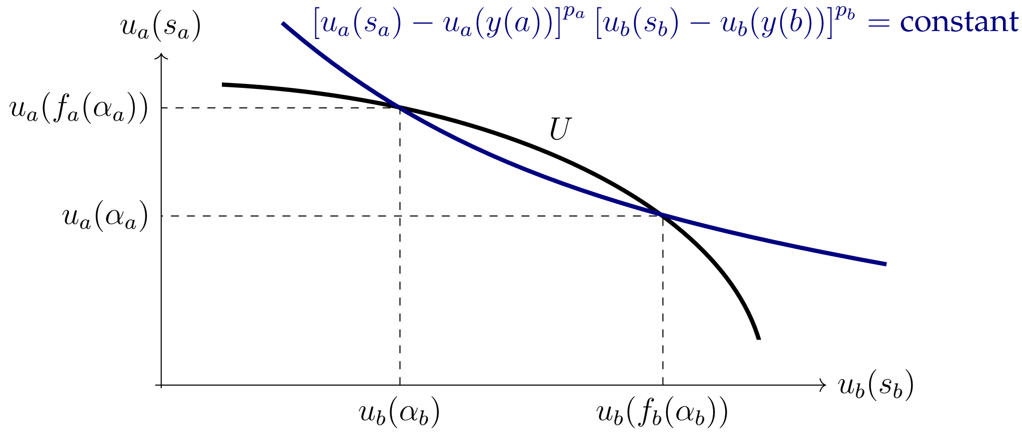

Figure 3: Illustration that if $(\alpha_i)_{i \in C}$ is such that, for any two types a and b with $\alpha_a > \max(\theta_a, y(a))$ and $\alpha_b > \max(\theta_b, y(b))$, then the pair (α_a, α_b) converges to the profile that maximizes (15). The graph U illustrates the set of utility pairs $(u_a(s_a), u_b(s_b))$ with $s_a + s_b \leq \alpha_a^* + \alpha_b^*$. This figure is analogous to Figure 4.2 in Osborne and Rubinstein (1990).

is to a first order equal to

$$[u_b(f_b(\alpha_b)) - u_b(y(b))]^{p_b}\,[u_a(\alpha_a) - u_a(y(a))]^{p_a}$$

As illustrated by Figure 3, the conclusion follows from the fact that, for each agent j and every x, $f_j(x)$ converges to x as q goes to zero.

A.6 Proof of Corollary 5.3

Letting $i < j$, we want to show that $\gamma_j^\star \geq \gamma_i^\star$. Since $\gamma_j^\star \geq y(j) = y(i)$, we can assume without loss of generality that $\gamma_i^\star > y(i)$. Let s_i and s_j be the steps of algorithm $\mathcal{A}^\star$ in which i and j's outside options are updated from 0 to a strictly positive number, respectively. Since i is more risk averse than j, the only way in which we can have $\gamma_i^\star > \gamma_j^\star$ is that $s_i < s_j$, so that outside options are higher in s_j than in s_i (i.e., $\gamma^{s_j-1} \geq \gamma^{s_i-1}$). Suppose for contradiction that $s_i < s_j$. Let C be a γ^{s_i-1}-Nash perfect coalition containing i. This coalition C cannot contain j and, since $\gamma_i^\star > y(i)$, C is not a singleton. Since $y(C + i - j) = y(C)$, the γ^{s_i-1}-Nash share of every agent in C other than j is strictly larger in $C + i - j$ than in C, which contradicts the fact that C is a γ^{s_i-1}-Nash perfect coalition.

A.7 Proof of Lemma 5.5

Suppose that there is no positive assortative matching. In particular, suppose that two agents $i < i'$ on the same side of the market and two agents $j < j'$ on another side of the market are such that i matches with j' and i' matches with j. This implies that

$$y(C + i' + j) = \gamma^\star_{i'} + \gamma^\star_j + \sum_{k \in C} \gamma^\star_k \text{ and } y(C + i + j') = \gamma^\star_i + \gamma^\star_{j'} + \sum_{k \in C} \gamma^\star_k.$$

Combining these two equations,

$$y(C + i' + j) + y(C + i + j') = \gamma^\star_i + \gamma^\star_{i'} + \gamma^\star_j + \gamma^\star_{j'} + 2\sum_{k \in C} \gamma^\star_k.$$

If the surplus function was supermodular, we would then have that

$$y(C + i + j) > \gamma^\star_i + \gamma^\star_j + \sum_{k \in C} \gamma^\star_k \text{ or } y(C + i' + j') > \gamma^\star_{i'} + \gamma^\star_{j'} + \sum_{k \in C} \gamma^\star_k,$$

which would be a contradiction.

References

BECKER, G. S. (1973): "A theory of marriage: Part I," *Journal of Political Economy*, 81, 813–846.

BINMORE, K., A. RUBINSTEIN, AND A. WOLINSKY (1986): "The Nash bargaining solution in economic modelling," *The RAND Journal of Economics*, 176–188.

BINMORE, K., A. SHAKED, AND J. SUTTON (1989): "An outside option experiment," *The Quarterly Journal of Economics*, 104, 753–770.

BINMORE, K. G. AND M. J. HERRERO (1988): "Matching and bargaining in dynamic markets," *The Review of Economic Studies*, 55, 17–31.

BURGUET, R. AND R. CAMINAL (2020): "Coalitional bargaining with consistent counterfactuals," *Journal of Economic Theory*, 187.

