



HAL
open science

Imperfect Social Learning on Matching Markets Implications for Stability

Laure Goursat

► **To cite this version:**

Laure Goursat. Imperfect Social Learning on Matching Markets Implications for Stability. Economics and Finance. 2019. dumas-02407515

HAL Id: dumas-02407515

<https://dumas.ccsd.cnrs.fr/dumas-02407515>

Submitted on 12 Dec 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



PARIS SCHOOL OF ECONOMICS
ÉCOLE D'ÉCONOMIE DE PARIS

MASTER THESIS N° 2019 – 01

Imperfect Social Learning on Matching Markets Implications for Stability

Laure Goursat

JEL Codes: D470, D830, D900

Keywords: Matching Markets, Imperfect Information, Bounded Rationality,
Social Learning, Pairwise Stability, Assortative Matching.



DISSERTATION - June 2019

**IMPERFECT SOCIAL LEARNING ON MATCHING MARKETS:
IMPLICATIONS FOR STABILITY**

Keywords: Matching Markets, Imperfect Information, Bounded Rationality, Social Learning, Pairwise Stability, Assortative Matching.
JEL codes: D470, D830, D900



PARIS SCHOOL OF ECONOMICS
ÉCOLE D'ÉCONOMIE DE PARIS

Supervisor:
Philippe JEHIEL
Paris School of Economics
Ecole Nationale des Ponts et Chaussées

Student:
Laure GOURSAT*
Paris School of Economics
M2 in Analysis and Policy in Economics
Applicant to PSE Ph.D. programme

*Contact details: laure.goursat@polytechnique.edu

Abstract

This dissertation introduces bounded rationality on matching markets, by way of imperfect cognition in social learning. On a two-sided agent-agent one-to-one matching market with Beckerian match utilities though forbidding transfers, a Live-Polarised-Unidimensional-Valuation (LPUV) rational agent m observes the surplus that a potential partner f is currently generating with her own match and uses it as an estimate of the surplus they (m and f) would jointly generate. The agent plugs this incorrect estimate of the surplus into the correct splitting rule, hence a coarse belief on match utilities. I compare the implicitly defined LPUV-stability to the usual (Gale-Shapley) stability concept, under odd or specific splitting rules, exogenous or endogenous surpluses. Quite remarkably, there exist splitting systems for which LPUV-stability is robust to the specification of surpluses. In addition, LPUV-stable matchings are disproportionately assortative with respects to GS-stable ones.

Acknowledgement

I cannot thank enough my supervisor Pr. Philippe JEHIEL for dedicating me time before and in the course of the dissertation work, for his helpful guidance and endless patience. I feel also grateful to Pr. Francis BLOCH for advice on the matching part and for endorsing my first research experience last year. I also would like to thank the audience of the December 18 2018 and April 9 2019 sessions of PSE micro-theory workgroup for constructive comments on the research question and model of this dissertation, first and foremost Pr. Olivier COMPTE, but also Prs. Catherine BOBTCHEFF, Bernard CAILLAUD, Antonin MACÉ.

Let me also mention the high-quality courses in Decision Theory (by Pr. Jean-Marc TALLON), Bounded Rationality (by Prs. Philippe JEHIEL and Olivier COMPTE) and Market Design (by Pr. Olivier TERCIEUX).

Finally, I am grateful to classmates Vivien CHBICHEB, Pierre BARDIER and Ph.D. students Euiyoung JUNG, Monica GOMEZ-OSPINA, Guilio IACOBELLI, Simon GLEYZE, Lucas GIRARD for general advice and support.

Contents

1	Introduction	3
1.1	Bounded rationality in matching: obvious combination but surprisingly blank field	3
1.2	Empirical puzzles on matching markets under the full rational theory	4
1.3	Research question and methodology	6
1.4	Modelling choices and matching vocabulary	7
1.5	Structure of the dissertation	8
2	The Model	8
2.1	Setting	8
2.1.1	Market structure	8
2.1.2	Information structure	9
2.2	The concept: Live-Polarised-Unidimensional-Valuation (LPUV)	10
2.2.1	LPUV-rationality	10
2.2.2	Interpretation	11
2.2.3	Two competing notions of pairwise stability	12
3	GS-stability and LPUV-stability: a comparative analysis	13
3.1	Existence and characterisation of stable matchings	13
3.1.1	Existence	13
3.1.2	Examples on a 2 x 2 market	14
3.1.3	Main characterisation of LPUV-stable matchings	15
3.1.4	Necessary conditions for LPUV-blocking pairs	16
3.2	Robust or crosswise conditions on splitting and surplus for stability	17
3.2.1	Robust or crosswise conditions under odd splitting rule	17
3.2.2	Robust or crosswise conditions under regular splitting	19
3.2.3	Graphical representations and simulations	22
3.3	Assortative matching stability and instability	24
3.3.1	Three special cases suggesting a LPUV bias towards assortativity	24
3.3.2	Assortative matching and endogenous surplus	26
4	Discussion	28
4.1	Value of the concept	28
4.2	Specification issues	29
4.3	Mirror cases	31
4.4	Relation to Valuation Equilibrium	32
5	Conclusion	33
6	Appendix	37
6.1	Proofs	37
6.2	Mirror cases: Results	45
6.2.1	Case 1: Egocentric estimation of match surplus	45
6.2.2	Case 2: Correct estimation of surplus, incorrect estimation of splitting rule	47
6.3	Link with Valuation Equilibrium	47
6.3.1	Valuation-Best matchings	47
6.3.2	Tree representation	50
6.4	Tool box	51
6.4.1	Matching belief matrices	51
6.4.2	Aggregate comparison of stabilities	52

1 Introduction

1.1 Bounded rationality in matching: obvious combination but surprisingly blank field

As a beginner in micro-economics, eager to explore the abundant theoretical literature, I have been struck by one discrepancy, that is the disconnection between research on matching markets and bounded rationality.

On the one hand, since Gale and Shapley (1962) [15], economists have been studying matching markets very thoroughly. The large matching literature can be accounted for by the complexity of the environment (results are sensitive to the specification of the market structure, agents preferences, objects priorities, and mechanism) and the variety of social objectives one can consider (whether it is incentive rationality, stability, efficiency, truthful reporting), objectives that incidentally prove to be partially antagonistic (see for example impossibility theorems in Roth (1982) [25], Papai (2000) [24], Ma (1994) [22], Kojima (2011) [19]). This already substantial literature is still expanding due to cross-discipline approaches (between economics, sociology, computer science, operations research), to new statistical techniques designed specifically for matching (see Chiappori and Salanié (2016) [10] for an extensive review of the econometrics of matching models), and to an empirical rise of centralised matching markets, primarily in public economics, calling for a mechanism design approach. Well-known examples of such markets include the matching of students to Universities which has been specifically explored by the school choice literature, initiated by Friedman (1955) [14] and revived by a group of mechanism designers in the early 21st century on the occasion of a school choice reform in Boston (Abdulkadiroğlu and Sönmez (2003) [2], Abdulkadiroğlu and al (2005) [3]).

On the other hand, following Simon (1962) [29], economists have questioned the full rationality assumption, namely the premise that the economic agent always makes the right decision (decision problem) or chooses the right strategy (game) to maximise her objective. Such evolution resulted from both normative and positive concerns. Indeed, on the normative side, full rationality seems unreasonably demanding for a human-being, even if bright, considering obvious and documented mind limitations. Pure rationality precisely means that the agent is able to maximise one objective given some expectations on her environment (instrumental rationality¹) and to form those expectations correctly, including on the other players strategies in a game, on the states of the world in a Bayesian setting (cognitive rationality¹). The cognitive rationality assumption is especially hard to buy in complex environments. On the positive side, full rationality is unsatisfactory as it does not account for empirical observations on how agents make decisions, play games (see Conlisk (1996) [12] for a review of direct evidence of bounded rationality in psychology and experimental economic literature). To tackle this two-fold concern, a wide array of bounded rationality notions of unequal normative and positive strengths have prospered since the 1960's and 1970's.

The puzzle is thorough, as a number of economists do show awareness. One clue is the value assigned to mechanisms with dominant strategies, that require no cognitive abilities. Even more disconnected from cognition is the concept of obviously dominant strategies (Li (2017) [20]); its behavioral interpretation is that a strategy is obviously dominant if and only if a cognitively disabled agent can recognise it as weakly dominant. Were it not for the knowledge of sharp human cognitive limitations, mechanism designers would surely not bother to design strategy-proof (obviously strategy-proof) mechanisms.

Yet, as far I know, no theoretical article has attempted to explicitly model bounded rationality in a matching environment. Liu and al (2014) [21] stands as the paper that is the closest to doing so. The authors set a two-sided one-to-one job market with transfers where firms have limited information and make inference on the match utilities using the knowledge that the match is

¹I will make use of this distinction between instrumental and cognitive rationalities in the following.

stable. They prove that under increasing and super-modular match surpluses, only the positive assortative matching can be stable. In the course of their proof, they distinguish several levels of sophistication of the inference made by firms, which could be interpreted as a hierarchy among several bounded rationality types.

Empirical literature on bounded rationality in matching is modest but supports my point. Using data on Australian college admissions, Artemov, Che and He (2017) [4] show that applicants choose not to apply to some college programs against their apparent interests, when doing so is unlikely to affect the outcome. Le Chapelain (2015) [9] performs rationality tests on students applying to colleges in Mexico and find that many of them do not correctly assess the various educational alternatives they are facing. But he fails to diagnose correlations between the degree or nature of irrationality and observable characteristics such as wealth (which he suggests could be positively correlated to information). Bounded rationality on matching markets appears as a wide-spread phenomenon.

The literature shortfall is even more surprising as classical arguments supporting bounded rationality are especially salient in matching environments, due to two aspects of such markets. First, data issues inherent to the matching market structure prevent learning. Indeed, at the core of the learning approach (either belief-based or performance-based), lies the idea that the agent plays the same game a significant number of times to perfectly apprehend her environment. By contrast, most matching markets allow for limited experience: when it comes to the marriage or job markets, one is "on the market" at most several times in her life. In relation to this point, Bayesian learning on matching markets requires multi-way inference on one's own preferences or match utilities (themselves depending on own characteristics, characteristics of the other side of the market, and on the match utility formation process) and on the equilibrium strategies of own and opposite sides of the market (themselves depending on the other agents' beliefs on match utilities and equilibrium strategies). Besides, the nature of the market, whether it is centralised or decentralised is problematic for different reasons. The size of centralised market is usually huge, which causes feedback data on states and strategies to exceed individual common storage and processing abilities. Conversely, the design of decentralised markets blocks feedback information as some parts of the market are hidden from each agent. In sum and stated in econometric terms, there is a problem of under-identification (few observations, many parameters).

Second, matching markets display a dual competition which agents find hard to understand and play. As an agent on side M of the market, I am competing against agents belonging to the same side M and to the opposite side F . This causes two issues. One issue stems from the difference in nature between individuals on each side (gender difference (male-female), collective difference (firm-employee), and so forth). The consequence is that introspection techniques to assess side F agents strategies cannot help. The second issue is that empirically agents tend to over-estimate the importance of competition within own side of the market compared to competition between sides, as emphasised by Roth and Sotomayor (1990) [27].

1.2 Empirical puzzles on matching markets under the full rational theory

Several empirical puzzles which the rational theory fails to explain remain pending at this point. Let me develop three of them.

The first puzzle lies in persistent instabilities. Indeed, matching markets seem to never reach a stable allocation. The marriage market provides a sound example. Observing data from early twentieth century, European countries, one could have thought that the marriage market was indeed stable. However, as soon as the marriage market was liberalised, approximately in the 1970's (legal as well as lifestyle evolutions), Europe experimented a sharp increase not only in the number of divorces but also in the rate of divorces (even taking into account the simulta-

neous drop in the marriage rate).² Yet, it is not that people finally got tired of marital life: a secondary market arouse. My interpretation is that liberalisation of the marriage market enabled to address intrinsic instability of the market: where the Incentive Rationality constraint was not satisfied, an individual would choose to divorce, and where the Incentive Compatibility constraint was not satisfied, a blocking pair would materialise in potential divorces followed by (re)marriage. This observation is at odds with theory predictions that under free bargaining the market should reach a stable allocation (Coase (1960) [11]). Candidate explanations for the instability of the marriage market include biological and psychological arguments. A popular view among biologists is that marriage is an institution that arouse accidentally, not fitted to human needs and desires. It could also be the case that people change identity and characteristics over the course of their lives so that preferences evolve in parallel. Those arguments I do not intend to discuss, I will rather search for a third explanation belonging to the field of economics and relevant for matching markets in general.

The second puzzle lies in the observation that self-declared matching preferences are very often assortative, including on markets where theory assesses the true preferences to be different. On the marriage market for instance, it is striking that a minority of men are appealing to a wide majority of women and conversely. Such observation is at odds with what mainstream theory in biology and psychology predicts. In psychological research, it remains an open question to determine which of similar or complementary individuals are the most likely to form a happy couple. Cornerstone papers by Newcomb (1961) [23], Byrne and Nelson (1965) [8] argued in favor of similar individuals. More recently, Cundiff and al (2015) [13] tested two interaction characters: what they called affiliation (warmth versus coldness) and control (dominance versus submissiveness). They concluded in favor of complementary agents for the specific case of interaction characters. Anyhow, conclusion for preferences is the same: the harmony and happiness of a couple result less of individual characteristics than of some interaction factor. There is not just a single man that is bound to lead to the highest satisfaction as husband for the whole set of women and conversely. So when people consider quality of partnership in daily life when matching, preferences are bound to be uncorrelated within a side of the market (possibly idiosyncratic, aligned...). From biology, since the discovery of the genotype (Johannsen (1909) [18]), it has been common knowledge that the building of individual characters of a child proceeds by a matching of alleles that is intrinsically complex (with the system of dominant or recessive alleles). So that if agents consider expected quality of offspring when matching, and even in the case when they agree on what this “quality of offspring” could mean, idiosyncratic preferences are bound to arise as well³. One may think that this marriage example is a little far-fetched or very specific. Let me take another example that may prove more telling within a circle of researchers⁴ as it deals with citations of research articles in scientific reviews: what one observes is that a minority of papers are cited by the majority of papers, so that some cited papers are very remote to the topic at study. This is likely to be at odds with the true distribution of papers with closely related topics or methodologies, and other more closely related work

²Eurostat, EU-28, census data: the crude divorce rate went from 0.8/1,000 inhabitants in 1965 to 2.0/1,000 inhabitants in 2015.

³This argument may remain opaque for a reader unfamiliar with genetics. Let me provide a simple illustration with the eye color. Imagine all people on the marriage market value blue-eye children more than those green-eye or brown-eye ones. The reproductive mechanism is such that someone with blue eye has a 50% chance of giving birth to a blue-eye child if he matches with a green-eye or brown-eye partner indifferently. Someone with brown eyes has a 12.5% chance of giving birth to a blue-eye child if he matches with a green-eye partner, but only 6.25% if he matches with a brown-eye partner. Ordinal preferences over green and brown eyes are: for blue-eye people $green \sim_{blue} brown$, for brown-eye people $green \succeq_{brown} brown$. Taking into account the grand-parents eye color or observing the alleles of parents for the genes ruling eye color, leads to even more idiosyncracies.

⁴Unlike the previous example on marriage which was my own puzzle, this second example is due to Pr. Philippe Jehiel himself. Let me clarify though that this description of assortativity in citations does not exclude explanations consistent with the full rationality theory such as: heterogeneity in quality, positive self-reinforcing loops operating on the most cited papers.

would probably deserve to end up in the literature review.

The third puzzle regards the obvious under-investment in information acquisition and strategy design on matching markets. For marriage, the mate searching period lasts between 10 and 15 years on average (from say, 18 to the early 30's⁵). This seems short considering on the one hand the stakes of this matching (married people usually have daily and intimate interactions) and the number of parameters relevant to the choice (due to multiplicity and variety of agents, attributes, competitors, as previously emphasised). This raises suspicion that either information acquisition and strategy design are highly costly to the agent (still obeying a strict cost-benefit analysis under perfect rationality) or that agents are acting under imperfect rationality.

Tentative explanations under full rationality remain unsatisfactory at this point. I give two examples, one for the second puzzle on assortativity, and one for the third puzzle on under-investment.

Gale and Shapley (1962) [15] proved that on a matching market with transfers and under super-modular match surplus function (equivalent to positive complementarities between characteristics of the matched agents), a stable matching is assortative. The intuition is simple. Suppose a matching is non-assortative, for example the two agents with the best individual characteristics are unmatched. Due to the super-modularity, if those two agents block and match together they increase the total surplus. Because the splitting of the surplus is not fixed, the pair can bargain on a splitting rule that increases utility of both partners, hence the incentive to block initially. One can repeat the reasoning through a descending induction. The issue with this argument is that it relies heavily on the possibility for monetary transfers. If the splitting rule is fixed, the perspective of an increased surplus for an unmatched pair does not necessarily fosters blocking.

In his *Theory of Marriage* (1973, 1974) [5] [6]⁶, Becker also models transfers, representing spouses as engaging in productive activities, jointly generating a surplus and then bargaining over a split of the surplus. Part 1 of Becker's theory is in complete information (on the surplus) but part 2 section §5 briefly describes a marriage market in incomplete information where agents get involved in optimal searching to estimate the match surpluses. In countries where divorce and remarriage are allowed, the incentive to stop the search early is greater than in countries where divorce is forbidden. Becker, Landes and Michael (1977) [7] develop a model to formalise the argument thus accounting for the sharp increase in the number of divorces and re-marriages in Western countries following the liberalisation of the marital institution. If I extrapolate the conclusion of the theory, it could be that the possibility for trials and errors provided by divorce pushes agent into under-investment in information acquisition and strategy design on the primary marriage market. But, generally speaking, although interesting in many aspects, this literature is irrelevant for other matching markets than marriage (as the authors rely on arguments from biology or sociology that cannot be generalised).

1.3 Research question and methodology

In a nutshell, bounded rationality makes a lot of sense in matching problems. Even more than this, matching seems to be the one environment where agents are the most likely to act with bounded rationality. Introducing bounded rationality may also shed lights on several empirical puzzles left unexplained under the full rationality assumption.

The research question I plan to address is thus the following: What is the nature of bounded rationality on matching markets? Does it reshuffle the matching outcome and if so, in what direction?

⁵Eurostat, France: In 2018, average age at (civil) marriage was 34 for men, 30.8 for women.

⁶In the Beckerian literature, marriage is considered very broadly: whether it is religious marriage, civil marriage, common-law marriage – it includes any mating of partners to experience marital life with the idea or ambition that the relationship should last for the rest of one's life. In the following, when appealing to the marriage market example, I will adopt the same broad definition.

The final purpose of introducing bounded rationality on matching markets is definitely not to fill a gap in the literature. It is rather: one, to account for the matching markets real mechanics (understanding stake); second, to renew the mechanism design literature so as to improve policy recommendations in light of modified behavioral modelling (implementation stake).

The methodology I intend to follow is elementary. I propose a cognitive procedure for agents participating to matching markets. I very carefully motivate the procedure as the one settling the conflict between matching market complexity and human cognitive limitations. I derive implications for stable allocations, with the ulterior motive to confront it to empirical market observations.

1.4 Modelling choices and matching vocabulary

Due to the complexity of any matching environment and to the novelty of the research question, modelling choices are essential. Regarding the matching environment, one has to decide whether to allow for transfers. On the pros side, matching with transfers problems are highly simplified by the key equivalence theorem in Shapley and Shubik (1971) [28] between the core and the set of Walrasian equilibrium allocations. On the cons side, this result is so potent that one cannot see how the market bargains to the equilibrium outcome. I personally would like to grasp what causes instability on the market: who is blocking and why. An additional argument stems from the existence of so-called “repugnant markets”, i.e. centralised matching markets where transfers are forbidden on fairness or ethical grounds. Examples include the assignment of social housing, of deceased donor organs for transplant.

Regarding the bounded rationality definition, one has to elect among a voluminous literature a concept likely to be relevant on matching markets or if not appropriate, to build a seminal bounded rationality concept. My above claim that matching markets have a peculiar structure demanding unique cognitive skills unambiguously calls for an original model. This being said, the choice remains between relaxing the assumption of full instrumental rationality or full cognitive rationality. As previously argued, the complexity of the matching environment hampers cognition more than optimisation so that my focus will be bounded cognitive rationality, keeping full instrumental rationality. Cognitive rationality embeds the ability to form correct expectations on both the states of the world in a Bayesian environment and on the other players’ strategies in a game. As my interest lies mainly on the concept of stability, that wipes out strategic aspects of the application process, I assume that the limited cognition applies to the inference on payoffs. Finally, let us remind that matching markets provide seldom personal experience but ample collective experience if the market size is large. Consequently, the temptation of an agent to reinforce his beliefs on payoffs not only considering his current payoff but others’ payoffs as well is high. Said differently, matching markets are a propitious environment for social learning. In this perspective, I choose to model imperfect social learning.

I take this opportunity to clarify the matching vocabulary. In particular, the use of the term assortativity in the theory literature is quite equivocal. Assortative preferences prevail when agents on each side of the market display positively correlated preferences on the opposite side of the market. They oppose for instance to aligned preferences (positive correlation across sides of the market) or idiosyncratic preferences (no correlation pattern). Differently, a matching is assortative if the agents sharing similar characteristics are matched together.

Throughout this report, I will constantly make use of the marriage market vocabulary, adopting the male point of view by default. Although this practise is bound to produce caricatural examples, I think resorting to a driving example still fosters understanding. This should not conceal the wider ambition of the concept and generality of the results.

1.5 Structure of the dissertation

The dissertation is structured as follows. Section §2 presents the model, defines a novel concept of bounded rationality on matching markets and revisits the pairwise stability definition accordingly. Section §3 compares the new stability notion to the standard concept defined by Gale and Shapley (1962) [15]. Section §4 discusses the value of the results and possible refinement or extensions. Section §5 summarises the findings and concludes. All proofs are provided in the appendix §6, in the order in which the results appear in the main body. The appendix also includes results for alternative but close model specifications, clarification on the relation between the model presented and the bounded rationality literature. In the appendix still, I propose new tools designed for theoretical analysis of bounded rationality on matching markets, which I happened to use throughout my work on this dissertation or plan to use in coming research.

2 The Model

2.1 Setting

2.1.1 Market structure

I set a two-sided one-to-one matching market. The two sides M and F are strategic and comprise respectively $n_M = \#M \in \mathbb{N}^*$, $n_F = \#F \in \mathbb{N}^*$ agents. On this market (M, F) , a matching is any function μ :

$$\begin{aligned} \mu &: M \cup F \longrightarrow M \cup F \cup \{\emptyset\} \\ \forall m \in M, f \in F &: \mu(m) \in F \cup \{\emptyset\}, \mu(f) \in M \cup \{\emptyset\} \\ \forall m \in M, f \in F &: \mu(m) = f \iff \mu(f) = m \end{aligned}$$

I adopt a Beckerian structure of match utilities, simplified to ban transfers. Matched agents $m \in M$ and $f \in F$ engage in productive activities and generate a surplus $s_{mf} > 0$. On a marriage market, the surplus is made of any household-produced commodity that is transferable among members of the same household but not marketable between different households. One implicit assumption is that these commodities can be combined into a single aggregate. Becker's own examples are: the quality of meals, the quality and quantity of children, prestige, recreation, companionship, love, health status. The inputs of the surplus function are k characteristics for each of the two agents c_m^1, \dots, c_m^k and c_f^1, \dots, c_f^k . To simplify notations, I store the k characteristics of an agent m in a vector $C_m = (c_m^1, \dots, c_m^k)$ (respectively $C_f = (c_f^1, \dots, c_f^k)$) for an agent f). In marriage, relevant characteristics most likely include fertility, cooperation capacities, and so on. Another simplification in comparison to Becker is that here no market goods or services enter the surplus function. It is equivalent to say they do but with zero complementarity with respect to agents' characteristics. Quite importantly, the surplus function is increasing in all arguments, which introduces a positive correlation between surpluses of all matches where a given individual is involved: $s_{mf} \not\ll s_{m'f}$, $m \neq m'$. To begin with, the precise shape of the surplus $s_{mf} = s(C_m, C_f)$ will not matter for the analysis.

Right after producing, matched agents agree on a split of the surplus: $s_{mf} = u_{mf}^m + u_{mf}^f$, where $u_{mf}^m, u_{mf}^f \in [0, s_{mf}]$ are the match utilities that agent m (respectively f), gets when he (she) is matched to f (m). To ease notations, I define a splitting rule:

$$\forall (m, f) \in M \times F : (\star) \begin{cases} u_{mf}^m = r_{mf} \cdot s_{mf} \\ u_{mf}^f = (1 - r_{mf}) \cdot s_{mf} \end{cases}$$

I define and interpret once and for all three special cases of splitting rule I will repeatedly call upon in the following:

Definition 2.1.

- (i) The splitting is **constant for both sides** M and F if: $\forall (m, f) \in M \times F, r_{mf} = r \in [0, 1]$.
- (ii) The splitting is **constant for side** M if: $\forall (m, f) \in M \times F, r_{mf} = r_m \in [0, 1]$.
- (iii) The splitting is **regular for both sides** M and F if:
 $\exists (\alpha_m)_{m \in M}, (\beta_f)_{f \in F} \forall (m, f) \in M \times F, r_{mf} = \frac{\alpha_m}{\alpha_m + \beta_f} \in [0, 1]$.

In all three cases, specifying the splitting rule for side M immediately determines the splitting rule for side F : (i) $\Rightarrow 1 - r_{mf} = 1 - r$, (ii) $\Rightarrow 1 - r_{mf} = 1 - r_m$, (iii) $\Rightarrow 1 - r_{mf} = \frac{\beta_f}{\alpha_m + \beta_f}$. The interpretation, stated for the marriage market is the following. (i) The bargaining power of a man m is constant whatever the woman he faces and similar across all men, and symmetrically for women. In other words, the individual gender is the unique determinant of the sharing of the surplus. (ii) The bargaining power of a man m is constant whatever the woman he faces. Mechanically, the bargaining power of a woman f is similar across all women. In other words, the identity of the man is the only thing that matters for the sharing of the surplus. The term r_m stands as a measure of m bargaining power over the female kind. (iii) The bargaining power of a man m against a woman f is determined by the relation between an absolute bargaining power score α_m for m and one β_f for f , and symmetrically for women.

When an agent $m \in M$ ($f \in F$) is unmatched he (she) receives a reservation utility $u_{m\emptyset}^m \geq 0$ ($u_{\emptyset f}^f \geq 0$). In marriage, it corresponds to the utility of remaining single. To unify notations, I define surpluses and splitting rules for the unmatched case: $s_{m\emptyset} = u_{m\emptyset}^m$, $r_{m\emptyset} = 1$ ($s_{\emptyset f} = u_{\emptyset f}^f$, $r_{\emptyset j} = 0$), so that the relation (\star) is verified $\forall m \in M \cap \{\emptyset\}, f \in F \cap \{\emptyset\}$.

The match utilities (equivalently the surpluses and splitting rules), and the reservation utilities are defined for all possible matches. In this work, they are considered as exogenous.

2.1.2 Information structure

I assume that information is limited. Agents know the market structure and in particular they know the surplus functional form $s_{mf} = s(\cdot, \cdot)$. Self-identity is perfect private knowledge: an agent m knows all his matching-relevant characteristics c_1^m, \dots, c_k^m .

Surpluses of currently matched agents are common knowledge. This means that any agent on the market perfectly observes the surpluses from all current matches, including the surplus from his own match if he is currently matched, but ignores the surpluses of unrealised matches, including the surpluses he would generate with an agent he is not currently matched to. This information structure makes sense in environments where the surpluses of matched agents are openly displayed. This is definitely the case in marriage regarding the quantity of children or prestige for instance, it could be discussed for health status, love. The surplus of not being matched is private knowledge unconditional on the matching. With splitting fixed to $r_{m\emptyset} = 1$, $r_{\emptyset f} = 0$, this means that in any case an agent knows his reservation utility. A possible story for this knowledge is that agents have memory. Indeed, any agent currently on the matching market was single at least once, at the precise moment in the past when he entered the market. To assess reservation utility, multi-inference is not an issue and, conditional on the reservation utility being constant over time, one observation enables perfect learning.

Splitting rules are private knowledge at the m level. That is, agents know their own splitting rules in any possible match, even if the match is not currently effective. They ignore the splitting rules of others. Justification for this ignorance could be that in most markets, the splitting rule depends on private bargaining performed away from prying eyes. In marriage, negotiation on whom will prepare diner or stop a professional career to look after the baby is intimate. On the job market, agreement on working hours and wages is laid down in a confidential contract.

In sum, using my notations, if the current match is μ , an agent $m \in M$ observes $s_{\mu(f)f}, \forall f \in F$, observes $s_{m\mu(m)}$ and knows $s_{m\emptyset}$. He also knows $r_{mf}, \forall f \in F$ and $r_{m\emptyset}$, but ignores $r_{m'f}, \forall m' \neq m$.

$m, f \in F$. Because he knows both the surplus and the splitting rule in current match $\mu(m)$ and of no match \emptyset , m is able to reconstruct $u_{m\mu(m)}^m$ and $u_{m\emptyset}^m$. Yet he cannot reconstruct $u_{\mu(f)f}^{\mu(f)}$, $f \in F$, u_{mf}^m , $f \neq \mu(m)$ due to ignorance of the splitting rule $r_{\mu(f)f}$, of the surplus s_{mf} . I denote the defined information structure by ψ and extend the matching market notation to include it: $(M, F; \psi)$.

Now the question arises: How would a fully (Bayesian) rational agent reason in such environment? At least one utility $u_{m\emptyset}^m$ is perfectly known, potentially two ($u_{m\mu(m)}^m$ if m is matched). But how would a fully rational agent estimate u_{mf}^m , $f \notin \{\emptyset, \mu(m)\}$? Two different cases can occur. In the first case, f is currently matched ($\mu(f) \neq \emptyset$). Then the observed surplus $s_{\mu(f)f}$ is a signal for m on C_f . Agent m should use his observation of $s_{\mu(f)f} = s(C_{\mu(f)}, C_f)$ to revise his prior q_f^m on agent f characteristics C_f and compute a posterior distribution p_f^m . Using his known true characteristics C_m and the posterior on f characteristics p_f^m , plus the knowledge of the surplus function, he should compute a posterior on the surplus s_{mf} . In the second case, f is currently unmatched ($\mu(f) = \emptyset$). Then, following the principle of indifference, the posterior p_f^m on s_{mf} should be a Dirac with full mass on the mean of the posterior distributions p_f^m , $\mu(f) \neq \emptyset$. Then, using the posterior distribution on surpluses and the knowledge of the splitting rules r_{mf} , $f \in F \cup \{\emptyset\}$, m should compute a posterior on the match utilities. The matching strategy should then be based on the comparison of $n_F + 1$ expected utilities $\mathbb{E}_{p_f^m}[u_{mf}^m]$, $f \in F \cup \{\emptyset\}$. Anytime a match is broken or created, Bayesian updating of the posterior on side F characteristics should happen. If the market size is large, learning on matching market operates mainly through the channel of social learning, direct reinforcement plays a minor role. A higher level of inference includes inference on the match status. Indeed, if an agent f is matched, it means that $\mu(f) \in M$ agreed to match with her in the previous period. So at least one agent on side M of the market assessed f to have good characteristics, which m values as well. This effect of inference on match status is expected to push all posteriors on match surpluses with currently matched agents up and all posteriors on surpluses with currently unmatched agents down. Overall, perfect cognition seems way too demanding in this environment we defined. That agents have the capacity to store $n_F + 1$ (vector of) posteriors at each period when n_F is large seems most unlikely. Besides, using the knowledge of the surplus functional form obliges to manipulate the k characteristic arguments for each f , causing a dimension issue. Finally, the inference on match status is quite indirect, requiring improbable sophistication.

2.2 The concept: Live-Polarised-Unidimensional-Valuation (LPUV)

2.2.1 LPUV-rationality

Unless otherwise specified, I locate on a matching market $(M, F; \psi)$, and μ is an odd matching on this market. For readability, I define the concept from the point of view of side M of the market. The concept is subsumed in the three coming definitions:

Definition 2.2. A *valuation* for side M is a function: $v : F \cup \{\emptyset\} \rightarrow \mathbb{R}^+$.

Definition 2.3. A valuation v_m is μ -*Live-Polarised-Unidimensional* (μ -LPU) for agent $m \in M$ if it writes:

$$v_m(f) = \begin{cases} u_{m\emptyset}^m & \text{if } f = \emptyset \\ r_{mf} \cdot s_{\mu(f)f} & \text{if } f \neq \emptyset, \mu(f) \neq \emptyset \\ \in \mathbb{R}_+ & \text{if } f \neq \emptyset, \mu(f) = \emptyset \end{cases}$$

Definition 2.4. An agent $m \in M$ is *Live-Polarised-Unidimensional-Valuation rational* (LPUV-rational) if for any current matching μ , there exists a μ -LPU valuation v_m such that m 's belief on match utilities is a Dirac with full mass on v_m : $\hat{u}_{mf}^m | \mu = \delta_{v_m}$.

In definition 2.2., I define a class of functions attached to one side of a matching market. With definition 2.3., I put structure on such functions. From this structure of function taken as beliefs, I build in definition 2.4. a bounded rationality type of agent on matching markets.

Let me detail each line of the formula in definition 2.3.

- First line: A LPUV-rational agent m perfectly assesses the reservation utility.
- Second line: If f is an agent ($f \neq \emptyset$), and f is currently matched ($\mu(f) \neq \emptyset$), m plugs the observed surplus generated in the match in which f is involved $s_{\mu(f)f}$ into the splitting rule that would prevail if m and f were matched r_{mf} . This line encompasses two cases:
 - **Perfect direct reinforcement:** f 's current match happen to be m , $\mu(f) = m$, and $v_m(f) = r_{mf} \cdot s_{mf} = u_{mf}^m$.
 - **Coarse indirect reinforcement or social learning:** f and m are not currently matched, $\mu(f) \neq m$, and in general $s_{\mu(f)f} \neq s_{mf}$. This case involves that m plugs an incorrect estimate of s_{mf} into the correct splitting rule r_{mf} . As a consequence, $v_m(f) \neq u_{mf}^m$ in general.
- Third line: If f is an agent ($f \neq \emptyset$), and f is currently unmatched $\mu(f) = \emptyset$, m is allowed to store any positive estimate of u_{mf}^m .

2.2.2 Interpretation

One may wonder: What is the mistake that a LPUV-rational agent does with respect to the Bayesian benchmark? Several layers of cognitive mistake are tangled in the definition.

First, a LPUV-rational agent only relies on the present market situation for inference on match utilities. If he happened to have a prior distribution on the utilities and a posterior belief forged over the course of previous market situations, he has forgotten it and he exploits nothing but the current observations and knowledge. The agent has no memory, or does not capitalise on it. Let me set an example on the marriage market. Suppose a woman f would like to assess the utility she would get from marrying man m who is currently married and has been married twice. As a LPUV-rational agent, she would only consider the surplus from this second (and still effective) marriage, for instance she would count the number of children born of this second union, overlooking children born from the first marriage. This corresponds to the “Live” part of LPUV.

Second, a LPUV-rational agent entirely omits the multiplicity of characteristics, as he stores (one) estimate of the match utility with each potential partner, rather than storing (k) estimates of the k characteristics of each potential partner. Resuming the marriage example, suppose agents are characterised by both their working capacities and thoughtfulness. If a married couple looks happy, a LPUV-rational woman infers good characteristics from the man, but does not know whether exceptional working capacities or outstanding thoughtfulness is driving this happiness. This corresponds to the “Unidimensional” part of LPUV rationality.

Third, a LPUV-rational agent does not manipulate probability distributions but point estimates. This corresponds to the “Valuation” part of LPUV-rationality.

Fourth, although direct reinforcement learning is perfect (considering a memory-less agent), social learning is coarse. While for $f \notin \{\emptyset, \mu(m)\}$, the splitting rule r_{mf} is perfectly known, the issue for a LPUV-rational agent m is rather the estimation of the surplus s_{mf} . What he takes as estimate for s_{mf} is $s_{\mu(f)f}$. As a function of individual characteristics, the true surplus and its estimate write respectively $s_{mf} = s(C_m, C_f)$ and $s_{\mu(f)f} = s(C_{\mu(f)}, C_f)$, so they differ in general. Thus, the mistake in social learning consists in neglecting the role of one’s own side of the market in the determination of the surplus. On the marriage market, imagine a woman observes a positive wealth outcome of her neighbours’ marriage: they have just fully refurbished their house, and afford holidays in expensive locations. The LPUV-rational woman credits this positive outcome exclusively to the husband: she immediately predicates that the

husband has a prominent position in a leading bank, and subsequently covets him. Imagine also that a man observes a negative offspring outcome of his best friend’s marriage: he knows from his best friend himself than he and his wife have been struggling to make a child for years, in vain. The LPUV-rational man blames bad fertility exclusively on his best friend’s wife and rejects her as a potential partner. This corresponds to the “Polarisation” part of LPUV-rationality. This neglect can be due either to a belief that there is no heterogeneity within one’s own side of the market ($\forall (m, m') \in M : C_m = C_{m'} \Rightarrow \forall (m, f) \in M \times F : C_m = C_{\mu(f)}$), or to the belief that the surplus function takes as argument only the characteristics of the agents belonging to the opposite side of the market ($s_{mf} = s(C_f)$). Extending the example, the woman described above credits good wealth outcome exclusively to her male neighbour, either because she thinks household expenses are born solely by the husband or because she believes all women earn the same wage. The man described above blames bad fertility on his best friend’s wife exclusively, either because he thinks that biologically women fertility is the only relevant parameter for the probability that a healthy child is born, or because he thinks all men feature the same fertility.

2.2.3 Two competing notions of pairwise stability

The notion of pairwise stability for matching markets (in full information) was coined by Gale and Shapley in 1962 [15]. Using my notations, it writes:

Definition 2.5. (Gale and Shapley (1962))

A matching μ is **stable** (henceforth **Gale-Shapley stable** (GS-stable)) if the two following conditions are satisfied:

- (i) (IR)^{GS}: $\forall m \in M, u_{m\emptyset}^m \leq u_{m\mu(m)}^m$, and $\forall f \in F, u_{\emptyset f}^f \leq u_{\mu(f)f}^f$.
- (ii) (IC)^{GS}: $\forall (m, f) \in M \times F, u_{mf}^m \leq u_{m\mu(m)}^m$ or $u_{mf}^f \leq u_{\mu(f)f}^f$.

The standard interpretation of GS-stability is two fold. One, as it eliminates incentives to deviate from an assignment, GS-stability promotes market peace or long-run sustainability. It also embeds some fairness criterion, often summarised as “no justified envy”. Stability is desirable both theoretically and empirically.⁷

From the notion of LPUV-rationality, I build a original concept of matching pairwise stability:

Definition 2.6. A matching μ is **Live-Polarised-Unidimensional-Valuation stable** (LPUV-stable) if there exist two families of valuations $(v_m)_{m \in M}$ and $(v_f)_{f \in F}$ such that:

- (i) $\forall m \in M, \forall f \in F, v_m, v_f$ are μ -Live-Polarised-Unidimensional for agents m, f respectively.
- (ii) (IR)^{LPUV}: $\forall m \in M, v_m(\emptyset) \leq v_m(\mu(m))$, and $\forall f \in F, v_f(\emptyset) \leq v_f(\mu(f))$.
- (iii) (IC)^{LPUV}: $\forall (m, f) \in M \times F, v_m(f) \leq v_m(\mu(m))$ or $v_f(m) \leq v_f(\mu(f))$.

Condition (i) puts the structure presented in definition 2.3. on all agents’ estimates of match utilities. Conditions (ii) and (iii) are simply the usual incentive rationality and incentive compatibility (no blocking pair) constraints under the coarse but structured estimation of match utilities. For LPUV-stability and whenever one defines stability as the usual (IR) and (IC) constraint under bounded rationality, interpretation is slightly different. The market peace or sustainability interpretation still prevails. But the fairness interpretation should now rather be “no feeling of justified envy”.

⁷In a paper from 1984 [26], Alvin Roth studied the algorithm used by the National Resident Matching Program (NMRP) clearinghouse for medical students in the US. He found that the mechanism used was closely related to Deferred Acceptance and stable. He then hypothesized that the fundamental reason for the success of the NRMP was that it produced stable matches. In the early 1990s, Roth went on to study similar medical markets in the U.K. There, he found that different regions had adopted different matching algorithms. The stable mechanisms which resulted in stable matches had turned out to be successful, whereas the other algorithms had broken down in various ways.

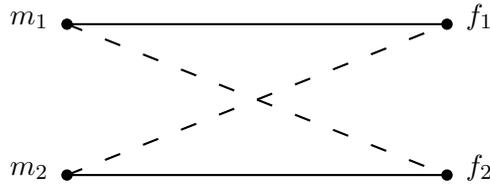
3 GS-stability and LPUV-stability: a comparative analysis

The purpose of this section is to compare the two available pairwise stability notions: Does a LPUV-stable matching always exist and is there a clear relation between GS-stable and LPUV-stable sets? When is one concept more stringent than the other, on which kind of market? Subsection §3.1 provides an answer to the former question. The two remaining subsections address the latter issue: subsection §3.2 focuses on a specific family of splitting rules delivering LPUV-stable matchings whatever the surpluses, while subsection §3.3 wonders whether the assortative matching under regular splitting is GS-stable or LPUV-stable.

In the following I call null matching and denote by μ_0 the allocation where no agent is matched: $\forall k \in M \cup F : \mu_0(k) = \emptyset$. I call full matching any allocation where all agents from at least one side of the market are matched: $\mu(F) = M$ or $\mu(M) = F$. I denote by Σ , Σ_F , Σ^{GS} , Σ^{LPUV} , Σ^{IR} respectively the set of matchings, full matchings, GS-stable matchings, LPUV-stable matchings, and incentive rational matchings. In the general case $n_M \neq n_F$, I keep on with the same notations. But whenever the market is balanced $n_M = n_F = n$, I lighten notations by numbering the market participants: $M = \{m_1, \dots, m_n\}$, $F = \{f_1, \dots, f_n\}$. I denote by $[(m_k, f_k)_{k \in \{1, \dots, n\}}]$ the matching μ where agents belonging to the same parentheses are matched together: $\forall k \in \{1, \dots, n\} : \mu(m_k) = f_k$ and where all the market participants who are not displayed are unmatched. Where unambiguous, I remove the m and f symbols from the match utility formulas: $\forall (i, j) \in \{1, \dots, n\}^2 : r_{ij} \equiv r_{m_i f_j}$, $s_{ij} \equiv s_{m_i f_j}$, $u_{i\emptyset}^i \equiv u_{m_i \emptyset}^{m_i}$, $u_{\emptyset j}^j \equiv u_{\emptyset m_j}^{m_j}$, and under regular splitting rule: $\forall (i, j) \in \{1, \dots, n\}^2 : \alpha_i \equiv \alpha_{m_i}$, $\beta_j \equiv \beta_{f_j}$.