CHATTERJEE, K., B. DUTTA, D. RAY, AND K. SENGUPTA (1993): "A noncooperative theory of coalitional bargaining," *The Review of Economic Studies*, 60, 463–477.

CHODOROW-REICH, G., J. COGLIANESE, AND L. KARABARBOUNIS (2018): "The macro effects of unemployment benefit extensions: a measurement error approach," *The Quarterly Journal of Economics*.

CHODOROW-REICH, G. AND L. KARABARBOUNIS (2016): "The cyclicality of the opportunity cost of employment," *Journal of Political Economy*, 124, 1563–1618.

COLLARD-WEXLER, A., G. GOWRISANKARAN, AND R. LEE (2019): ""Nash-in-Nash" Bargaining: A Microfoundation for Applied Work," *Journal of Political Economy*, 127, 163–195.

COMPTE, O. AND P. JEHIEL (2010): "The coalitional Nash bargaining solution," *Econometrica*, 78, 1593–1623.

DE FRAJA, G. AND J. SÁKOVICS (2001): "Walras retrouvé: Decentralized trading mechanisms and the competitive price," *Journal of Political Economy*, 109, 842–863.

ELLIOTT, M. AND F. NAVA (2019): "Decentralized bargaining in matching markets: Efficient stationary equilibria and the core," *Theoretical Economics*, 14, 211–251.

GALE, D. (1987): "Limit theorems for markets with sequential bargaining," *Journal of Economic Theory*, 43, 20–54.

HAGEDORN, M., F. KARAHAN, I. MANOVSKII, AND K. MITMAN (2013): "Unemployment benefits and unemployment in the great recession: the role of macro effects," Tech. rep., National Bureau of Economic Research.

HALL, R. E. (2017): "High discounts and high unemployment," *American Economic Review*, 107, 305–30.

HALL, R. E. AND P. R. MILGROM (2008): "The limited influence of unemployment on the wage bargain," *American Economic Review*, 98, 1653–74.

HARSANYI, J. (1959): "A bargaining model for the cooperative n-person games," *Contributions to the theory of Games IV, ed. by A. W. Tucker and R. D. Luce*, 4, 325–355.

HORN, H. AND A. WOLINSKY (1988): "Bilateral monopolies and incentives for merger," *The RAND Journal of Economics*, 408–419.

JÄGER, S., B. SCHOEFER, S. YOUNG, AND J. ZWEIMÜLLER (2020): "Wages and the value of nonemployment," *Quarterly Journal of Economics*, 135, 1905–1963.

KALAI, E. AND M. SMORODINSKY (1975): "Other solutions to Nash's bargaining problem," *Econometrica*, 43, 513–518.

KRUSELL, P., T. MUKOYAMA, AND A. ŞAHIN (2010): "Labour-market matching with precautionary savings and aggregate fluctuations," *The Review of Economic Studies*, 77, 1477–1507.

LAUERMANN, S. (2013): "Dynamic matching and bargaining games: A general approach," *American Economic Review*, 103, 663–89.

LJUNGQVIST, L. AND T. J. SARGENT (2017): "The fundamental surplus," *American Economic Review*, 107, 2630–65.

MANEA, M. (2011): "Bargaining in stationary networks," *American Economic Review*, 101, 2042–2080.

——— (2017): "Steady states in matching and bargaining," *Journal of Economic Theory*, 167, 206–228.

NGUYEN, T. (2015): "Coalitional bargaining in networks," *Operations Research*, 63, 501–511.

OKADA, A. (2011): "Coalitional bargaining games with random proposers: Theory and application," *Games and Economic Behavior*, 73, 227–235.

OSBORNE, M. J. AND A. RUBINSTEIN (1990): *Bargaining and markets*, Academic Press, San Diego.

PISSARIDES, C. A. (2000): *Equilibrium unemployment theory*, MIT press.

POLANSKI, A. AND F. VEGA-REDONDO (2018): "Bargaining frictions in trading networks," *The BE Journal of Theoretical Economics*, 18.

PYCIA, M. (2012): "Stability and preference alignment in matching and coalition formation," *Econometrica*, 80, 323–362.

RAY, D. (2007): *A game-theoretic perspective on coalition formation*, Oxford University Press.

RAY, D. AND R. VOHRA (1999): "A theory of endogenous coalition structures," *Games and Economic Behavior*, 26, 286–336.

ROCHFORD, S. C. (1984): "Symmetrically pairwise-bargained allocations in an assignment market," *Journal of Economic Theory*, 34, 262–281.

RUBINSTEIN, A. AND A. WOLINSKY (1985): "Equilibrium in a market with sequential bargaining," *Econometrica*, 1133–1150.

SHIMER, R. (2005): "The cyclical behavior of equilibrium unemployment and vacancies," *American Economic Review*, 95, 25–49.

SORKIN, I. (2015): "Are there long-run effects of the minimum wage?" *Review of Economic Dynamics*, 18, 306–333.

SUTTON, J. (1986): "Non-cooperative bargaining theory: An introduction," *The Review of Economic Studies*, 53, 709–724.

TALAMÀS, E. (2019): "Price dispersion in stationary networked markets," *Games and Economic Behavior*, 115, 247–264.

THOMSON, W. AND T. LENSBERG (1989): *Axiomatic Theory of Bargaining With a Variable Population*, Cambridge University Press.

WOLINSKY, A. (1987): "Matching, search, and bargaining," *Journal of Economic Theory*, 42, 311–333.

——— (1990): "Information revelation in a market with pairwise meetings," *Econometrica*, 1–23.