3.1 Existence and characterisation of stable matchings

Here, I recall existence of GS-stable matching and derive existence and a convenient characterisation for the brand new concept of LPUV-stability. This may prove tedious at some point, but this stage is unavoidable to master the comparison between GS and LPUV stabilities. Before proceeding to the analysis, let me make a last general remark. To be precise I introduce notations, setting two pairs of matched agents (m_1, f_1) , $\mu(m_1) = f_1$ and (m_2, f_2) , $\mu(m_2) = f_2$. On the figure just below (plus in all the following), I bind the matched pairs with full line, and the potential blocking pairs with dotted line:



From the figure, it is obvious that $(IC)_{(m_1, f_2)}^{GS}$, $(IC)_{(m_2, f_1)}^{GS}$, $((IC)_{(m_1, f_2)}^{LPUV}$, $(IC)_{(m_2, f_1)}^{LPUV})$ feature common parameters of splitting and surpluses, and as such, they are dependent. This dependence between incentive compatibility constraints is what will drive characterisation of stable matchings.

3.1.1 Existence

From Gale and Shapley (1962) [15], it is common knowledge that a GS-stable matching always exists. The proof is constructive; the authors propose the Deferred Acceptance mechanism, which, applied to any structure of preferences, delivers a stable matching. The exact formulation is:

Theorem 3.1. (Gale and Shapley (1962))

The Deferred Acceptance mechanism applied to any matching market results in a stable matching.

Corollary 3.2. (Gale and Shapley (1962))

There always exists a GS-stable matching: $\Sigma^{GS} \neq \emptyset$.

Oppositely, existence of LPUV-stable matching is supported by a non-constructive argument:

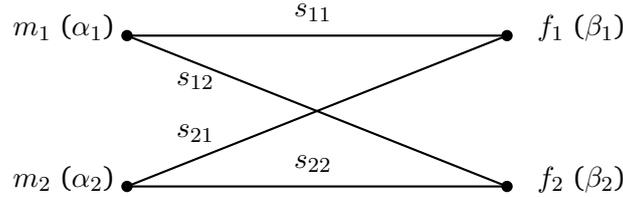
Theorem 3.3. μ_0 is LPUV-stable: $\mu_0 \in \Sigma^{LPUV}$.

The proof is immediate. Because no woman is matched under the null matching, $(\mu_0(F) = \emptyset)$, any valuation $v_m : F \cup \{\emptyset\} \rightarrow \mathbb{R}_+$ s.t. $v_m(\emptyset) = u_{m\emptyset}^m$ is LPU. In particular v_m s.t. $v_m(\emptyset) = u_{m\emptyset}^m, \forall f \in F : v_m(f) = 0$ is a LPU-valuation. Under v_m , m is not willing to match. Generalising the argument, not a single man is willing to match and there is no blocking pair. The (IC) condition is satisfied. The (IR) for either side is not limiting and trivially satisfied. This directly implies existence:

Corollary 3.4. There always exists a LPUV-stable matching: $\Sigma^{LPUV} \neq \emptyset$.

3.1.2 Examples on a 2 x 2 market

My aim here is simply to illustrate that the relation between the sets of GS-stable and LPUV-stable matchings is far from being trivial. In particular, no inclusion or separation result can be found. In order to back this statement, I display a series of counter-examples on a 2×2 market under regular splitting rule. The notations are recalled in the following graph, where the reader will find bargaining scores in parenthesis next to the individual identifiers, and surpluses on the lines connecting two agents:



I name the two full matchings: $\mu = [(m_1, f_1); (m_2, f_2)]$ and $\mu' = [(m_1, f_2); (m_2, f_1)]$. I set the four reservation utilities to zero ($\forall (i, j) \in \{1, 2\}^2 u_{i\emptyset}^i = u_{\emptyset j}^j = 0$), whereas all four surpluses are strictly positive ($\forall (i, j) \in \{1, 2\}^2, s_{ij} > 0$), and a matched agent always receives a strictly positive share of the generated surplus ($\forall (i, j) \in \{1, 2\}^2, \alpha_i, \beta_j > 0$).

A few general observations on the specificities of this market are worth making. They derive from the two assumptions we made, namely the normalisation of reservation utilities to zero, and the equal number of agents on each side of the market. The following two propositions will considerably simplify the treatment of examples:

Proposition 3.5. Set $\#M = \#F = n, \forall (i, j) \in \{1, \dots, n\}^2 u_{i\emptyset}^i = u_{\emptyset j}^j = 0, s_{ij} > 0$.

Then, GS-stable matchings are full: $\Sigma^{GS} \subseteq \Sigma_F$.

This result goes without formal proof. Due to the normalisation of reservation utilities to zero, match utilities are always strictly greater than reservation utilities. There is a positive value of being matched whoever the partner may be. Because the number of agents is equal on both sides of the market, a single agent will always find a partner. So: $\Sigma^{GS} \subseteq \Sigma_F$. Proposition 3.5. applied to a 2×2 matching market claims that μ GS-unstable $\Rightarrow \mu'$ GS-stable, and by contraposition, μ' GS-unstable $\Rightarrow \mu$ GS-stable.

Proposition 3.6. Set $\#M = \#F = n, \forall (i, j) \in \{1, \dots, n\}^2 u_{i\emptyset}^i = u_{\emptyset j}^j = 0, s_{ij} > 0$.

Any matching where at most one pair of agents is matched is LPUV-stable.

I already proved that $\mu_0 \in \Sigma^{LPUV}$. Now suppose one pair of agents, say (m^*, f^*) is matched while all other agents are unmatched, using my notations, the match is $[(m^*, f^*)]$. Set a negative

prejudice of all agents on unmatched agents: $\forall (m, f) \in M \times F, \forall m' \neq m^*, f' \neq f^* : v_m(f') = v_f(m') = 0$. m^* and f^* never engage in a blocking pair. The matching $[(m^*, f^*)]$ is LPUV-stable. Proposition 3.6. applied to a 2×2 market claims that any non-full matching is LPUV-stable: $\Sigma - \{\mu, \mu'\} \subseteq \Sigma^{LPUV}$.

All in all the two latter properties suggest that LPUV-stability is biased towards matchings where a low number of agents are matched in comparison to GS-stability. This being said, the stability of full matchings remains a pending question. Under full matching, the reservation utilities do not matter and the effect of the distortion of LPUV-stability towards the null matching vanishes.

Below, I give two numerical examples, specifying bargaining scores and surpluses, where the two sets of full stable matchings are disjoint: $\Sigma_F^{GS} \cap \Sigma_F^{LPUV} = \emptyset$.

- a. $(\alpha_1, \alpha_2, \beta_1, \beta_2) = (7, 1, 1, 1), (s_{11}, s_{22}, s_{12}, s_{21}) = (5, 1, 3, 2)$
 In μ , m_2 gets $\frac{1}{1+1} \cdot 1 = \frac{1}{2}$, f_1 gets $\frac{1}{7+1} \cdot 5 = \frac{5}{8}$. By matching together, m_2 and f_1 would get respectively $\frac{1}{1+1} \cdot 2 = 1 > \frac{1}{2}$, $\frac{1}{1+1} \cdot 2 = 1 > \frac{5}{8}$. (m_2, f_1) is a GS-blocking pair. μ is GS-unstable. By proposition 3.5. for 2×2 markets, μ' is GS-stable. In μ still, m_1 and f_1 get respectively $\frac{7}{7+1} \cdot 5 = \frac{35}{8}$ and $\frac{5}{8}$. If they are LPUV-rational, m_1, f_1 believe they would get respectively $\frac{7}{7+1} \cdot 1 = \frac{7}{8} < \frac{35}{8}$ and $\frac{1}{1+1} \cdot 1 = \frac{1}{2} < \frac{5}{8}$ by blocking with f_2, m_2 respectively. m_1 and f_1 never want to block, equivalently there is no LPUV-blocking pair and μ is LPUV-stable. In μ' , m_2 and f_2 get respectively $\frac{1}{1+1} \cdot 2 = 1$ and $\frac{1}{7+1} \cdot 3 = \frac{3}{8}$. If they are LPUV-rational, they believe they would get respectively $\frac{1}{1+1} \cdot 3 = \frac{3}{2} > 1$ and $\frac{1}{1+1} \cdot 2 = 1 > \frac{3}{8}$ by matching together. (m_2, f_2) is a LPUV-blocking pair. μ' is unstable.
 $\Sigma_F^{GS} = \{\mu'\}, \Sigma_F^{LPUV} = \{\mu\}$.
- b. $(\alpha_1, \alpha_2, \beta_1, \beta_2) = (2, 1, 1, 1), (s_{11}, s_{22}, s_{12}, s_{21}) = (5, 4, 4, 3)$.
 In μ , m_2 and f_1 get respectively $\frac{1}{1+1} \cdot 4 = 2$ and $\frac{1}{2+1} \cdot 5 = \frac{5}{3}$. If they are LPUV-rational, they believe they would get $\frac{1}{1+1} \cdot 5 = \frac{5}{2} > 2$ and $\frac{1}{1+1} \cdot 4 = 2 > \frac{5}{3}$ by matching together. (m_2, f_1) is a LPUV-blocking pair. μ is LPUV-unstable. In μ' , m_2 and f_2 get respectively $\frac{1}{1+1} \cdot 3 = \frac{3}{2}$ and $\frac{1}{2+1} \cdot 4 = \frac{4}{3}$. If they are LPUV-rational, they believe they would get $\frac{1}{1+1} \cdot 4 = 2 > \frac{3}{2}$ and $\frac{1}{1+1} \cdot 3 = \frac{3}{2} > \frac{4}{3}$. (m_2, f_2) is a blocking pair. μ' is LPUV-unstable.
 $\Sigma_F^{GS} \neq \emptyset, \Sigma_F^{LPUV} = \emptyset$

This case can be contrasted with the situation when full GS-stable and LPUV-stable sets exactly coincide:

- c. $(\alpha_1, \alpha_2, \beta_1, \beta_2) = (1, 2, 1, 2), (s_{11}, s_{22}, s_{12}, s_{21}) = (1, 1, 1, 1)$.
 In μ' , m_1 and f_1 both get $\frac{1}{1+2} \cdot 1 = \frac{1}{3}$. By matching together, they would both get $\frac{1}{1+1} \cdot 1 = \frac{1}{2} > \frac{1}{3}$. So (m_1, f_1) is a GS-blocking pair. μ' is GS-unstable. By proposition 3.5. for 2×2 markets, μ is GS-stable. In μ' , if m_1 and f_1 are LPUV-rational, they rightly believe they would both get $\frac{1}{1+2} \cdot 1 = \frac{1}{2} > \frac{1}{3}$ by matching together. (m_1, f_1) is a blocking pair. μ' is LPUV-unstable. In μ , m_1 and f_1 both get $\frac{1}{1} \cdot 1 = \frac{1}{2}$. If they are LPUV-rational, they rightly believe that would get $\frac{1}{1+2} \cdot 1 = \frac{1}{3} < \frac{1}{2}$ by blocking with f_2, m_2 respectively. m_1 and f_1 never want to block, equivalently there is no LPUV-blocking pair and μ is LPUV-stable.
 $\Sigma_F^{GS} = \Sigma_F^{LPUV} = \{\mu\}$.

I generalise example c. in proposition 3.14. in subsection §3.3.1.

3.1.3 Main characterisation of LPUV-stable matchings

The next proposition provides a convenient characterisation of LPUV-stable matchings:

Proposition 3.7.

$$\mu \in \Sigma^{LPUV} \iff \begin{cases} \mu \text{ satisfies } (IR)^{GS}. \\ \mu \text{ satisfies } (IC)^{LPUV} \text{ for pairs where both agents are currently matched.} \end{cases}$$

Proof

The first part of the proof relies on the knowledge of the reservation utility (line 1 of definition 2.3.) and of the utility of current match, causing $(IR)^{GS}$ and $(IR)^{LPUV}$ to be equivalent. The second part relies on the unstructured estimate of the surplus that would be generated with a currently unmatched agent: line 3 of definition 2.3. imposes no constraint. Indeed, considering a pair (m, f) where f is unmatched, $v_m(f) \leq r_{m\mu(m)}u_{m\mu(m)}$ is a sufficient condition for (m, f) not to block.

The nice implications are: one, Σ^{GS} and Σ^{LPUV} can be compared only based on the (IC) condition, second, $(IR)^{LPUV}$ needs to be checked only for potential blocking pairs where both agents are currently matched. The (IC) condition for LPUV-stability now writes:

$$(IC)^{LPUV} : \forall (m, f) \in \mu(F) \times \mu(M) : \begin{cases} r_{m\mu(m)} \cdot s_{m\mu(m)} \geq r_{mf} \cdot s_{\mu(f)f} \\ \text{or} \\ r_{\mu(f)f} \cdot s_{\mu(f)f} \geq (1 - r_{mf}) \cdot s_{m\mu(m)} \end{cases}$$

From the formula just above, it is worth noticing that $(IC)^{LPUV}$ compares only the two realised surpluses $s_{m\mu(m)}$ and $s_{\mu(f)f}$. $(IC)^{GS}$ on the contrary compares three distinct surpluses: the two realised surplus $s_{m\mu(m)}$ and $s_{\mu(f)f}$ plus the unrealised surplus s_{mf} , $\mu(m) \neq f$. This has again two consequences. One, matchings where the realised surpluses are the highest will be more often GS-stable than matchings where the unrealised surpluses are the highest, whereas LPUV-stability remains insensitive to this comparative statistic. Thus, the sum of surpluses is on average higher under GS-stability than under LPUV-stability. Taking a shortcut, I expect GS-stable matchings to be welfare superior to LPUV-stable matchings. Second, imagine a matching where all agents are matched. Because the $(IC)^{GS}$ features the same number of inequalities than in $(IC)^{LPUV}$ with additional parameters of unrealised surpluses, it is on expectation ⁸ more easily satisfied than $(IC)^{LPUV}$. For this reason, matchings where all agents are matched tend to be more GS-stable than LPUV-stable. Said differently and in comparative terms, GS-stability leans towards full matchings, whereas LPUV-stability leans towards the null matching.

3.1.4 Necessary conditions for LPUV-blocking pairs

I now characterise LPUV-blocking pairs:

Proposition 3.8.

- (i) (m, f) is a LPUV-blocking pair for $\mu \Rightarrow r_{mf} \in (0, r_{\mu(f)f}) \cup (r_{m\mu(m)}, 1)$.
- (ii) (m, f) is a LPUV-blocking pair for μ , $s_{m\mu(m)} > s_{\mu(f)f} \Rightarrow r_{m\mu(m)} < r_{mf}$
- (iii) $r_{m\mu(m)} < r_{mf} < r_{\mu(f)f} \Rightarrow (m, f)$ is not a LPUV-blocking pair for μ .

Proof

(i) gives a necessary condition for the pair (m, f) to be blocking in μ . Rewriting the condition as $r_{mf} > r_{m\mu(m)}$ or $1 - r_{mf} > 1 - r_{\mu(f)f}$ provides an immediate interpretation: in a blocking pair, at least one of the two agents improves his splitting rule. The intuition is quite direct. From $(IC)^{LPUV}$, one sees that m and f make their blocking decision as if they would exchange their current surpluses $s_{m\mu(m)}$ and $s_{\mu(f)f}$. Among the 2 surpluses, one is higher than the other. So the one agent in the blocking pair who agrees on getting a smaller surplus than he currently has must at least increase his share of the surplus to be incentivised to block. This is exactly what (ii) states. (iii) is just the contraposition of (i). It means that if both agents m and f decrease

⁸On expectation if the surpluses are identically and independently distributed (i.i.d.).

their splitting rule by mating together, they cannot form a blocking pair (whatever the surpluses may be).

Proposition 3.8. applied to the three families of splitting rules I introduced in definition 2.1., subsection §2.1.1 delivers the following:

Corollary 3.9.

- (i) The splitting is constant for both sides $\Rightarrow \Sigma^{LPUV} = \Sigma^{IR}$.
- (ii) The splitting is constant for side M , (m, f) LPUV blocking pair $\Rightarrow r_m < r_{\mu(f)}$.
- (iii) The splitting is regular with bargaining power scores $(\alpha_m)_{m \in M}$, $(\beta_f)_{f \in F}$.
 (m, f) LPUV blocking pair $\Rightarrow \alpha_m < \alpha_{\mu(f)}$ or $\beta_f < \beta_{\mu(m)}$.
 If additionally $s_{m\mu(m)} > s_{\mu(f)f}$, then: $\beta_f < \beta_{\mu(m)}$.

Proof The intuition is quite direct. (i) Under constant splitting for both sides, the unique motivation to deviate is the perspective of a higher surplus. So for any potential blocking pair, there is exactly one agent who wishes to deviate (the one with the lowest current surplus) and one who does not want. There is no blocking pair where both agents are matched ((IC)^{LPUV}), which, according to Proposition 3.7. and conditional on (IR), guarantees LPUV-stability. (ii) Remind the general principle that for a pair to be blocking, at least one agent should increase his share of the surplus. Because m splitting rule is r_m across $f \in F$, m keeps his share constant. So f must increase her share of the surplus, that is $1 - r_{\mu(f)} < 1 - r_m \iff r_m < r_{\mu(f)}$. (iii) Under regular splitting rule, from m 's point of view α_m is fixed. In contrast, $\frac{\alpha_m}{\alpha_m + \beta_f}$ is decreasing in β_f . The only way m can increase his surplus is by matching to a weaker agent than her current match $\mu(m)$: $\beta_f < \beta_{\mu(m)}$. This condition needs always be realised when m goes to a lower surplus $s_{m\mu(m)} > s_{\mu(f)f}$ by blocking.

3.2 Robust or crosswise conditions on splitting and surplus for stability

From now on, I consider only the case when the market is balanced: $\#M = \#F = n \in \mathbb{N}$. I also focus on the set of full matchings, which is often the most interesting case empirically. Unless otherwise specified, I always consider the matching $\mu = [(m_k, f_k)_{k \in \{1, \dots, n\}}]$.

3.2.1 Robust or crosswise conditions under odd splitting rule

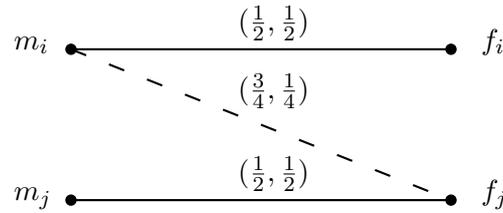
The coming theorem gives necessary but also sufficient conditions for LPUV-stability:

Theorem 3.10.

- (i) μ is LPUV-stable $\iff \forall 1 \leq i < j \leq n : \frac{s_{ii}}{s_{jj}} \in \mathbb{R}_+^* \setminus \left(\frac{1 - r_{jj}}{1 - r_{ij}}, \frac{r_{ij}}{r_{ii}} \right) \cap \left(\frac{r_{jj}}{r_{ji}}, \frac{1 - r_{ji}}{1 - r_{ii}} \right)$
- (ii) μ is LPUV-stable if and only if, $\forall 1 \leq i < j \leq n$, one of the three conditions below is satisfied:
 - (1) $\begin{cases} r_{ii}(1 - r_{jj}) > r_{ij}(1 - r_{ij}) \\ r_{jj}(1 - r_{ii}) > r_{ji}(1 - r_{ji}) \end{cases}$
 - (2) a. $\begin{cases} r_{ii}(1 - r_{jj}) > r_{ij}(1 - r_{ij}) \\ r_{jj}(1 - r_{ii}) \leq r_{ji}(1 - r_{ji}) \end{cases}$ and $\frac{s_{ii}}{s_{jj}} \in \left(0, \frac{r_{jj}}{r_{ji}} \right] \cup \left[\frac{1 - r_{ji}}{1 - r_{ii}}, +\infty \right)$
 - b. $\begin{cases} r_{ii}(1 - r_{jj}) \leq r_{ij}(1 - r_{ij}) \\ (1 - r_{ii})r_{jj} > (1 - r_{ij})r_{ji} \end{cases}$ and $\frac{s_{ii}}{s_{jj}} \in \left(0, \frac{1 - r_{jj}}{1 - r_{ij}} \right] \cup \left[\frac{r_{ij}}{r_{ii}}, +\infty \right)$
 - (3) $\begin{cases} r_{ii}(1 - r_{jj}) \leq r_{ij}(1 - r_{ij}) \\ r_{jj}(1 - r_{ii}) \leq r_{ji}(1 - r_{ji}) \end{cases}$ and $\frac{s_{ii}}{s_{jj}} \in \left(0, \min \left(\frac{1 - r_{jj}}{1 - r_{ij}}, \frac{r_{jj}}{r_{ji}} \right) \right] \cup \left[\max \left(\frac{r_{ij}}{r_{ii}}, \frac{1 - r_{ji}}{1 - r_{ii}} \right), +\infty \right)$

Proof

The disjunction of cases in (ii) is essential, even if tough to interpret in the most general setting. What is striking is that case (1) features two inequalities involving only the splitting rules but not the surpluses. In other words, there exist systems of splitting rules that guarantee matching μ LPUV-stability whatever the surpluses. Said differently, under a splitting system complying with case (1), μ LPUV-stability is robust to any change in the surplus system. Looking more closely at the two inequalities, I take note of the following. On the left (right) side of the inequalities appear the realised (unrealised) splitting. The direction of the inequality is intuitive. On the first (second line) of the conditions, appear the shares of the surpluses that pairs (m_i, f_j) $((m_j, f_i))$ currently get, and would get if blocking. That is to say, one line corresponds to one potential blocking pair, which again makes sense. What is trickier to interpret is the multiplicative pattern: $r_{ii}(1 - r_{jj})$ multiplies the shares of the surpluses that m_i and f_j currently get. If one of the two terms is close to zero, $r_{ii}(1 - r_{jj}) \simeq 0$ and (m_i, f_j) is likely to block. The intuition is that at least one agent in (m_i, f_j) , say m_i is currently suffering from a very bad deal, he or she is willing to block to get a better deal, even if it involves switching to a lower surplus. There exists some system of surpluses (intuitively, s_{ii} (s_{jj}) must be high (low)) for which f_j would also be willing to block. If one of the two terms approaches one, (m_i, f_j) may not be blocking but $r_{jj}(1 - r_{ii}) \simeq 0$ and (m_j, f_i) is likely to block. It results that $r_{ii}, r_{jj} \simeq \frac{1}{2}$ favors LPUV-stability. Similarly, $r_{ij}, r_{ji} \simeq 0$ or 1 favors stability. Let me give a numerical example. I draw below 4 agents on a matching market and indicate the splitting rules $(r_{ij}, 1 - r_{ij})$ on the line connecting two agents:



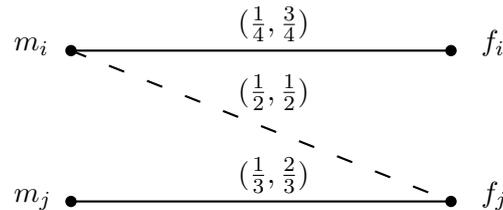
It can be checked that the splitting system does comply with case 1. I write the expected benefit of forming a blocking pair as a function of the surpluses:

$$\begin{aligned} \Delta_{m_i \rightarrow f_j} &= \frac{3}{4}s_{jj} - \frac{1}{2}s_{ii} & \Delta_{f_j \rightarrow m_i} &= \frac{1}{4}s_{ii} - \frac{1}{2}s_{jj} \\ \Delta_{m_i \rightarrow f_j} > 0 &\Rightarrow \Delta_{f_j \rightarrow m_i} < 0 \end{aligned}$$

One of the expected benefit of blocking is negative, and this is true independently of the surpluses.

In sum, condition (1) means that a matching where the split of the surpluses between mates is fairer than the split between potential blocking pairs is LPUV-stable.

Otherwise, if in at least one matched pair the split is unfair, the squeezed agent wants to block and there exists a surplus system for which he will find a willing blocking partner. This is exactly what condition (2) states. In condition (3), in both potential blocking pairs, one agent is squeezed. I again give a numerical example:



This example falls into case (3). Let me consider the potential blocking pair (m_i, f_j) . m_i feels squeezed in his current match and is willing to block with f_j , even if it involves switching to a

slightly lower surplus. To incentivise f_j to block which entails giving up a share of the surplus, one indeed needs $s_{ii} > s_{jj}$. One can check $(s_{ii}, s_{jj}) = (3, 2)$ works.

Let me further notice that condition (2) and (3) are always satisfied as soon as the ratio $\frac{s_{ii}}{s_{jj}}$ goes to zero or to the infinity. The rationale is that if there is a large imbalance between the two current surpluses, the two agents generating the higher surplus will never block, which is a sufficient condition for LPUV-stability.

3.2.2 Robust or crosswise conditions under regular splitting

In this subsection, I produce a more precise interpretation for the special case of regular splitting rules. Let me justify once for all my appealing to this specific but telling splitting structure. On the analytical side, regular splitting much simplifies the writing of necessary and sufficient conditions for stability. On the descriptive side, it displays desirable properties such as symmetry, and monotonicity:

$$\frac{\partial r_{ij}}{\partial \alpha_i} = \frac{\beta_j}{(\alpha_i + \beta_j)^2} > 0 \quad \frac{\partial r_{ij}}{\partial \beta_j} = -\frac{\beta_j}{(\alpha_i + \beta_j)^2} < 0$$

The share of the surplus that comes down to agent i is strictly increasing (strictly decreasing) in i 's (in mate j 's) bargaining score. Finally, as the reader will soon find out, LUPV-stability under regular splitting has a promising connection with assortativity.

Henceforth, and unless otherwise specified, “assortative matching”, refers to “perfect positive assortative full matching with respect to the bargaining score”. I denote it μ^a . Namely, if the two each side of the market is rank according to bargaining scores: $\alpha_1 > \dots > \alpha_n$, $\beta_1 > \dots > \beta_n$, then $\mu^a = [(m_k, f_k)_{k \in \{1, \dots, n\}}]$. By default, μ without subscript remain an odd matching.

GS-stable matchings must comply with $\frac{(n-1)n}{2}$ inequalities:

Theorem 3.11.

μ is GS-stable $\iff \forall 1 \leq i < j \leq n : \max\left(\frac{s_{ij}}{\alpha_i + \beta_j}; \frac{s_{ji}}{\alpha_j + \beta_i}\right) \leq \max\left(\frac{s_{ii}}{\alpha_i + \beta_i}; \frac{s_{jj}}{\alpha_j + \beta_j}\right)$.

Proof

This characterisation already provides a clear insight on the requirements for GS-stability. On the left (right) side of the \leq sign are the unrealised (realised). The realised and unrealised surpluses do not compare in a direct fashion but through maximum. A consequence is that a large positive gap between (at least one of) the realised surpluses, say s_{ii} and both unrealised surpluses (s_{ij}, s_{ji}) favors GS-stability. The splitting rule strengthens or qualifies this condition. If $s_{ii} > s_{jj}$ for instance, $\alpha_i + \beta_i < \alpha_j + \beta_j$ favors GS-stability. That is, it is desirable that the agents with the higher realised surpluses are also the weaker ones. This ensures that in at least one matched pair (here (α_i, β_i)), no agent is willing to deviate. This is enough to guarantee that neither (α_i, β_j) nor (α_j, β_i) are GS-blocking. Conversely, if $s_{ij} > s_{ji}$, then $\alpha_i + \beta_j > \alpha_j + \beta_i$ favors GS-stability. That is, it is desirable that the blocking pair with the higher (unrealised) surplus associate the stronger agents. This balances the incentives to block for each blocking pair, settling them at a low level. In a nutshell, GS-stability requires that the gap between one of the realised surpluses and the unrealised surpluses is positive, and/or⁹ that the agents generating the largest realised surpluses are on average weaker than the agents producing the lowest realised surpluses, and/or⁹ that the potential blocking pairs with the highest surpluses comprise the strongest agents. Unsurprisingly, the characterisation for LPUV-stability demands a more thorough analysis:

⁹The three constraints are obviously dependent, as opposed to case (1) in LPUV-stability robust to any surplus system. If one condition is loosely satisfied (for example, all surpluses are close in values), then the others need to be tightly enforced to guarantee GS-stability.

Theorem 3.12.

(i) μ is LPUV-stable $\iff \forall 1 \leq i < j \leq n : \frac{s_{ii}}{s_{jj}} \in \mathbb{R}_+^* \setminus \left(\frac{\alpha_j + \beta_i}{\alpha_j + \beta_j}, \frac{\alpha_i + \beta_i}{\alpha_j + \beta_i} \right) \cap \left(\frac{\alpha_i + \beta_j}{\alpha_j + \beta_j}, \frac{\alpha_i + \beta_i}{\alpha_i + \beta_j} \right)$

(ii) μ is LPUV-stable if and only if $\forall 1 \leq i < j \leq n$, one of the two conditions below is satisfied:

(1) Robust splitting conditions: $(\alpha_i + \beta_i)(\alpha_j + \beta_j) \leq \min((\alpha_i + \beta_j)^2, (\alpha_j + \beta_i)^2)$

- Positive assortativity in bargaining scores within matched pairs
- “Intermediate” bargaining scores within sides of the market, neither too homogeneous, nor too heterogeneous
- ...

Then: $(1^*) : \begin{cases} \alpha_i \leq \alpha_j \\ \beta_i \leq \beta_j \end{cases} \text{ or } \begin{cases} \alpha_j \leq \alpha_i \\ \beta_j \leq \beta_i \end{cases}$

Equivalently: $(1^{**}) : (\alpha_i + \beta_i)(\alpha_j + \beta_j) < (\alpha_j + \beta_i)(\alpha_i + \beta_j)$

(2) Crosswise splitting and surplus conditions:

- Negative assortativity in bargaining scores within matched pairs
- Imbalance in surpluses within matched pairs
- ...

a. $\begin{cases} (\alpha_i + \beta_i)(\alpha_j + \beta_j) \leq (\alpha_j + \beta_i)^2 \\ (\alpha_i + \beta_i)(\alpha_j + \beta_j) > (\alpha_i + \beta_j)^2 \end{cases} \text{ and } \frac{s_{ii}}{s_{jj}} \in \left(0, \frac{\alpha_i + \beta_j}{\alpha_j + \beta_j} \right] \cup \left[\frac{\alpha_i + \beta_i}{\alpha_i + \beta_j}, +\infty \right)$

Then: $(2a^*) : \begin{cases} \alpha_i \leq \alpha_j \\ \beta_j \leq \beta_i \end{cases}$

b. $\begin{cases} (\alpha_i + \beta_i)(\alpha_j + \beta_j) \leq (\alpha_i + \beta_j)^2 \\ (\alpha_i + \beta_i)(\alpha_j + \beta_j) > (\alpha_j + \beta_i)^2 \end{cases} \text{ and } \frac{s_{ii}}{s_{jj}} \in \left(0, \frac{\alpha_j + \beta_i}{\alpha_j + \beta_j} \right] \cup \left[\frac{\alpha_i + \beta_i}{\alpha_j + \beta_i}, +\infty \right)$

Then $(2b^*) : \begin{cases} \alpha_j \leq \alpha_i \\ \beta_i \leq \beta_j \end{cases}$

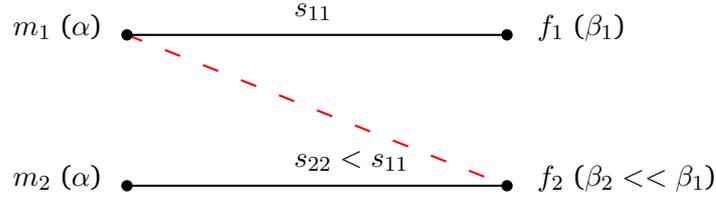
Proof

The proof is done by applying the general result in theorem 3.10. to the regular splitting case, and showing that case (3) in the mentioned theorem is impossible in this setting, hence the two remaining cases. The reason is that regular splitting rule imposes a tighter dependence structure on the splitting rule, and forbids pairs $(m_i, f_j), (m_j, f_i)$ to simultaneously block.

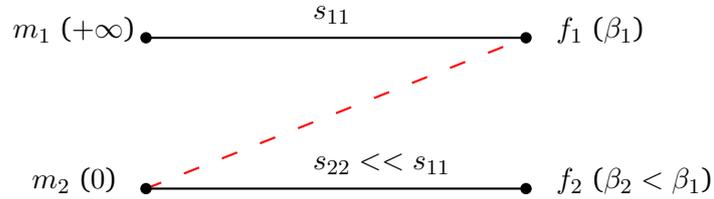
In case (1), the splitting system ensures LPUV-stability whatever the surpluses may be. Necessary conditions (1^*) and (1^{**}) are essential for interpretation of this robust case. They impose (weak) positive assortativity in individual bargaining scores: considering any set of two mached pairs, the stronger man is matched to the stronger woman and conversely. The fact that the fairness requirement in the general case translates into positive assortativity in the regular case is sensible. Indeed, given a profile of bargaining power scores, positive assortativity minimises the imbalance in splitting between matched pairs. Actually, the constraints (1^*) and (1^{**}) are equivalent (proof available [here](#)). In spite of this documented equivalence, it is worth rewriting (1^*) as (1^{**}) . Indeed, (1^*) is literal positive assortativity, whereas (1^{**}) shows the implicit constraints that assortativity imposes on the pattern of bargaining scores. In appendix §6.1, I develop a functional analysis (available [here](#)) based on (1^{**}) . Positive assortativity imposes “not-too-positive” assortativity within potential blocking pairs, and heterogeneity in average bargaining scores between the different currently matched pairs.

Now the question arises: because (1^*) is only a necessary, not sufficient condition for case (1), what else than positive assortativity is expected of bargaining scores for robust LPUV-stability? Considering relative positions of bargaining scores within the same sides of the market, I look at the two extreme cases when scores are perfectly homogeneous ($\alpha_i = \alpha_j$), and when they

are perfectly heterogeneous ($\alpha_i = 0$, $\alpha_j = +\infty$). Algebraic proof (available [here](#)) shows that in either case, condition (1) is not satisfied. I propose an intuition on a 2×2 market where the current matching is $\mu = [(m_1, f_1); (m_2, f_2)]$. Without loss of generality, I order the female bargaining scores, $\beta_1 > \beta_2$. Suppose first that men scores are perfectly homogeneous: $\exists \alpha \in (0, 1)$ s.t. $\alpha_1 = \alpha_2 = \alpha$. Here assortative matching is meaningless and positive assortativity is trivially satisfied. Then, for any of the two women f_j , $j \in \{1, 2\}$, whatever her mate m_i , $i \in \{1, 2\}$, she gets a fixed share of the surplus, $1 - r_{ij} = \frac{\beta_j}{\alpha + \beta_j}$. So the women only consider discrepancies in surpluses when blocking. Suppose s_{11} is slightly above s_{22} . Then f_2 wants to block with m_1 because she believes she increases her surplus by a little. m_1 is also willing to block because he believes he increases his share of the surplus by a lot while decreasing his surplus only by a little. I have therefore found a surplus system for which (m_1, f_2) form a blocking pair, and μ is LPUV-unstable. I summarise the splitting and surplus pattern on the following figure:



Now suppose that men scores are perfectly heterogeneous: $\alpha_1 = 0$, $\alpha_2 = +\infty$, and thanks to $\beta_1 > \beta_2$, positive assortativity is verified. Because her bargaining score is negligible compared to her match m_1 's score, f_1 currently gets a zero share of the surplus s_{11} . She would thus agree on any block where the new partner has a finite bargaining score, in particular, she is willing to block with m_2 . Set $s_{11} \gg s_{22}$. Then m_2 wants to block with f_1 , because he expects an major increase in his surplus jointly with a minor decrease of his share of the surplus. I have therefore found a surplus system for which (m_2, f_1) form a blocking pair, and μ is LPUV-unstable. I summarise the splitting and surplus pattern on the following figure:



The intuition for the robust case (1) is clear if one considers the two surplus specification extreme cases. Suppose that in $\mu = [(m_1, f_1); (m_2, f_2)]$, the two realised surpluses are very discrepant, say $s_{m_1 f_1} = +\infty$ and $s_{m_2 f_2} = 0$. Then, for rather “intermediate” bargaining scores within sides of the market, neither too homogeneous, nor too heterogeneous, and in particular for finite scores, m and f never engage in a blocking pair. If, quite the contrary, the two realised surpluses are equal $s_{mf} = s_{m'f'}$, then assortativity guarantees that the two weakest agents never engage in a blocking pair. For intermediate surplus specification, the splitting are such that at least on of either effect (intermediate scores within sides of the market or assortativity is at play).

I claim that my interest in case (1) is justified as this situation is far from being anecdotal. To understand why, let me remark that theorem 3.12. authorises the absence of blocking pairs to be sustained, for some groups of 2×2 agents by the robust condition (1), for some other groups by crosswise conditions (2a) or (2b). Conceptually, it could be that LPUV-stability of a matching on a market of odd size $n \times n$, $n \geq 3$ would never be fully sustained by condition (1), but only by alternating conditions (1) and (2a) or (2b). A simple example should wash out such doubts. When $\alpha_1 = \beta_1 < \dots < \alpha_n = \beta_n$, condition (1) is verified for any potential blocking pairs in

the assortative matching $\mu^a = [(\alpha_k, \beta_k)_{k \in \{1, \dots, n\}}]$. This example I develop in proposition 3.14., along with formal proof and succinct intuition.

Let me now consider case (2), for example sub-case (2a). Necessary condition (2*) conveys (weak) negative assortativity: considering any set of two matched pairs, the stronger man α_j is matched to the weaker woman β_j and conversely. There is thus one potential blocking pair (α_i, β_j) bringing together weak agents $\alpha_i + \beta_j < \alpha_j + \beta_i$, which is incentivised to form under some surplus system. The constraint on the realised surplus ratio ensures that this does not occur, by way of a large imbalance between surpluses.

3.2.3 Graphical representations and simulations

To better understand the conditions for LPUV-stability I represent the set of bargaining scores and surpluses complying with each of the three cases graphically. Needless to say, parameters are too numerous to be all represented so that I must arbitrarily choose to normalise some of them. I first aim at visualising the array of bargaining scores that satisfy condition (1) on a 2×2 market. Because condition (1) ensures LPUV-stability that is robust to any surplus system, it suffices to consider the four bargaining scores as parameters. To be able to plot points in three dimensions, I normalise one of the bargaining score, say $\alpha_1 = 1$, and restrict the three others, $\alpha_2, \beta_1, \beta_2 \in [0, 3]$. I browse the $[0, 3]^3$ space with a given step resulting from the trade off between precision and computation time. For any quadruplet complying with condition (1), I color the corresponding dot on a three-dimensional graph. The result in dynamic version is available in the attached file [Goursat-Dissertation-Graph-Condition_1.html](#)¹⁰. For security purposes, I still display in figure 1 a static version of the dynamic graph, captured from four different angles:

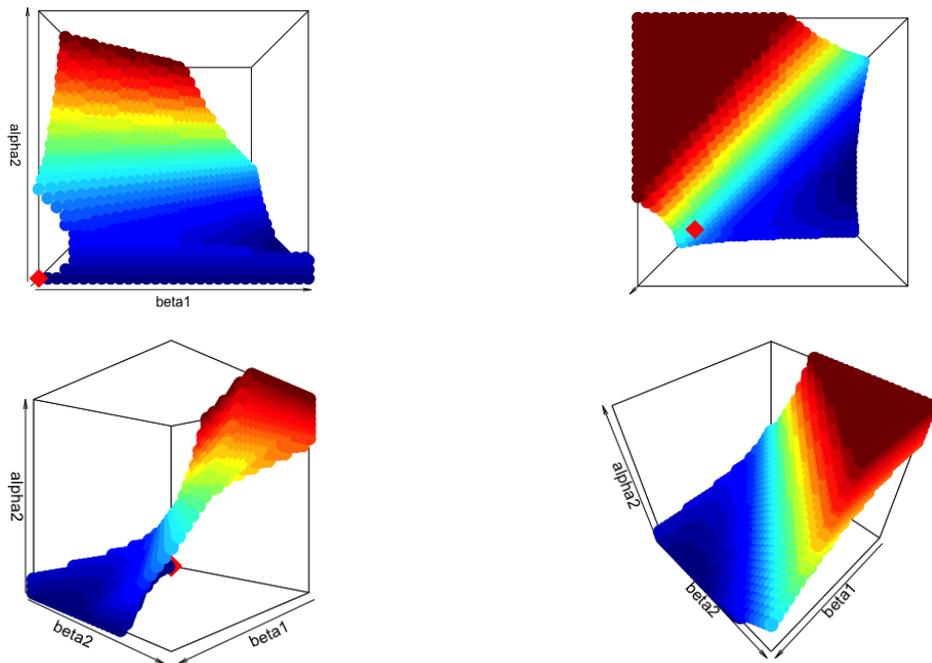


Figure 1: $\alpha_1 = 1, \alpha_2, \beta_1, \beta_2 \in [0, 3]^3$ - Colored dots for (1) - The red square labels the origin

I recognise none standard solid. At this stage, the interesting thing to do is to compare representations of (1) and (1*). Using the same normalisation and parameter domain, I plot the triplets

¹⁰The reader is encouraged to download the file. By clicking on the downloaded file, it should open on a web page. If it does not, one can copy and paste the full path and file name with the .html extension in the URL bar of a standard web browser such as Google Chrome.

$(\alpha_2, \beta_1, \beta_2)$ complying with condition (1^*) . The dynamic graph is available in the attached file: [Goursat-Dissertation-Graph-Condition_1etoile.html](#).¹⁰ The static version captured from two different angles delivers figure 2:



Figure 2: $\alpha_1 = 1, \alpha_2, \beta_1, \beta_2 \in [0, 3]^3$ - Colored dots for (1^*) - The red square labels the origin

Unsurprisingly, one observes two triangular prisms. Remind constraint (1^*) is a necessary, non sufficient condition for (1) . Estimating the share of points complying with both conditions among all the points complying with (1^*) provides a clue on how far (1^*) is from being sufficient. The estimated share amounts to 30% and is little sensitive to the choice of the normalisation and parameter domain; hence a disappointingly low percentage. It means assortativity captures roughly 30% of the constraining power of condition (1) . Looking more closely at the two graphs, one sees that going from conditions (1^*) to (1) , the majority of the point loss happens for $\beta_1 \simeq \alpha_1 = 1$ or very heterogeneous or homogeneous α_2, β_2 . This observation confirms the second necessary condition for (1) , that I could not write analytically, but proved through two extreme cases in subsection §3.2.2, and that I summarised as “intermediate” bargaining scores within sides of the market, neither too homogeneous, nor too heterogeneous.

Condition $(2a)$ obliges the bargaining scores to lie in the non-colored area below the colored solid on figure 1. It also constrains the realised surpluses to be distant enough from each other. The condition $\frac{s_{ii}}{s_{jj}} \in (0, \frac{\alpha_j + \beta_i}{\alpha_j + \beta_j}] \cup [\frac{\alpha_i + \beta_i}{\alpha_j + \beta_i}, +\infty)$ can be rewritten as: $\frac{\alpha_i + \beta_i}{\alpha_j + \beta_i} \cdot s_{jj} \leq s_{ii} \leq \frac{\alpha_j + \beta_i}{\alpha_j + \beta_j} \cdot s_{jj}$. The constraint is obviously linear with ordered slopes $\frac{\alpha_i + \beta_i}{\alpha_j + \beta_i} \leq \frac{\alpha_j + \beta_i}{\alpha_j + \beta_j}$. Consequently, the surpluses can locate anywhere in the $(\mathbb{R}_+^*)^2$ space but on an infinite triangle having as the finite vertex the origin. On figure 3, I color this infinite triangle in red, based on two different bargaining score systems complying with condition $(2a)$:

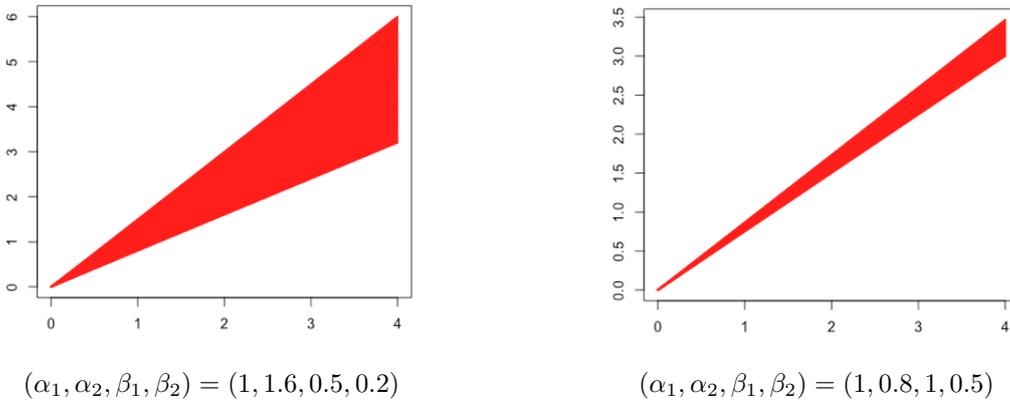


Figure 3: Uncolored dots for $(2a)$

3.3 Assortative matching stability and instability

In the whole subsection §3.2, I have produced an introductory comparison between GS-stability and LPUV-stability, on grounds of the analytical necessary and sufficient conditions for stability. Theorem 3.12. in subsection §3.2.2 exhibits sorting conditions on the individual bargaining power scores for LPUV-stability under regular splitting. What I would like to clarify now is the connection between assortativity and either notion of stability. For tractability reasons, I turn over the question and wonder: When is the assortative matching GS-stable or LPUV-stable? I expect conditions on surpluses and on the cardinality of bargaining scores. I begin with three special but thought-provoking cases (§3.3.1). I then endogenise the surpluses by introducing a surplus function taking as argument matched agents' bargaining scores (§3.3.2).

In all the subsequent work, bargaining scores are ordered: $\alpha_1 > \dots > \alpha_n$, $\beta_1 > \dots > \beta_n$ s.t. $\mu^a = [(m_k, f_k)_{k \in \{1, \dots, n\}}]$ is the assortative matching.

3.3.1 Three special cases suggesting a LPUV bias towards assortativity

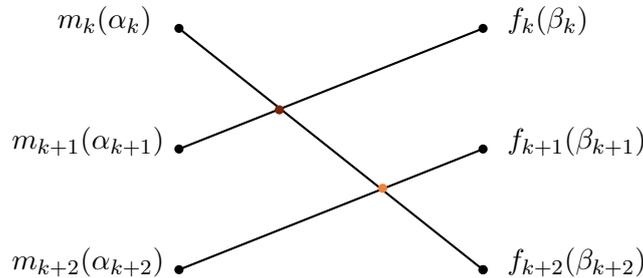
Equal surpluses, odd bargaining scores

Proposition 3.13. Suppose $\exists s > 0$ s.t. $\forall (i, j) \in \{1, \dots, n\}^2 : s_{ij} = s$.

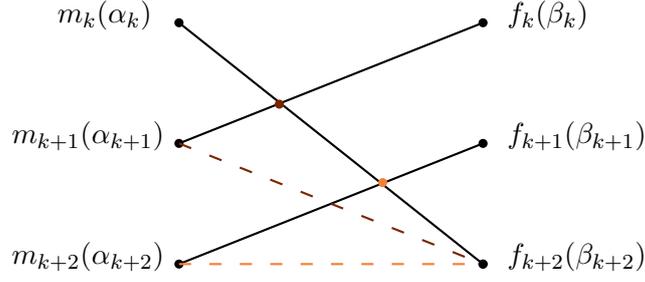
Then there exists a unique GS-stable full matching, and it is the assortative one: $\Sigma^{GS} = \{\mu^a\}$.

Proof

An algebraic brief proof relying on the necessary and sufficient conditions for GS-stability (theorem 3.12. in subsection §3.2.2) is provided in appendix §6.1. However, I would like to give the intuition of the result. GS-stability of the assortative matching is immediate. Due to the perfect equality in all realised and unrealised surpluses, the bargaining score of the contemplated partner is the only relevant aggregate for the matching decision. As a consequence, an agent will always prefer to be matched to a weak partner. Now, take any potential GS-blocking pair in the assortative matching (m_i, f_j) , $i \neq j$, and assume without loss of generality that $i < j$. This implies that: $\mu^a(j) = j > i \Rightarrow \alpha_{\mu^a(j)} < \alpha_i$, so f_j is not willing to block with m_i . Said differently, for any potential blocking pair, there is exactly one agent who is willing to deviate and one who is not. There is no GS-blocking pair and $\mu^a \in \Sigma^{GS}$. I provide a graphical intuition for the reverse statement. Set $\mu \neq \mu^a$ a full non assortative matching, and consider the bipartite graph representing μ . Because the market is balanced ($\#M = \#F = n$) and by basic properties of balanced bipartite graphs, there are at minimum two lines connecting matched agents that cross each other. Suppose for example that one observes the following scheme, for some $k \in \{1, n - 2\}$:



Let us look at one point where two edges intersect (for instance at the point colored in brown on the figure). Let us spot the four dots connected to this point (here, $m_k, m_{k+1}, f_k, f_{k+2}$) and consider the two dots at the bottom (here, m_{k+1} and f_{k+2}). It turns out that the corresponding agents form a blocking pair. Indeed, the location of the two bottom dots below the intersection point ensures that by matching together, both agents deviate to a weaker agent than their current matches. On the next figure, I represent in dashed colored lines the edges connecting the GS-blocking pairs:



I seek the same kind of insight for LPUV-stability:

Proposition 3.14. *Let μ be a matching. Suppose $\exists s_\mu > 0$ s.t. $\forall i \in \{1, \dots, n\} : s_{m_i \mu(m_i)} = s_\mu$. Then: μ is LPUV-stable $\iff \mu = \mu^a$.*

Proof

The proof is based on the necessary and sufficient conditions for LPUV-stability I explicated in theorem 3.12. in subsection §3.2.2. The intuition is that when realised surpluses are equal, beliefs on unrealised surpluses are alike, thus the graphical reasoning exposed above for GS-stability applies.

Corollary 3.15. *Suppose $\exists s > 0$ s.t. $\forall (i, j) \in \{1, \dots, n\}^2 : s_{ij} = s$. Then there exists a unique LPUV-stable full matching, and it is the assortative one: $\Sigma^{LPUV} = \{\mu^a\}$.*

When all surpluses are equal, for any matching μ , $\exists s_\mu = s > 0$ s.t. $\forall i \in \{1, \dots, n\} : s_{m_i \mu(m_i)} = s_\mu$. We are definitely in the environment of proposition 3.14.. Propositions 3.13. and corollary 3.15. combined deliver the proposition below:

Proposition 3.16. *Suppose $\exists s > 0$ s.t. $\forall (i, j) \in \{1, \dots, n\}^2 : s_{ij} = s$. Then there exist a unique GS-stable full matching and a unique LPUV-stable full matching and in both cases it is the assortative one: $\Sigma_F^{GS} = \Sigma_F^{LPUV} = \{\mu^a\}$.*

This is the general statement underlying example c. in subsection §3.1.2.

Equal bargaining scores across the market, odd surpluses

Proposition 3.17. *Suppose $\forall (i, j) \in \{1, \dots, n\}^2 : \alpha_i = \beta_j$. Then, the assortative matching is LPUV-stable: $\mu^a \in \Sigma^{LPUV}$.*

Proof

The algebraic proof is simplistic as based on the necessary and sufficient conditions for LPUV-stability reported in theorem 3.12.. Condition (1) is verified; equivalently any bargaining score system constant across sides $\forall i \in \{1, \dots, n\}$ ensures robust to surplus system LPUV-stability. The intuition I find a little more intricate. Under equal bargaining scores, in μ^a , any agent receives half of the surplus. Set (m_i, f_j) , $i \neq j$ a potential blocking pair. There are two possibilities. The trivial case is when $\alpha_i = \alpha_j$, so that $\frac{\alpha_i}{\alpha_i + \alpha_j} = \frac{\alpha_j}{\alpha_i + \alpha_j} = \frac{1}{2}$, no agent increases his surplus and by corollary 3.9. in subsection §3.1.4, (m_i, f_j) is not a blocking pair. In the non-trivial case, $\alpha_i \neq \alpha_j$. Let me assume, without loss of generality than $i < j$, which induces $\alpha_i > \alpha_j$. The new splitting rule is $(\lambda, (1 - \lambda)) \equiv (\frac{\alpha_i}{\alpha_i + \alpha_j}, \frac{\alpha_j}{\alpha_i + \alpha_j})$. The total welfare before blocking was: $W^0 = \frac{1}{2} \cdot s_{ii} + \frac{1}{2} \cdot s_{jj}$. The total LPUV-expected welfare after blocking is $W^1 = \lambda \cdot s_{jj} + (1 - \lambda) \cdot s_{ii}$. By elementary properties of convex combinations, with $\lambda > \frac{1}{2} > (1 - \lambda)$: $W^0 > W^1$. This induces that at least one agent anticipates a decrease in utility by blocking, equivalently (m_i, f_j) is not a LPUV-blocking pair. $\mu^a \in \Sigma^{LPUV}$.

GS-stability of the assortative matching under equal bargaining scores across the market depends on non-assortative surplus specification.

Aligned ordinality between bargaining scores and surpluses

Proposition 3.18.

Set bargaining scores for side M , F respectively to be $\alpha_1 \geq \dots \geq \alpha_n$, $\beta_1 \geq \dots \geq \beta_n$. Suppose part of the surpluses are ranked in the following order: $s_{11} \leq \dots \leq s_{nn}$. Then, the assortative matching is LPUV-stable: $\mu^a \in \Sigma^{LPUV}$.

Proof

The proof is done by strong induction on $k \in \{1, \dots, n\}$. The weaker agents α_n, β_n are not willing to deviate as they are each matched to the weakest agent belonging to the opposite side of the market, and by LPUV-rationality they (potentially wrongly) believe that they would go to a lower surplus by deviating. α_n, β_n never engage in a LPUV-blocking pair. Conditional on this, $\alpha_{n-1}, \beta_{n-1}$ do not wish to deviate either. Because the only agent with which α_{n-1} (for instance) could have thought of blocking (a large ratio $\frac{s_{nn}}{s_{ii}}$ would still be required, according to theorem 3.12. in subsection §3.2.2) is β_n .

This setting contains none ordinal condition involving the non-assortative surpluses. That is why again nothing can be said on the GS-stability of the assortative matching.

3.3.2 Assortative matching and endogenous surplus

In all the foregoing, surpluses were taken to be exogenous, and as such, independent from the splitting rules (from the bargaining power scores in the regular splitting case). Though, endogenising the surplus is a decisive step to take for several reasons I briefly explain. From an analytical perspective, the characterisation of GS-stable and LPUV-stable matchings has been constantly plagued by the multiplicity of parameters to consider. In addition, the settings that delivered valuable results were partly counter-intuitive. Proposition 3.18. in subsection §3.3.1 well illustrates this point. It considers a situation where the weakest agents would jointly generate the largest surplus. Finally, endogenous surplus enables to model the effect of an outside variable bearing dual causality on the splitting and surpluses. Let me be precise here. Suppose this outside variable, denoted e impacts the individual bargaining scores (respectively surpluses) through a bijective invertible function invariant across different individuals (pairs):

$$\forall (i, j) \in \{1, \dots, n\}^2 : \begin{cases} \alpha_i = \alpha(e_i) \\ \beta_j = \beta(e_j) \\ s_{ij} = f(e_i, e_j) \end{cases}$$

By inversion and composition, the surplus can be written as a function of the outside variable:

$$\forall (i, j) \in \{1, \dots, n\}^2 : s_{ij} = f(\alpha^{-1}(\alpha_i), \beta^{-1}(\beta_j)) \equiv s(\alpha_i, \beta_j)$$

Empirically, a good candidate for this outside variable could be, for the labor market, education. Common sense suggests that a skilled worker will, one, generate significant added value in the production process. Second, he will also be more aware of the value of his output, of the current macroeconomic conditions affecting the labor market equilibrium and so on, so that he is likely to wield a preferential bargaining position with respect to the employer.

Additive endogenous surplus

I consider a special case where the surplus is additive in the bargaining power scores. In particular, there is no complementary between agents in the surplus production process. More precisely, an agent i recovers exactly the surplus he generates while being matched:

Proposition 3.19. Suppose the surplus is additive with respect to bargaining power scores:

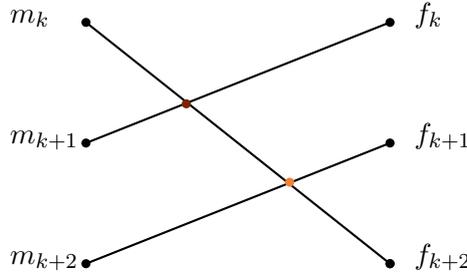
$\forall (i, j) \in \{1, \dots, n\}^2 : s_{ij} = s(\alpha_i, \beta_j) = \alpha_i + \beta_j$. Then:

(i) Any full matching is GS-stable: $\Sigma_F \subseteq \Sigma^{GS}$.

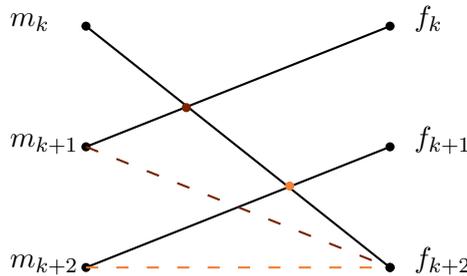
(ii) There is a unique full LPUV-stable matching and it is the assortative one: $\Sigma_F^{LPUV} = \{\mu^a\}$.

Proof

The intuition for (i) is very straightforward. The true match utility writes: $u_{ij} = \frac{\alpha_i}{\alpha_i + \beta_j} \cdot (\alpha_i + \beta_j) = \alpha_i$, which depends only on i . From the agent m_i perspective, the matching utility is fixed, hence inconsequential matching decision. Regarding (ii), A LPUV-rational agent m_i is willing to deviate with a partner f_j if and only if f_j has a stronger partner $\mu(f_j)$ than herself: $\alpha_{\mu(f_j)} \leq \alpha_i$. This because the strength of the partner of i 's potential match enters the valuation $v_i(j)$ in the surplus part (thus in an increasing fashion) and does not enter the splitting part. In sum, the bargaining score of the current partner of the contemplated partner is the only aggregate relevant for the matching decision. This gives a criteria for deviation which is assortative, and should ring a bell. The kind of proof performed for proposition 3.13. in subsection §3.3.1 applies here. Indeed, take any potential blocking pair in the assortative matching (m_i, f_j) , $i \neq j$ and assume without loss of generality that $i < j$. This implies: $\mu^a(j) = j \Rightarrow \alpha_{\mu^a(j)} < \alpha_i$, so m_i is not willing to block with f_j . Said differently, for any potential blocking pair, there is exactly one agent who is willing to deviate and one who is not. There is no LPUV-blocking pair and $\mu^a \in \Sigma^{LPUV}$. I provide a graphical intuition for the reverse statement. I set $\mu \neq \mu^a$ and pose the same scheme as in subsection §3.3.1, $k \in \{1, \dots, n - 2\}$:



Let us look at one point where two edges intersect (for instance at the point colored in blue on the figure). Let us consider the four dots connected to this point (here, $m_k, m_{k+1}, f_k, f_{k+2}$) and consider the two dots at the bottom (here, m_{k+1} and f_{k+2}). It turns out that the corresponding agents form a LPUV-blocking pair. Indeed, the location of the two bottom dots below the intersection point ensures that both agents deviate to an agent whose current partner is stronger than them. On the next figure, I represent in dashed colored lines the edges connecting the LPUV-blocking pairs:



Let me remark I followed the same proof outline in both subsection §3.3.1 for proposition 3.13. (GS-stability under equal surpluses across the market) and here for proposition 3.19. (LPUV-stability under endogenous additive surplus). This reasoning is valid as soon as the ordinal preferences resulting from the estimation of matched utilities are assortative.

General endogenous surplus

I finally seek to characterise the stability of the assortative matching in the general endogenous

surplus case, that is, I would like to find conditions on the surplus function s that ensure that μ^a is GS-stable or LPUV-stable. I have to say I developed the additive case because of the nice simplification allowing full characterisation of stable matchings. But in Becker's idea for example, marriage production function displays complementarities in spouses input factors. It may be, for instance a Cobb-Douglas function: $s(\alpha_i, \beta_j) = c \cdot \alpha_i^a \cdot \beta_j^b$, $a, b, c > 0$.

Theorem 3.20. *Set a symmetric differentiable surplus function $\forall (i, j) \in \{1, n\} : s_{ij} = s(\alpha_i, \alpha_j)$.*

$$\text{If } \forall (x, y) \in (\mathbb{R}_+^*)^2, \frac{\partial s(x, y)}{\partial x}(x + y) - s(x, y) < 0$$

$$\text{or } \forall (x, y) \in (\mathbb{R}_+^*)^2, \frac{\partial s(x, y)}{\partial x}(x + y) - s(x, y) > 0$$

Then: $\mu^a \in \Sigma^{GS}$.

Proof

$\frac{\partial s(x, y)}{\partial x}(x + y) - s(x, y) > 0$ ensures that considering any potential GS-blocking pair (m_i, f_j) , $i < j$ (so m_i is higher ranked on the criterion of bargaining power score within his own side of the market than f_j is), m_i does not want to deviate with f_j . Conversely, $\frac{\partial s(x, y)}{\partial x}(x + y) - s(x, y) < 0$ ensures that f_j (lower ranked than m_i on the criterion of bargaining power score within own side of the market) does not want to deviate with m_i .

The analysis of the LPUV case is a little more sophisticated as, even when the favorable case when the surplus function is symmetric, the valuation function is not symmetric in bargaining power scores. The ambition though, should be to derive analytical conditions on the surplus function that would be directly comparable with theorem 3.20. .

4 Discussion

4.1 Value of the concept

As far as I know, the theoretical assessment of any bounded rationality concept is common whatever the environment the modeller has chosen to consider. A quality heuristic is expected to much simplify individual reflection (in its numerous dimensions: computation, inference, memory, reasoning), while leading to low-consequential mistakes. Stated in utilitarian terms, a convincing heuristic produces a significant decrease in reflection costs, while an insignificant shrinkage in payoff. Pushed to extremes, such heuristic should appear as optimal within a cost-benefit analysis featuring a utility function subtracting reflection costs to real payoffs, and a utility-maximising agent. This kind of argument I feel very attached to is much in the vein of rational inattention models (originally proposed by Sims (2003) [30] for macroeconomic models).

I examine (only qualitatively) the strength of LPUV-rationality under this filter. Namely, I advocate for each of the four components of LPUV-rationality as solving the trade off between cost of cognition and cost of doing strategic mistakes, and discuss the matching environments where LPUV-rationality is bound to be the most helpful. First, the "Live" attitude is advantageous under high memory costs, on markets akin to my motivating examples of marriage or job markets, where one is "on the market" only a few times in his life, and for short periods, and when individual characteristics hence match utilities evolve at fast pace over time. Second and third, the "Valuation" and "Unidimensional" components partly solve the under identification issue I have emphasised in subsection §1.1. More precisely for the information structure I have specified, remember the difficulties that a Bayesian agent faces. With only one observation (of the current surplus $s_{\mu(f)f}$), he must build a posterior distribution on k variables (on f 's vector of characteristics C_f) to finally recover a posterior distribution on the surplus when matching with f , s_{mf} . In econometric terms, the Bayesian-estimate is unbiased but very noisy. In

contrast, a LPUV-rational agent uses one observation (of the current surplus $s_{\mu(f)f}$), to build a point estimate of one variable (the surplus when matching with f , s_{mf}). In econometric terms still, the LPUV-estimate is biased but is more precise. Intuition suggests there is a trade-off between bias and noise, between having *too much* data or *too little* data per observation, so that LPUV-rationality may be preferable than a pure Bayesian approach in this setting.¹¹ Fourth and final, the “Polarisation” is especially beneficial when there are numerous agents on one’s own side of the market, few on the opposite side. Here I should acknowledge that any lower level partition (for instance, bundling of several characteristics of potential partners) of higher level partitioning (for instance, bundling of several potential partners current surpluses) would also be a candidate. Yet, valuation function imposes itself as a convenient thus natural level of partition.

The empirical value of the concept lies mainly in its relation to assortativity. Indeed, the robust case in subsection §3.2 as well as examples in subsection §3.3 hint that LPUV-stable matchings are disproportionately assortative and conversely. LPUV-stability shows potential to account for the long-established empirical puzzle on assortativity in matching markets. Furthermore, very preliminary evidence provided in appendix §6.4.2 also suggests that full matchings are less often LPUV-stable than GS-stable, at least for some well-behaved distribution of match utilities. Last but not least, each of the four components relates to psychological biases that are well documented empirically (for illustration: prior neglect for “Live”, anthropocentric thinking for “Polarisation”).

In any case, direct empirical testing of the theory needs to be performed. Field data on centralised matching markets does exist.¹² This being said, disentangling bounded rationality from misunderstanding of the mechanism is virtually impossible as Artemov, Che and He (2017) [4] and Le Chapelain (2015) [9] found out. Laboratory experiments quite the opposite enable to test the understanding of participants on the information and allocation mechanisms before going through rationality tests. Besides, I think laboratory testing of the bounded rationality notions I introduced above is direct¹³.

4.2 Specification issues

So far, I have detailed the mistake that a LPUV-rational agent makes when estimating the utility he would get from matching with a currently matched agent. This coincides with the second line in definition 2.3.. A number of assumptions are also made implicit in the first and third lines of the definition.

According to the first line $v_m(\emptyset) = u_{m\emptyset}^m$, that is a LPUV-rational agent perfectly estimates the reservation utility. This is not demanding, since reservation utility is known under ψ . Yet it is not very consistent either, because the assessment of the reservation utility is not “Live” as for the assessment of the match utilities. If I posed an information structure ψ' similar to ψ except

¹¹From what I have understood, this trade-off between *too much* data or *too little* data, as obvious in the game of Chess, has been a core motivation for Analogy Based Expectation Equilibrium by Jehiel (2005) [16] and Valuation Equilibrium by Jehiel and Samet (2007) [17].

¹²In school choice, most famous empirical mechanism design work include Abdulkadiroğlu and al (2005) [3] and Abdulkadiroğlu and al (2009) [1], based on data sets on choices and characteristics of students and schools in the Boston, respectively New York City Public School systems).

¹³Here is an outline for: (1) an experimental protocol for testing LPUV-rationality, (2) exploitation of the collected data. Participants are split into two sides. After an understanding test stage, an initial allocation is performed. Surpluses of realised matches are publicly displayed together with series of characteristics of matched agents (eg.: level of education). Agents are also privately informed on the shares of the surplus they would receive if matched with any agent of the opposite side. Then agents are asked to report their estimates of match utilities, where the payment scheme should incentivise them to produce good assessment (payoff could be a decreasing function of the distance between estimated and true match utilities). I would then regress the estimated utilities separately on the true utilities, on the live valuations, and on the characteristics (several specifications can be tested: characteristics of the opposite side, distance between own and other characteristics) to spot correlation.

that reservation utilities are unknown, a Valuation that would makes sense is:

$$v_m(f) = \begin{cases} \frac{1}{\#\{k \in M | \mu(k) = \emptyset\}} \sum_{k \in M | \mu(k) = \emptyset} u_{k\emptyset} & \text{if } f = \emptyset \\ r_{mf} \cdot s_{\mu(f)f} & \text{if } f \neq \emptyset, \mu(f) \neq \emptyset \\ \in \mathbb{R}_+ & \text{if } f \neq \emptyset, \mu(f) = \emptyset \end{cases}$$

This modified definition states that to estimate his reservation utility, an agent computes an arithmetic mean of all reservation utilities of currently unmatched agents. This discussion transposed to the marriage market delivers the following. Definition 2.3. assumes that after twenty years of marriage, a woman still knows how it feels to be single. Quite the opposite, the definition just above implies the married woman has long forgotten the value of single life. To assess single life value, she relies on observation of all single women satisfaction. Let me be caricatural again: Do single women seem to entertain a lot going out in nightclubs? This is what a married women will base her estimate of the value of single status on.

The third line in the definition allows for any positive estimate of the utility of being unmatched. In marital terms, it allows for any negative or positive prejudice against single individuals. Fantasy enters the cognition framework defined here but only where no signal is available. A more classic modelling choice consistent with both the principle of insufficient reason and the “Live” denomination of the function, would model:

$$v_m(f) = \begin{cases} u_{m\emptyset}^m & \text{if } f = \emptyset \\ r_{mf} \cdot s_{\mu(f)f} & \text{if } f \neq \emptyset, \mu(f) \neq \emptyset \\ r_{mf} \cdot \frac{1}{\#\{k \in F | \mu(k) \neq \emptyset\}} \sum_{k \in F | \mu(k) \neq \emptyset} s_{\mu(k)k} & \text{if } f \neq \emptyset, \mu(f) = \emptyset \end{cases}$$

This modified definition states that to estimate the utility of being matched to a currently unmatched agent, an agent computes an arithmetic mean of all currently produced surpluses and plugs it into the correct splitting rule.

Yet, as obvious in the proof of theorem 3.4., assuming perfect assessment of reservation utility allows for simplified comparison with usual stability, and not imposing structure on the estimate of match utilities with unmatched agents guarantees existence of a LPUV stable matching. I thus keep on with definition 2.3..

Smooth generalisation to many-to-one markets is another asset of the concept. Let us locate on a two-sided many-to-one matching market where agents on side F can be matched to more than one agent on side M . On this market, a matching is any correspondence μ :

$$\begin{aligned} \mu &: M \cup F \rightrightarrows M \cup F \cup \{\emptyset\} \\ \forall m \in M, \mu(m) &\subseteq F \cup \{\emptyset\} \text{ and } \#\mu(m) \leq 1, \forall f \in F, \mu(f) \subseteq M \cup \emptyset \\ \forall m \in M, f \in F &: \mu(m) = \{f\} \iff m \in \mu(f) \end{aligned}$$

The definitions of a valuation and of a LPUV-rational on side F need not be changed. The cogent definition of a LPUV for an agent on the “many” side F is the following:

Definition 4.1. Let μ be a matching on a many-to-one market where M is the “one” side, F is the “many” side, and ψ is the information structure.

A valuation is μ -Live-Polarised-Unidimensional for agent $m \in M$ if it writes:

$$v_m(f) = \begin{cases} u_{m\emptyset}^m & \text{if } f = \emptyset \\ u_{mf}^m & \text{if } f \neq \emptyset, \mu(f) \neq \emptyset, m \in \mu(f) \\ r_{mf} \cdot \frac{1}{\#\mu(f)} \sum_{k \in \mu(f)} s_{kf} & \text{if } f \neq \emptyset, \mu(f) \neq \emptyset, m \notin \mu(f) \\ \in \mathbb{R}_+ & \text{if } f \neq \emptyset, \mu(f) = \emptyset \end{cases}$$

The interpretation is straightforward, I formulate it with a job market terminology. Suppose two well established firms, say firms 1 and 2, employ 10 workers each, and a start-up, say firm 3, was just born and has no employee yet. An employee working in firm 1 perfectly knows his reservation utility (first line), meaning the utility of being unemployed. He also perfectly knows his utility when working at firm 1 (second line). To assess the utility of working at firm 2, the employee computes an arithmetic mean of the 10 observed surpluses of the contract between firm 2 and each of her 10 employees and plugs this mean into the right splitting rule (third line). He has a positive unstructured estimate of the utility of working in the start-up (fourth line), for instance I allow him to fantasise on start-up lifestyle.

4.3 Mirror cases

In the model described above, and as discussed in considerable details in subsection §2.2.2, the mistake in social learning consists in neglecting the role of one's own side of the market in determining the surplus. A mirror agent would assume that his side of the market fully determines the match surpluses. In this case, LPU valuation¹⁴ would write:

$$w_m(f) = \begin{cases} u_{mf}^m & \text{if } f = \emptyset \\ r_{mf} \cdot s_{m\mu(m)} & \text{if } f \neq \emptyset, \mu(f) \neq \emptyset \text{ and } \mu(m) \neq \emptyset \\ \in \mathbb{R}_+ & \text{if } f \neq \emptyset, \mu(f) = \emptyset \text{ or } \mu(m) = \emptyset \end{cases}$$

For the moment, I remain unsure about which way is the most relevant empirically. On the one hand, the latest definition depicts an agent who overestimates his individual might, which is consistent with evidence on egocentric bias in psychological literature. On the other hand, if this is indeed the way agents think, then heterogeneity in surplus estimates only comes from heterogeneous splitting rules. This should facilitate inference that the surplus estimation method is void making the LPU-valuation non sustainable over time. Anyway, this case is very direct to solve, since any matched agent on the market have the same ordinal preferences on potential partner, ranking the weakest agents on top. This mechanically generates that the unique full LPUV-stable matching is the assortative one.

My conviction is that empirically agents rather use a combination of both the classic LPU-valuation v and the egocentric LPU-valuation w . In the simplest form, the estimate of a match utility could be a non weighted arithmetic mean of the two criteria:

$$\begin{aligned} \hat{u}_{mf}^m &= \frac{1}{2}(v_m(f) + w_m(f)) \\ &= \frac{1}{2}r_{mf}(s_{\mu(f)f} + s_{m\mu(m)}) \text{ if } f \neq \emptyset, \mu(f) \neq \emptyset \end{aligned}$$

Another issue deals with the assumed information structure. Under information structure ψ , surpluses of currently matched agents are common knowledge, and splitting rules are private knowledge at the m level. That is, if the current match is μ , an agent $m \in M$ observes $s_{m'\mu(m')}$, $\forall m' \in M$ and knows $s_{m\emptyset}$. He also knows r_{mf} , $\forall f \in F$, but ignores $r_{m'f}$, $\forall m' \neq m$, $f \in F$. As previously argued, this setting is especially relevant for the marriage market. It is supported by the fact that part of the surplus of a marriage is directly observable (children), whereas the splitting rule depends of some intimate negotiation between spouses. It seems less adequate for the job market for instance. The surplus of a contract of employment is in general unknown to the public, whereas the splitting is more easily uncovered¹⁵. In

¹⁴Despite risk for confusion, my labelling of mirror cases is invariant: "LPU-valuation", and "LPUV-rationality". The reason is that each of the four components of LPUV-rationality remains valid. What changes is the side of the polarisation (mirror case 1) or object of the valuation (mirror case 2)

¹⁵Think about this acquaintance bragging in every possible occasion what a good deal he has with Goldman Sachs, or about this other one who complains to be squeezed by an employer who pays him a low wage while filling his pockets

this latter example, it would make more sense to assume an alternative information structure ψ' where splitting rules of currently matched agents are common knowledge, and surpluses are private knowledge at the m level. That is, if the current match is μ , an agent $m \in M$ observes $r_{m'\mu(m')}$, $\forall m' \in M$, and knows $r_{m\emptyset}$. He also knows s_{mf} , $\forall f \in F$, but ignores $s_{m'f}$, $\forall m' \neq m, f \in F$. Under ψ' , a LPUV-rational¹⁴ agent would assess his match utilities in the following way:

$$v_m(f) = \begin{cases} u_{m,f}^m & \text{if } f = \emptyset \\ r_{\mu(f)f} \cdot s_{mf} & \text{if } f \neq \emptyset, \mu(f) \neq \emptyset \\ \in \mathbb{R}_+ & \text{if } f \neq \emptyset, \mu(f) = \emptyset \end{cases}$$

I do not intend to chose between the original and latest modelling as I illustrate each of them can fit different markets.

Key results for both “mirror” cases are included in appendix §6.2.

4.4 Relation to Valuation Equilibrium

As the reader may have noticed already, the concept of LPUV-rationality shares similar traits with Valuation Equilibrium by Jehiel (2005) [16], and I have to acknowledge inspiration. The analogous to Valuation Equilibrium on a (one-to-one) matching market would be the following:

Definition 4.2. μ is **optimal** for side M under the family of valuation $(v_m)_{m \in M}$ if:

$$\forall m \in M, \mu(m) \in \arg \max_{f \in F \cup \emptyset} v_m(f)$$

Definition 4.3.

μ is a **LPUV-best** matching for side M if there exists a family of valuations $(v_m)_{m \in M}$ s.t.:

- (i) μ is optimal for side M under $(v_m)_{m \in M}$.
- (ii) $\forall m \in M, v_m$ is μ -LPU.

A characterisation of LPUV-best matchings is provided in appendix §6.3, as well as the link with LPUV-stability. What stands out of this analysis is that the concept is too strong to be viable on real markets. In appendix §6.3 still, I clarify the relation between the latter concept and Valuation Equilibrium in terms of valuation partition and using a tree representation of matching markets.

5 Conclusion

Motivations for this dissertation has been two-fold. On the one hand, research on matching markets and bounded rationality has remained very disconnected (§1.1). This is bewildering if one considers the recent boom in both the matching and the bounded rationality literatures, plus the fact that matching markets are the one environment where the full rationality theory is the most unlikely to hold. On the other hand, the rational theory has failed to explain a number of phenomena observed on real matching markets (§1.2). In the introduction, I have emphasised three of them: the persistent instabilities, the assortativity in self-declared matching preferences, and the under-investment in information acquisition and strategy design. That is why I have investigated the following research question (§1.3): What is the nature of bounded rationality on matching markets? Does it reshuffle the matching outcome and if so, in what direction?

This paper has been an exploratory work to address this question, on a two-sided agent-agent one-to-one market featuring Beckerian match utilities, and a fixed information structure (§2.1). That is, matched agents jointly generate a surplus and share the surplus according to a (here exogenous) splitting rule. With respect to information, splitting rules are private knowledge, the realised surpluses are common knowledge whereas the unrealised surpluses are unknown. I have extensively argued that the structure of matching markets calls for a brand new concept of bounded rationality, and consistently proposed a heuristic that agents on matching markets may follow, as a reasonable trade-off between matching markets complexity and human cognitive limitations (§2.2). To make matching decisions, an agent m needs to form beliefs about the match utilities he would get when matched with any potential partner f . A Live-Polarised-Unidimensional-Valuation (LPUV) rational m does so by plugging the surplus $s_{\mu(f)f}$ currently generated by f and her current partner $\mu(f)$, which is an incorrect estimate of the surplus s_{mf} , into the correct splitting rule r_{mf} . I also assume the agent knows his reservation utility $u_{m\emptyset}^m$, perfectly observes the utility of his current match $u_{m\mu(m)}^m$ and can store any estimate of the utility he would get from mating with a currently unmatched agent u_{mf} , $\mu(f) = \emptyset$. Under LPUV-rationality, social learning on the match surpluses is definitely non-Bayesian: the mistake made amounts to neglecting the role of one's own side of the market in the determination of the surplus. LPUV-rationality has enabled me to define a new LPUV pairwise stability notion, consisting in the usual incentive rationality and incentive compatibility constraints under the coarse estimation of match utilities.

I have studied the implications of LPUV-rationality on stability of matching markets, in other words, I have performed a comparative analysis between the usual stability notion (as introduced by Gale and Shapley (1962) [15]) and the new one (LPUV-stability). Based on examples, I have shown that a LPUV-stable matching always exists and that there exists no simple inclusion or disjunction relation between the sets of GS-stable and LPUV-stable matchings (§3.1). Yet, LPUV-stability leans towards matchings where a low number of agents are matched, due to the unconstrained specification for the valuation of currently unmatched potential partners. I have characterised GS-stable and LPUV-stable full matchings under odd and regular splitting rules. What has proven rather striking is that GS-stability invariably depends on both the splitting and surpluses systems, whereas there exist some splitting systems for which LPUV-stability is robust to any surplus system (§3.2). In this latter robust case, and under regular splitting rule, assortativity in bargaining power scores is a necessary (not sufficient) condition. Consequently and lastly, I have sought to clarify the relation between assortativity in bargaining scores and GS-stability or LPUV-stability, in the regular splitting case (§3.3). I have done so by providing three examples: in one of them, the assortative matching is the unique GS-stable and LPUV-stable matching, in the other two, the assortative matching is LPUV-stable but not necessarily GS-stable. I have finally made a first tentative to endogenise the surplus. In the painless additive case for instance, assortativity and LPUV-stability are again equivalent whereas GS-stability

is undemanding.

In my opinion, the main contributions of this paper are, (i) the novelty and relevance of the issue, (ii) the discussion over a tailor-made LPUV-rationality for matching markets, (iii) the fact that the proposed LPUV-stability concept can account for the long-established empirical puzzle on assortativity on markets without transfers. Nonetheless, I remain conscious that the results presented in the paper need to be consolidated and that the examples set need to be generalised. This could be done using new designed tools for matching markets under limited information and bounded cognitive rationality I present in appendix §6.4.

Future research should go in five directions. (1) The specification of LPUV-rationality has to be questioned through more numerous empirical examples of matching markets (mine were mainly related to the marriage and job markets), and complementary analysis on the so-called “mirror cases”. In particular, and as stated in subsection §4.3, I would like to explore the case when agents use a combination of the classic LPU-valuation and egocentric valuation for inference on match utilities. The convex, unweighted case for instance would write $\hat{u}_{mf}^m = \frac{1}{2}r_{mf}(s_{\mu(f)f} + s_{m\mu(m)})$ if $f \neq \emptyset$, $\mu(f) \neq \emptyset$. Whatever the bounded rationality definition I pose, I should not compromise on the personal intuition that bounded rationality on matching markets proceeds from data size¹⁶ and competition complexity challenging human cognitive limitations. (2) I also would like to pursue endogenising the surplus, in functional form $s_{mf} = s(\alpha_m, \beta_f)$. As a priority, I would like to test Cobb-Douglas surplus functions: $s(\alpha_m, \alpha_f) = c \cdot \alpha_m^a \cdot \beta_f^b$, $a, b, c > 0$. Making bargaining scores continuous, I expect to obtain conditions on the shape of the surplus function s for GS-stability and LPUV-stability¹⁷, which hopefully could be compared with each other. This methodology has two virtues. One, it reduces the number of parameters, thus simplifying the writing of results (modelling concern). Second, it allows to model dependence between spitting and surplus, possibly through a third outside variable, which is undoubtedly more realistic (empirical concern). (3) I should perform a welfare analysis. Remember that GS-stability favors matchings where the realised surpluses are high with respect to unrealised surpluses, whereas LPUV-stability remains insensitive to this comparison. Yet, LPUV-stability leans towards the null and assortative matchings, which are welfare superior when respectively reservation utilities are high and when there is positive complementarity in between agents scores in the surplus production function (endogenous case). Which effect dominates the other remains to be seen to conclude on welfare (as measured by Pareto-dominance, maximisation of the sum of utilities, and so forth). (4) I also aim at shifting from a concept of LPUV-stable matchings to the level of LPUV-stable mechanisms. As for now, I remain concerned that it would compel to manipulate reallocation (rather than allocation) mechanisms, thus preventing the comparison between GS-stable (allocation) mechanisms and LPUV-stable (reallocation) mechanisms. (5) On the longer run, I would also find interesting to shift to the population level. What happens when one side of the market is perfectly rational and the other one is LPUV-rational? When part of one side of the market is perfectly rational and the other part of the same side is LPUV-rational? I would like to determine who benefits and who suffers from rationality asymmetries on matching markets. (6) Last in my list but ideally first in time, empirical testing of the theory should be performed. Indeed, to assess the validity of the approaches described above, going to the data is imperative. As discussed above, I envision generating my own data through laboratory experiments rather than exploiting fields data: because I would like to test the understanding of participants on the information and allocation mechanisms before going through rationality tests, and because at first glance, the concept seems to well abide by the experimental protocol format.

¹⁶Implicit in this rule is the choice to maintain full instrumental rationality and to relax cognitive rationality, plus the assumption that coarse cognition results of limited data processing abilities rather than of individual reasoning flaws. Let me thus once more acknowledge inspiration from class by, reading of, and discussion with Pr. Jehiel.

¹⁷In the model of theorem 3.20. for GS-stability.

References

- [1] Atila Abdulkadiroğlu, Parag A Pathak, and Alvin E Roth. “Strategy-proofness versus efficiency in matching with indifference: Redesigning the NYC high school match”. In: *American Economic Review* 99.5 (2009), pp. 1954–78.
- [2] Atila Abdulkadiroğlu and Tayfun Sönmez. “School choice: A mechanism design approach”. In: *American economic review* 93.3 (2003), pp. 729–747.
- [3] Atila Abdulkadiroğlu et al. “The Boston public school match”. In: *American Economic Review* 95.2 (2005), pp. 368–371.
- [4] Georgy Artemov, Yeon-Koo Che, and Yinghua He. *Strategic ‘Mistakes’: Implications for Market Design Research*. Tech. rep. mimeo, 2017.
- [5] Gary S Becker. “A theory of marriage: Part I”. In: *Journal of Political economy* 81.4 (1973), pp. 813–846.
- [6] Gary S Becker. “A theory of marriage: Part II”. In: *Journal of political Economy* 82.2, Part 2 (1974), S11–S26.
- [7] Gary S Becker, Elisabeth M Landes, and Robert T Michael. “An economic analysis of marital instability”. In: *Journal of political Economy* 85.6 (1977), pp. 1141–1187.
- [8] Donn Byrne and Don Nelson. “Attraction as a linear function of proportion of positive reinforcements.” In: *Journal of personality and social psychology* 1.6 (1965), p. 659.
- [9] Alexis Le Chapelain. “Estimating students’ preferences and bounded rationality in Mexico City”. 2015.
- [10] Pierre-André Chiappori and Bernard Salanié. “The econometrics of matching models”. In: *Journal of Economic Literature* 54.3 (2016), pp. 832–61.
- [11] Ronald H Coase. “The problem of social cost”. In: *Classic papers in natural resource economics*. Springer, 1960, pp. 87–137.
- [12] John Conlisk. “Why bounded rationality?” In: *Journal of economic literature* 34.2 (1996), pp. 669–700.
- [13] Jenny M Cundiff et al. “Affiliation and control in marital interaction: Interpersonal complementarity is present but is not associated with affect or relationship quality”. In: *Personality and Social Psychology Bulletin* 41.1 (2015), pp. 35–51.
- [14] Milton Friedman. *The role of government in education*. Vol. 13. Rutgers University Press New Brunswick, NJ, 1955.
- [15] D. Gale and L. S. Shapley. “College Admissions and the Stability of Marriage”. In: *The American Mathematical Monthly* 69.1 (1962), pp. 9–15.
- [16] Philippe Jehiel. “Analogy-based expectation equilibrium”. In: *Journal of Economic Theory* 123.2 (2005), pp. 81–104.
- [17] Philippe Jehiel and Dov Samet. “Valuation equilibrium”. In: *Theoretical Economics* 2.2 (2007), pp. 163–185.
- [18] Wilhelm Ludvig Johannsen. *The elements of an exact theory of heredity*. 1909.
- [19] Fuhito Kojima. “Robust stability in matching markets”. In: *Theoretical Economics* 6.2 (2011), pp. 257–267.
- [20] Shengwu Li. “Obviously strategy-proof mechanisms”. In: *American Economic Review* 107.11 (2017), pp. 3257–87.
- [21] Qingmin Liu et al. “Stable matching with incomplete information”. In: *Econometrica* 82.2 (2014), pp. 541–587.

- [22] Jinpeng Ma. “Strategy-proofness and the strict core in a market with indivisibilities”. In: *International Journal of Game Theory* 23.1 (1994), pp. 75–83.
- [23] Theodore M Newcomb. “The acquaintance process as a prototype of human interaction.” In: (1961).
- [24] Szilvia Pápai. “Strategyproof assignment by hierarchical exchange”. In: *Econometrica* 68.6 (2000), pp. 1403–1433.
- [25] Alvin E Roth. “Incentive compatibility in a market with indivisible goods”. In: *Economics letters* 9.2 (1982), pp. 127–132.
- [26] Alvin E Roth. “The evolution of the labor market for medical interns and residents: a case study in game theory”. In: *Journal of political Economy* 92.6 (1984), pp. 991–1016.
- [27] Alvin E Roth and Marilda A Oliveira Sotomayor. *Two-sided matching, volume 18 of Econometric Society Monographs*. 1990.
- [28] Lloyd S Shapley and Martin Shubik. “The assignment game I: The core”. In: *International Journal of game theory* 1.1 (1971), pp. 111–130.
- [29] Herbert A. Simon. “Models of Man, Social and Rational: Mathematical Essays on Rational Human Behavior in a Social Setting”. In: *Journal of Philosophy* 59.7 (1962), pp. 177–182.
- [30] Christopher A Sims. “Implications of rational inattention”. In: *Journal of monetary Economics* 50.3 (2003), pp. 665–690.

6 Appendix

6.1 Proofs

Proof. 3.7. Summarising the conditions for stability, one gets:

	GS-stability	LUPV-stability
(IR)	$\begin{cases} \forall m \in M : u_{m\mu(m)}^m \geq u_{m\emptyset}^m \\ \forall f \in F : u_{\mu(f)f}^f \geq u_{\emptyset f}^f \end{cases}$	$\begin{cases} \forall m \in M : v_m(f) \geq v_m(\emptyset) \\ \forall f \in F : v_f(m) \geq v_f(\emptyset) \end{cases}$
(IC)	$\forall (m, f) \in M \times F : \begin{cases} u_{m\mu(m)}^m \geq u_{mf}^m \\ \text{or} \\ u_{\mu(f)f}^f \geq u_{mf}^f \end{cases}$	$\forall (m, f) \in M \times F : \begin{cases} v_m(\mu(m)) \geq v_m(f) \\ \text{or} \\ v_f(\mu(f)) \geq v_f(m) \end{cases}$

I would like to rewrite this table, expliciting the formulas for the utility u and valuation v functions, and deleting immaterial inequalities.

A first remark is that (IR) constraints are identical in both stability definition. GS and LPUV stabilities differ only by the (IC) constraint.

Second, the $(IC)^{LPUV}$ can be decomposed into two subcases:

1. Both agents in the blocking pair are currently matched: $(\mu(m) \neq \emptyset) \cap (\mu(f) \neq \emptyset)$

$$\text{Then: } (IC)^{LPUV} \iff \begin{cases} r_{m\mu(m)} \cdot s_{m\mu(m)} \geq r_{mf} \cdot s_{\mu(f)f} \\ \text{or} \\ r_{\mu(f)f} \cdot s_{\mu(f)f} \geq (1 - r_{mf}) \cdot s_{\mu(f)f} \end{cases}$$

2. At least one agent in the blocking pair is currently unmatched: $\mu(m) = \emptyset$

Suppose for example that $\mu(m) = \emptyset, \mu(f) \neq \emptyset$.

$$\text{Then: } (IC)^{LPUV} \iff \begin{cases} r_{m\mu(m)} \cdot s_{m\mu(m)} \geq r_{mf} \cdot s_{\mu(f)f} \\ \text{or} \\ r_{\mu(f)f} \cdot s_{\mu(f)f} \geq v_f(m) \end{cases}$$

Choosing $v_f(m) = 0$ ensures that the second condition is verified, hence (m, f) not a blocking pair. In this second case, $(IC)^{LPUV}$ imposes no constraint.

	GS-stability	LUPV-stability
(IR)	$\begin{cases} \forall m \in \mu(F) : r_{m\mu(m)} \cdot s_{m\mu(m)} \geq u_{m\emptyset} \\ \forall f \in \mu(M) : r_{\mu(f)f} \cdot s_{\mu(f)f} \geq u_{\emptyset f} \end{cases}$	$\begin{cases} \forall m \in \mu(F) : r_{m\mu(m)} \cdot s_{m\mu(m)} \geq u_{m\emptyset} \\ \forall f \in \mu(M) : r_{\mu(f)f} \cdot s_{\mu(f)f} \geq u_{\emptyset f} \end{cases}$
(IC)	$\forall (m, f) \in M \times F : \begin{cases} u_{m\mu(m)}^m \geq u_{mf}^m \\ \text{or} \\ u_{\mu(f)f}^f \geq u_{mf}^f \end{cases}$	$\forall (m, f) \in \mu(F) \times \mu(M) : \begin{cases} r_{m\mu(m)} \cdot s_{m\mu(m)} \geq v_m(f) \\ \text{or} \\ r_{\mu(f)f} \cdot s_{\mu(f)f} \geq v_f(m) \end{cases}$

□

Proof. 3.8.

(i)

$$(m, f) \text{ blocking } (\mu(m), \mu(f) \neq \emptyset) \iff \begin{cases} r_{m\mu(m)} \cdot s_{m\mu(m)} < r_{mf} \cdot s_{\mu(f)f} & (\overline{IC})_m \\ \text{and} \\ (1 - r_{\mu(f)f}) \cdot s_{\mu(f)f} < (1 - r_{mf}) \cdot s_{m\mu(m)} & (\overline{IC})_f \end{cases}$$

Multiplying both conditions:

$$\begin{aligned}
(\overline{IC})_m \times (\overline{IC})_f &: r_{m\mu(m)} \cdot (1 - r_{\mu(f)f}) \cdot s_{m\mu(m)} \cdot s_{\mu(f)f} < r_{mf} \cdot (1 - r_{mf}) \cdot s_{\mu(f)f} \cdot s_{m\mu(m)} \\
&\Rightarrow r_{mf} \cdot (1 - r_{mf}) > r_{m\mu(m)} \cdot (1 - r_{\mu(f)f}) \\
&\Rightarrow r_{mf} > r_{m\mu(m)} \text{ or } 1 - r_{mf} > 1 - r_{\mu(f)f} \\
&\Rightarrow r_{mf} > r_{m\mu(m)} \text{ or } r_{mf} < r_{\mu(f)f}
\end{aligned}$$

(ii) Without loss of generality, suppose: $s_{m\mu(m)} > s_{\mu(f)f}$ (\star).

$$(IC)_m \cap (\star) \Rightarrow r_{m\mu(m)} < r_{mf}.$$

(iii) By contraposition of (i). □

Proof. 3.9.

(i) Constant bargaining power for both sides: $(IC)_m \times (IC)_f : r(1 - r) > r(1 - r)$. This condition is never verified. So there is no blocking pair in which both agents are matched. Following characterisation of LPUV-stable matchings in proposition 3.7., this is enough for $(IC)^{LPUV}$ to be satisfied and LPUV-stability boils down to (IR).

For (ii) and (iii), the proof is made by applying proposition 3.8.. For each case, the necessary condition for (m, f) to be a blocking pair $r_{mf} \in (0, r_{\mu(f)f}) \cup (r_{m\mu(m)}, 1)$ writes:

(ii) Constant bargaining power for side M : $r_m \in (0, 1 - (1 - r_{\mu(f)f}) \cup (r_m, 1) \iff r_m < r_{\mu(f)}$.

(iii) Regular bargaining power:

$$\begin{aligned}
\frac{\alpha_m}{\alpha_m + \beta_f} &\in \left(0, \frac{r_{\mu(f)}}{\alpha_{\mu(f)} + \beta_f}\right) \cup \left(\frac{\alpha_m}{\alpha_m + \beta_{\mu(m)}}, 1\right) \\
\frac{\beta_f}{\alpha_m + \beta_f} &> \frac{\beta_f}{\alpha_{\mu(f)} + \beta_f} \text{ or } \frac{\alpha_m}{\alpha_m + \beta_f} > \frac{\alpha_m}{\alpha_m + \beta_{\mu(m)}} \\
\alpha_m &< \alpha_{\mu(f)} \text{ or } \beta_f < \beta_{\mu(m)}
\end{aligned}$$
□

Proof. 3.10.

$$\begin{aligned}
(m_i, f_j) \text{ is a LPUV-blocking pair} &\iff \begin{cases} r_{ii} \cdot s_{ii} < r_{ij} \cdot s_{jj} \\ (1 - r_{jj}) \cdot s_{jj} < (1 - r_{ij}) \cdot s_{ii} \end{cases} \\
&\iff \frac{1 - r_{jj}}{1 - r_{ij}} < \frac{s_{ii}}{s_{jj}} < \frac{r_{ij}}{r_{ii}}
\end{aligned}$$

Symmetrically:

$$\begin{aligned}
(m_j, f_i) \text{ is a LPUV-blocking pair} &\iff \frac{1 - r_{ii}}{1 - r_{ji}} < \frac{s_{jj}}{s_{ii}} < \frac{r_{ji}}{r_{jj}} \\
&\iff \frac{r_{jj}}{r_{ji}} < \frac{s_{ii}}{s_{jj}} < \frac{1 - r_{ji}}{1 - r_{ii}}
\end{aligned}$$

Finally:

$\mu = [(m_k, f_k)]_{k \in \{1, \dots, n\}}$ is LPUV-stable

$$\iff \forall i, j \in \{1, \dots, n\}, i \neq j : (m_i, f_j) \text{ is not a LPUV-blocking pair}$$

$$\iff \forall 1 \leq i < j \leq n : (m_i, f_j) \text{ and } (m_j, f_i) \text{ are not LPUV-blocking pairs}$$

$$\iff \forall 1 \leq i < j \leq n : \frac{s_{ii}}{s_{jj}} \in \mathbb{R}_+^* \setminus \left(\frac{1 - r_{jj}}{1 - r_{ij}}, \frac{r_{ij}}{r_{ii}}\right) \cap \left(\frac{r_{jj}}{r_{ji}}, \frac{1 - r_{ji}}{1 - r_{ii}}\right)$$

$$\iff \forall 1 \leq i < j \leq n : \frac{s_{ii}}{s_{jj}} \in \left(0, \min\left(\frac{1 - r_{jj}}{1 - r_{ij}}, \frac{r_{jj}}{r_{ji}}\right)\right] \cup \left[\max\left(\frac{r_{ij}}{r_{ii}}, \frac{1 - r_{ji}}{1 - r_{ii}}\right), +\infty\right)$$

I denote by I the union of the intervals written above. I can be empty or non-empty depending on the relation between the splitting rules.

I make a disjunction of cases:

$$1. \begin{cases} r_{ii}(1 - r_{jj}) > r_{ij}(1 - r_{ij}) \\ r_{jj}(1 - r_{ii}) > r_{ji}(1 - r_{ji}) \end{cases} \text{ Then:}$$

$$\frac{1 - r_{jj}}{1 - r_{ij}} > \frac{r_{ij}}{r_{ii}}, \quad \frac{r_{jj}}{r_{ji}} > \frac{1 - r_{ji}}{1 - r_{ii}} \\ \Rightarrow I = \mathbb{R}_+^*$$

$$2. \quad \text{a. } r_{ii}(1 - r_{jj}) > r_{ij}(1 - r_{ij}) \text{ and } \frac{s_{ii}}{s_{jj}} \in \left(0, \frac{r_{jj}}{r_{ji}}\right] \cup \left[\frac{1 - r_{ji}}{1 - r_{ii}}, +\infty\right)$$

Then:

$$\frac{1 - r_{jj}}{1 - r_{ij}} < \frac{r_{ij}}{r_{ii}} \\ \Rightarrow I = \left(0, \frac{r_{jj}}{r_{ji}}\right] \cup \left[\frac{1 - r_{ji}}{1 - r_{ii}}, +\infty\right)$$

$$\text{b. } (1 - r_{ii})r_{jj} > (1 - r_{ij})r_{ji} \text{ and } \frac{s_{ii}}{s_{jj}} \in \left(0, \frac{1 - r_{jj}}{1 - r_{ij}}\right] \cup \left[\frac{r_{ij}}{r_{ii}}, +\infty\right)$$

Then:

$$\frac{r_{jj}}{r_{ji}} < \frac{1 - r_{ji}}{1 - r_{ii}} \\ \Rightarrow I = \left(0, \frac{1 - r_{jj}}{1 - r_{ij}}\right] \cup \left[\frac{r_{ij}}{r_{ii}}, +\infty\right)$$

$$3. \begin{cases} r_{ii}(1 - r_{jj}) \leq r_{ij}(1 - r_{ij}) \\ r_{jj}(1 - r_{ii}) \leq r_{ji}(1 - r_{ji}) \end{cases} \text{ Then:}$$

$$\frac{1 - r_{jj}}{1 - r_{ij}} < \frac{r_{ij}}{r_{ii}}, \quad \frac{r_{jj}}{r_{ji}} < \frac{1 - r_{ji}}{1 - r_{ii}} \\ \Rightarrow I = \left(0, \min\left(\frac{1 - r_{jj}}{1 - r_{ij}}, \frac{r_{jj}}{r_{ji}}\right)\right] \cup \left[\max\left(\frac{r_{ij}}{r_{ii}}, \frac{1 - r_{ji}}{1 - r_{ii}}\right), +\infty\right)$$

□

Proof. 3.11.

$$\begin{aligned} (\alpha_i, \beta_j) \text{ is a GS blocking pair} &\iff \begin{cases} \frac{\alpha_i}{\alpha_i + \beta_i} s_{ii} < \frac{\alpha_i}{\alpha_i + \beta_j} s_{ij} \\ \frac{\beta_j}{\alpha_i + \beta_j} s_{jj} < \frac{\beta_j}{\alpha_i + \beta_j} s_{ij} \end{cases} \\ &\iff \begin{cases} \frac{s_{ii}}{\alpha_i + \beta_i} < \frac{s_{ij}}{\alpha_i + \beta_j} \\ \frac{s_{jj}}{\alpha_j + \beta_j} < \frac{s_{ij}}{\alpha_i + \beta_j} \end{cases} \\ &\iff \max\left(\frac{s_{ii}}{\alpha_i + \beta_i}; \frac{s_{jj}}{\alpha_j + \beta_j}\right) < \frac{s_{ij}}{\alpha_i + \beta_j} \end{aligned}$$

By contraposition:

$$(\alpha_i, \beta_j) \text{ is not a GS blocking pair} \iff \frac{s_{ij}}{\alpha_i + \beta_j} \leq \max\left(\frac{s_{ii}}{\alpha_i + \beta_i}; \frac{s_{jj}}{\alpha_j + \beta_j}\right)$$

By definition,

$\mu = ((\alpha_1, \beta_1); \dots; (\alpha_n, \beta_n))$ is GS-stable.

$$\iff \forall (i, j) \in \{1, \dots, n\}^2, i \neq j : (\alpha_i, \beta_j) \text{ is not a GS blocking pair.}$$

$$\iff \forall 1 \leq i < j \leq n : (\alpha_i, \beta_j), (\alpha_j, \beta_i) \text{ are not GS blocking pairs.}$$

$$\iff \forall 1 \leq i < j \leq n : \begin{cases} \frac{s_{ij}}{\alpha_i + \beta_j} \leq \max\left(\frac{s_{ii}}{\alpha_i + \beta_i}; \frac{s_{jj}}{\alpha_j + \beta_j}\right) \\ \frac{s_{ji}}{\alpha_j + \beta_i} \leq \max\left(\frac{s_{ii}}{\alpha_i + \beta_i}; \frac{s_{jj}}{\alpha_j + \beta_j}\right) \end{cases}$$

$$\iff \forall 1 \leq i < j \leq n : \max\left(\frac{s_{ij}}{\alpha_i + \beta_j}; \frac{s_{ji}}{\alpha_j + \beta_i}\right) \leq \max\left(\frac{s_{ii}}{\alpha_i + \beta_i}; \frac{s_{jj}}{\alpha_j + \beta_j}\right)$$

□

Proof. 3.12.

For case 1. and 2., I simply apply theorem 3.12. and go through uninteresting algebra.

Case 3 is defined by: $(\alpha_i + \beta_i)(\alpha_j + \beta_j) > \max((\alpha_j + \beta_i)^2, (\alpha_i + \beta_j)^2)$.

Let me formalise the relation between the four bargaining scores, I denote: $x = \alpha_i + \beta_i$, $y = \alpha_j + \beta_j$, $z = \alpha_i + \beta_i$.

$$(\alpha_i + \beta_i)(\alpha_j + \beta_j) > \max((\alpha_j + \beta_i)^2, (\alpha_i + \beta_j)^2) \iff \begin{cases} xy > z^2 \\ xy > (x + y - z)^2 \end{cases}$$

I set:

$$\begin{aligned} g_{xy} : [0, x + y] &\rightarrow \mathbb{R}_+ & h_{xy} : [0, x + y] &\rightarrow \mathbb{R}_+ \\ z &\mapsto xy - z^2 & z &\mapsto xy - (x + y - z)^2 \\ g'_{xy}(z) &= -2z < 0 & h'_{xy}(z) &= 2(x + y - z) > 0 \end{aligned}$$

One gets the following variations and signs for the two functions defined:

z	0	\sqrt{xy}	$x + y - \sqrt{xy}$	$x + y$
$g_{xy}(z)$	xy	0		$xy - (x + y)^2$
$g_{xy}(z)$	+	0	-	-
z	0	\sqrt{xy}	$x + y - \sqrt{xy}$	$x + y$
$h_{xy}(z)$	$xy - (x + y)^2$		0	xy
$h_{xy}(z)$	-	-	0	+

So, it is never the case that $\begin{cases} g_{xy}(z) > 0 \\ h_{xy}(z) > 0 \end{cases}$. Case 3. is absurd. □

Proof. $(1^*) \iff (1^{**})$

- Set (1^*) , and suppose without loss of generality that: $\begin{cases} \alpha_i \leq \alpha_j \\ \beta_i \leq \beta_j \end{cases}$. This implies (using the notations and results of the proof just below):

$$\alpha_i + \beta_i \leq \alpha_j + \beta_j \Rightarrow x \leq y \Rightarrow |x - y| = y - x \Rightarrow \begin{cases} r = x \\ r' = y \end{cases}$$

$$\alpha_i + \beta_j \leq \alpha_j + \beta_i \leq \alpha_j + \beta_j \Rightarrow x \leq z \leq y$$

$$\Rightarrow f_{xy}(z) \geq 0$$

So (1^{**}) is verified.

- Set (1^{**}) and $\alpha_i \leq \alpha_j$. Suppose $\beta_j \leq \beta_i$ (negative assortativity). Then one would have:

$$\alpha_i \leq \alpha_j \Rightarrow \alpha_i + \beta_j \leq \alpha_j + \beta_j$$

$$\beta_j \leq \beta_i \Rightarrow \alpha_i + \beta_j \leq \alpha_i + \beta_i$$

$$\Rightarrow (\alpha_i + \beta_j)(\alpha_j + \beta_i) \leq (\alpha_i + \beta_i)(\alpha_j + \beta_j) \quad \overline{(1^{**})} \text{ (absurd)}$$

So $\beta_i \leq \beta_j$, (1^*) is verified. Symmetrically: $\alpha_i \geq \alpha_j$ and $(1^{**}) \Rightarrow \beta_i \geq \beta_j$. □

Proof. Functional analysis (1^{**})

The condition I would like to interpret writes: $(1^{**}) : (\alpha_1 + \beta_1)(\alpha_2 + \beta_2) < (\alpha_1 + \beta_2)(\alpha_2 + \beta_1)$. Using the same notations as in the preceding proof: $(1^{**}) : xy < z(x + y - z)$. I define:

$$f_{xy} : \begin{cases} [0, x + y] \rightarrow \mathbb{R}_+^* \\ z \mapsto (x + y - z)z - xy \end{cases}$$

$$= x + y - 2z, \quad f'_{xy}(z) = 0 \iff z = \frac{x + y}{2}$$

$$f_{xy}\left(\frac{x + y}{2}\right) = \frac{(x + y)^2}{4} - xy = \frac{1}{4}(x^2 + 2xy - y^2 - 4xy) = \frac{(x - y)^2}{4} > 0$$

I get the following variations and sign:

z	0	r	$\frac{x+y}{2}$	r'	$x + y$
$f_{xy}(z)$	$-xy$	0	$\frac{(x-y)^2}{4}$	0	$-xy$
$f_{xy}(z)$	-	0	+	+	0

r and r' are the solutions of: $f_{xy}(z) = 0 \iff z^2 - (x + y)z + xy = 0$. I find:

$$r = \frac{1}{2}(x + y - \sqrt{|x - y|}) \quad r' = \frac{1}{2}(x + y + \sqrt{|x - y|})$$

Finally:

$$\begin{aligned}
f(z) \geq 0 &\iff r \leq z \leq r' \\
&\iff \frac{1}{2}(x + y - \sqrt{|x - y|}) \leq z \leq \frac{1}{2}(x + y + \sqrt{|x - y|}) \\
&\iff \beta_1 + \alpha_2 - \sqrt{|(\alpha_1 - \beta_1) - (\alpha_2 + \beta_2)|} \leq \alpha_1 + \beta_2 \leq \beta_1 + \alpha_2 + \sqrt{|(\alpha_1 - \beta_1) - (\alpha_2 + \beta_2)|}
\end{aligned}$$

This happens when:

- $(1^{**})'$ $\alpha_1 + \beta_2 \simeq \alpha_2 + \beta_1$
 $(1^{**})'$ means that on average, the two potential blocking pairs wield comparable bargaining powers. Combined with positive assortativity within matched pairs, this gives a kind of negative assortativity within potential blocking pairs. This comes as no surprise. Imagine one potential blocking pair would match two very weak agents, and the other blocking pair would match two very strong agents. Then, even if assortativity in matched pairs is enforced, fairness of the split is not satisfactory enough. Then for not to wild surpluses, the weak agents would block for sure. This second condition though is not strict negative assortativity, using vague vocabulary on purpose, I would say it imposes “not-too-positive assortativity” within potential blocking pairs.
- $(1^{**})''$ $|(\alpha_1 + \beta_1) - (\alpha_2 + \beta_2)| \gg 0$
 $(1^{**})''$ means that there is large heterogeneity in bargaining scores between different matched pairs. This makes a lot of sense. Suppose there is high heterogeneity and the two weakest matched agents are far much weaker than the rest of the agents. Then, for a wide array of surplus systems, none of those agents want to deviate. The $(1^{**})'$ and $(1^{**})''$ complement one another so that the constraining forces of each of the two conditions are interdependent.

□

Proof. Extreme homogeneity / heterogeneity in bargaining scores within sides.

- Set $\alpha \in (0, 1)$ s.t. $\alpha_1 = \alpha_2 = \alpha$. Then:

$$\begin{aligned}
(\alpha_1 + \beta_2)^2 - (\alpha_1 + \beta_1)(\alpha_2 + \beta_2) &= \alpha^2 + 2\alpha\beta_2 + \beta_2^2 - \alpha^2 - \alpha\beta_1 - \alpha\beta_2 - \beta_1\beta_2 \\
&= \alpha\beta_2 + \beta_2^2 - \alpha\beta_1 - \beta_1\beta_2 \\
&= (\alpha + \beta_2)(\beta_2 - \beta_1)
\end{aligned}$$

$$(\alpha_2 + \beta_1)^2 - (\alpha_1 + \beta_1)(\alpha_2 + \beta_2) = (\alpha + \beta_1)(\beta_1 - \beta_2)$$

So: $(\alpha_1 + \beta_2)^2 > (\alpha_1 + \beta_1)(\alpha_2 + \beta_2) \Rightarrow (\alpha_2 + \beta_1)^2 < (\alpha_1 + \beta_1)(\alpha_2 + \beta_2)$ and reciprocally

(1) is not verified.

- Set $\alpha_1 \simeq +\infty$, $\alpha_2 \simeq 0$. Then:

$$\begin{aligned}
(\alpha_1 + \beta_1)(\alpha_2 + \beta_2) &\simeq (+\infty + \beta_1)(\beta_2) \simeq +\infty \\
(\alpha_2 + \beta_1)^2 &\simeq \beta_1^2 \\
\Rightarrow (\alpha_1 + \beta_1)(\alpha_2 + \beta_2) &< (\alpha_2 + \beta_1)^2
\end{aligned}$$

(1) is not verified.

□

Proof. 3.13. I normalise: $s = 1$. By theorem 3.11.:

$$\begin{aligned}
\mu \in \Sigma^{GS} &\iff \forall 1 \leq i < j \leq n : \max\left(\frac{s}{\alpha_i + \beta_j}; \frac{s}{\alpha_j + \beta_i}\right) \leq \max\left(\frac{s}{\alpha_i + \beta_i}; \frac{s}{\alpha_j + \beta_j}\right) \\
&\iff \forall 1 \leq i < j \leq n : \max\left(\frac{1}{\alpha_i + \beta_j}; \frac{1}{\alpha_j + \beta_i}\right) \leq \max\left(\frac{1}{\alpha_i + \beta_i}; \frac{1}{\alpha_j + \beta_j}\right) \\
&\iff \forall 1 \leq i < j \leq n : \min(\alpha_i + \beta_j, \alpha_j + \beta_i) \geq \min(\alpha_i + \beta_i, \alpha_j + \beta_j) \\
&\iff \forall 1 \leq i < j \leq n : \min(\alpha_i + \beta_j, \alpha_j + \beta_i) \geq \alpha_j + \beta_j \\
&\iff \forall (i, j) \in \{1, \dots, n\}^2 : \alpha_i + \beta_i > \alpha_j + \beta_j \Rightarrow \begin{cases} \alpha_i \geq \alpha_j \\ \beta_i \geq \beta_j \end{cases} \\
&\iff \mu = \mu^a
\end{aligned}$$

□

Proof. 3.14. Set $\mu \in \Sigma_F$. Suppose $\exists s_\mu > 0$ s.t. $\forall i \in \{1, \dots, n\} : s_{m_i \mu(m_i)} = s_\mu$.

- Suppose $\mu \in \Sigma^{LPUV}$. By theorem 3.12., we are in case (1), (2a) or (2b). Now suppose, without loss of generality that we are in case 2.a.. Under equal realised surpluses $\forall (i, j) \in \{1, \dots, n\}$, case 2.a. writes:

$$\begin{cases} (\alpha_i + \beta_i)(\alpha_j + \beta_j) < (\alpha_j + \beta_i)^2 \\ (\alpha_i + \beta_i)(\alpha_j + \beta_j) \geq (\alpha_i + \beta_j)^2 \end{cases} \text{ and } 1 \in \left(0, \frac{\alpha_i + \beta_j}{\alpha_j + \beta_j}\right] \cup \left[\frac{\alpha_i + \beta_i}{\alpha_i + \beta_j}, +\infty\right)$$

So: $1 \leq \frac{\alpha_i + \beta_j}{\alpha_j + \beta_j}$ or $1 \geq \frac{\alpha_i + \beta_i}{\alpha_i + \beta_j} \Rightarrow \alpha_j \leq \alpha_i$ or $\beta_i \leq \beta_j$

Suppose, without loss of generality that: $\alpha_j \leq \alpha_i$. By (2a*), $\alpha_i \leq \alpha_j$, so $\alpha_i = \alpha_j$. Then:

$$\begin{aligned}
\alpha_j = \alpha_i &\Rightarrow (\alpha_j + \beta_i)(\alpha_j + \beta_j) = (\alpha_i + \beta_i)(\alpha_j + \beta_j) \\
&\Rightarrow (\alpha_j + \beta_i)(\alpha_j + \beta_j) < (\alpha_j + \beta_i)^2 \\
&\Rightarrow \beta_j < \beta_i
\end{aligned}$$

This contradicts (2a*) : $\beta_i \leq \beta_j$. So case (2a) never arises when $\mu \in \Sigma^{LPUV}$. By symmetry, case (2b) never arises either.

We are always in case (1) By theorem 3.12. still, and necessary condition (1*): $\mu = \mu^a$.

- Take μ^a , and (m_i, f_j) a potential blocking pair. Assume, without loss of generality that $i < j$. By LPUV-rationality:

$$\begin{aligned}
v_{f_j}(m_i) &= \frac{\beta_j}{\alpha_i + \beta_j} \cdot s \quad v_{f_j}(m_j) = \frac{\beta_j}{\alpha_j + \beta_j} \cdot s \\
i < j &\Rightarrow \alpha_i > \alpha_j \Rightarrow v_{f_j}(m_i) < v_{f_j}(m_j)
\end{aligned}$$

f_j does not want to block with m_i . $\forall i \neq j$, (m_j, f_j) is not a LPUV-blocking pair. So $\mu^a \in \Sigma^{LPUV}$.

□

Proof. 3.17. Suppose $\forall i \in \{1, \dots, n\} : \alpha_i = \beta_i$.

$$\begin{aligned}
\forall 1 \leq i < j \leq n, & (\alpha_i + \beta_i)(\alpha_j + \beta_j) \leq \min((\alpha_i + \beta_j)^2; (\alpha_j + \beta_i)^2) \\
&\iff 4\alpha_i\alpha_j < (\alpha_i + \alpha_j)^2 \\
&\iff 2\alpha_i\alpha_j < \alpha_i^2 + \alpha_j^2 \\
&\iff 0 < (\alpha_i - \alpha_j)^2 \text{ true}
\end{aligned}$$

So condition (1) is verified for any potential blocking pair. $\mu^a \in \Sigma^{LPUV}$.

□

Proof. 3.18. Set $\alpha_1 \geq \dots \geq \alpha_n$, $\beta_1 \geq \dots \geq \beta_n$, and $s_{11} \leq \dots \leq s_{nn}$.

I define an induction hypothesis: $H(k)$: m_k, f_k are part of no blocking pairs. And I make a (strong descending) induction proof.

- Base case.

$$\forall k \in \{1, \dots, n-1\} : v_{m_n}(f_k) = \frac{\alpha_n}{\alpha_n + \beta_k} \cdot s_{kk} \leq \frac{\alpha_n}{\alpha_n + \beta_n} \cdot s_{nn} = v_{m_n}(f_n)$$

So m_n is part of no blocking pair. Symmetrically, f_n is part of no blocking pair. $H(n)$ is verified.

- Induction step. $l \in \{1, n-1\}$. Suppose $H(n), \dots, H(l+1)$ is true.

$$\forall k \in \{1, \dots, l-1\} : v_{m_l}(f_k) = \frac{\alpha_l}{\alpha_l + \beta_k} \cdot s_{kk} \leq \frac{\alpha_n}{\alpha_n + \beta_n} \cdot s_{kk} = v_{m_l}(f_l)$$

So (m_l, f_k) , $k \in \{1, \dots, l-1\}$ are not blocking pairs. By $H(n), \dots, H(l+1)$, (m_l, f_k) , $k \in \{l+1, \dots, n\}$ are not blocking pair either. So m_l is part of no blocking pair. Symmetrically, f_l is part of no blocking pair. $H(l)$ is verified. □

Proof. 3.19. $\forall (i, j) \in \{1, \dots, n\} : s_{ij} = s(\alpha_i, \beta_j) = \alpha_i + \beta_j$.

- (i) $u_{ij} = \alpha_i$. So, u_{ij} is constant with respects to j . In a full matching, there is never any profitable deviation hence no blocking pairs. $\Sigma_F \subset \Sigma^{GS}$.
- (ii) – Take the assortative matching μ^a . Set a potential blocking pair (m_i, f_j) . Incentive compatibility constraints write:

$$\begin{aligned} (IC)_{m_i} : v_{m_i}(f_j) > v_{m_i}(f_i) &\iff \frac{\alpha_i}{\alpha_i + \beta_j} \cdot (\alpha_i + \beta_j) > \frac{\alpha_i}{\alpha_i + \beta_i} \cdot (\alpha_i + \beta_i) \iff \alpha_j < \alpha_i \\ (IC)_{f_j} : v_{f_j}(m_i) > v_{f_j}(f_j) &\iff \beta_i < \beta_j \\ i < j \Rightarrow \beta_i > \beta_j \Rightarrow \overline{(IC)_{m_i}}, \quad i > j \Rightarrow \alpha_i < \alpha_j \Rightarrow \overline{(IC)_{f_j}} \end{aligned}$$

So (m_i, f_j) is not a blocking pair. $\mu^a \in \Sigma^{LPUV}$.

- Set $\mu \in \Sigma_F$, $\mu \neq \mu^a$. By definition of μ , $\exists i \neq j$ s.t. $\mu(m_i) = f_j$. By basic properties of balanced (full) bipartite graph, $\exists i < i', j > j'$, s.t. $\mu(i) = j$, $\mu(j) = j'$. Let me consider the pair made of the weakest agent on each of the two sides: $(m_{i'}, f_j)$. One has that:

$$\begin{aligned} i < i' \Rightarrow \alpha_i > \alpha_{i'} \Rightarrow \alpha_{\mu(j)} > \alpha_{i'} \quad (IC)_{m_{i'}} \\ j > j' \Rightarrow \beta_j < \beta_{j'} \Rightarrow \beta_j < \beta_{\mu(i')} \quad (IC)_{f_j} \end{aligned}$$

So $(m_{i'}, f_j)$ is a blocking pair. μ is LPUV-unstable. $\Sigma_F^{LPUV} \subseteq \{\mu^a\}$.

By two-way inclusion: $\Sigma_F^{LPUV} = \{\mu^a\}$. □

Proof. 3.20. Set $(i, j) \in \{1, \dots, n\}^2$.

$$\begin{aligned} (m_i, f_j) \text{ GS-blocking pair} &\iff \begin{cases} \frac{\alpha_i}{\alpha_i + \beta_j} \cdot s(\alpha_i, \beta_j) > \frac{\alpha_i}{\alpha_i + \beta_{\mu(i)}} \cdot s(\alpha_i, \beta_{\mu(i)}) \\ \frac{\beta_j}{\alpha_i + \beta_j} \cdot s(\alpha_i, \beta_j) > \frac{\beta_j}{\alpha_{\mu(i)} + \beta_j} \cdot s(\alpha_{\mu(i)}, \beta_j) \end{cases} \\ &\iff \frac{s(\alpha_i, \beta_j)}{\alpha_i + \beta_j} > \max \left(\frac{s(\alpha_i, \beta_{\mu(i)})}{\alpha_i + \beta_{\mu(i)}}, \frac{s(\alpha_{\mu(i)}, \beta_j)}{\alpha_{\mu(i)} + \beta_j} \right) \end{aligned}$$

Set: $f(x, y) = \frac{s(x, y)}{x + y}$. Then:

$$(m_i, f_j) \text{ GS-blocking pair} \iff f(\alpha_i, \beta_j) > \max(f(\alpha_i, \beta_{\mu(i)}), f(\alpha_{\mu(j)}, \beta_j))$$

Finally:

$$\mu \text{ is GS-stable} \iff \forall 1 \leq i < j \leq n : \max(f(\alpha_i, \beta_j), f(\alpha_j, \beta_i)) > \max(f(\alpha_i, \beta_{\mu(i)}), f(\alpha_{\mu(j)}, \beta_j))$$

$$\mu^a \text{ is GS-stable} \iff \forall 1 \leq i < j \leq n : \max(f(\alpha_i, \beta_j), f(\alpha_j, \beta_i)) > \max(f(\alpha_i, \beta_i), f(\alpha_j, \beta_j))$$

Suppose f is increasing in both arguments. Then, $\forall 1 \leq i < j \leq n$:

$$\begin{aligned} \beta_j < \beta_i &\Rightarrow f(\alpha_i, \beta_j) < f(\alpha_i, \beta_i) \\ \alpha_i < \alpha_j &\Rightarrow f(\alpha_j, \beta_i) < f(\alpha_i, \beta_i) \\ \max(f(\alpha_i, \beta_i), f(\alpha_j, \beta_j)) &= f(\alpha_i, \beta_i) \end{aligned}$$

Suppose f is decreasing in both arguments. Then, $\forall 1 \leq i < j \leq n$:

$$\begin{aligned} \beta_j < \beta_i &\Rightarrow f(\alpha_j, \beta_i) < f(\alpha_j, \beta_j) \\ \alpha_i < \alpha_j &\Rightarrow f(\alpha_i, \beta_j) < f(\alpha_j, \beta_j) \\ \max(f(\alpha_i, \beta_i), f(\alpha_j, \beta_j)) &= f(\alpha_j, \beta_j) \end{aligned}$$

So, f is monotone in one argument (by symmetry of f , implies same monotonicity in the second argument). By the inequalities above, $\mu^a \in \Sigma^{GS}$.

If f is differentiable (equivalent to s differentiable):

$$\frac{\partial f(x, y)}{\partial x} = \frac{1}{(x + y)^2} \left(\frac{\partial s(x, y)}{\partial x} \cdot (x + y) - s(x, y) \right)$$

Thus, in analytical terms:

$$\mu^a \in \Sigma^{GS} \iff \frac{\partial s(x, y)}{\partial x} \cdot (x + y) - s(x, y) \text{ has constant sign.}$$

□

6.2 Mirror cases: Results

6.2.1 Case 1: Egocentric estimation of match surplus

Suppose that the mistake a LPUV-rational agent is that he supposes his side of the market fully determines the surplus. The corresponding valuation function writes:

$$w_m(f) = \begin{cases} u_{mf}^m & \text{if } f = \emptyset \\ r_{mf} \cdot s_{m\mu(m)} & \text{if } f \neq \emptyset, \mu(f) \neq \emptyset \\ \in \mathbb{R}_+ & \text{if } f \neq \emptyset, \mu(f) = \emptyset \end{cases}$$

Set a matching market such that $\#M = \#F = n$ and number the individuals $(m_i)_{i \in \{1, \dots, n\}}$, $(f_j)_{j \in \{1, \dots, n\}}$. I give necessary and sufficient conditions for the modified LPUV-stability:

Proposition 6.1. Set a matching $\mu = [(m_k, f_k)_{k \in \{1, \dots, n\}}]$.

(i) μ is LPUV-stable if and only if one of the four following conditions is satisfied:

$$(1) \begin{cases} r_{ii} \geq r_{ij} \\ \text{and} \\ r_{jj} \geq r_{ji} \end{cases}, \quad (2) \begin{cases} r_{jj} \leq r_{ij} \\ \text{and} \\ r_{ii} \leq r_{ji} \end{cases}, \quad (3) r_{ij} \leq r_{ii} \leq r_{ji}, \quad (4) r_{ji} \leq r_{jj} \leq r_{ij}$$

(ii) Suppose the splitting is regular. μ is LPUV-stable if and only if μ is (weakly) positively assortative with respects to bargaining scores: $\forall (i, j) \in \{1, \dots, n\} : \alpha_i \geq \alpha_j \iff \beta_i \geq \beta_j$.

Two elements are worth noting. First, LPUV-stability of a matching depends only on splitting rules, equivalently, it is robust to any surplus system. The rationale behind this result is that the splitting rule is the only source of heterogeneity on the estimated match utilities from a given agent's perspective. In cases (1), (2), (3), (4) respectively, the man side, the woman side, the pair (m_i, f_i) , the pair (m_j, f_j) are not willing to block. This is enough to guarantee LPUV-stability. Second, under regular splitting rule, positive assortativity is a necessary and sufficient condition for LPUV-stability. The intuition is very simple. Because an agent m_i assesses the surplus to be invariant across different partner, the only aggregate relevant for matching is the bargaining score of the potential partner. Any agent wants to deviate to a weaker partner. Now suppose there is a negative assortative couple of matched pairs: $\alpha_i \geq \alpha_j, \beta_i \leq \beta_j$. Then the two weakest agents m_j and f_i block together.

Proof. 6.1. As usual, I begin by characterising LPUV-blocking pairs:

$$\begin{aligned} (m_i, f_j) \text{ form a LPUV-blocking pair} &\iff \begin{cases} r_{ii} \cdot s_{ii} < (1 - r_{ij}) \cdot s_{ii} \\ r_{jj} \cdot s_{jj} < (1 - r_{ij}) \cdot s_{jj} \end{cases} \\ &\iff \begin{cases} r_{ii} < r_{ij} \\ r_{jj} < 1 - r_{ij} \end{cases} \end{aligned}$$

This allows me to characterise LPUV-stable matchings in the general case:

$$\begin{aligned} \mu \text{ is LPUV-stable} &\iff \forall (i, j) \in \{1, \dots, n\} : (m_i, f_j), (m_j, f_i) \text{ are not LPUV-blocking pairs} \\ &\iff \begin{cases} r_{ii} \geq r_{ij} \\ \text{or} \\ r_{jj} \leq r_{ij} \end{cases} \text{ and } \begin{cases} r_{jj} \geq r_{ji} \\ \text{or} \\ r_{ii} \leq r_{ji} \end{cases} \\ &\iff (1) \begin{cases} r_{ii} \geq r_{ij} \\ \text{and} \\ r_{jj} \geq r_{ji} \end{cases} \text{ or } (2) \begin{cases} r_{jj} \leq r_{ij} \\ \text{and} \\ r_{ii} \leq r_{ji} \end{cases} \text{ or } (3) r_{ij} \leq r_{ii} \leq r_{ji} \text{ or } (4) r_{ji} \leq r_{jj} \leq r_{ij} \end{aligned}$$

This proves (i). Applying result (i) to regular splitting rule, one gets:

$$\begin{aligned} (1) \begin{cases} \frac{\alpha_i}{\alpha_i + \beta_i} \geq \frac{\alpha_i}{\alpha_i + \beta_j} \\ \text{and} \\ \frac{\alpha_j}{\alpha_j + \beta_j} \geq \frac{\alpha_j}{\alpha_j + \beta_i} \end{cases} &\iff \begin{cases} \beta_i \leq \beta_j \\ \text{and} \\ \beta_j \leq \beta_i \end{cases} \iff \beta_i = \beta_j \\ (2) \begin{cases} \frac{\alpha_j}{\alpha_j + \beta_j} \leq \frac{\alpha_i}{\alpha_i + \beta_j} \\ \text{and} \\ \frac{\alpha_i}{\alpha_i + \beta_i} \leq \frac{\alpha_j}{\alpha_j + \beta_i} \end{cases} &\iff \begin{cases} \alpha_j \leq \alpha_i \\ \text{and} \\ \alpha_i \leq \alpha_j \end{cases} \iff \alpha_i = \alpha_j \\ (3) \frac{\alpha_i}{\alpha_i + \beta_j} \leq \frac{\alpha_i}{\alpha_i + \beta_i} \leq \frac{\alpha_j}{\alpha_j + \beta_i} &\iff \begin{cases} \alpha_i \leq \alpha_j \\ \beta_i \leq \beta_j \end{cases} \\ (4) \frac{\alpha_j}{\alpha_j + \beta_i} \leq \frac{\alpha_j}{\alpha_j + \beta_j} \leq \frac{\alpha_i}{\alpha_i + \beta_j} &\iff \begin{cases} \alpha_j \leq \alpha_i \\ \beta_j \leq \beta_i \end{cases} \end{aligned}$$

This proves (ii). □

6.2.2 Case 2: Correct estimation of surplus, incorrect estimation of splitting rule

Suppose the mistake that a LPUV-rational agent makes is that he assumes the opposite side of the market fully determines the splitting rule. The corresponding valuation writes:

$$v(f) = \begin{cases} u_{mf}^m & \text{if } f = \emptyset \\ r_{\mu(f)f} \cdot s_{mf} & \text{if } f \neq \emptyset, \mu(f) \neq \emptyset \\ \in \mathbb{R}_+ & \text{if } f \neq \emptyset, \mu(f) = \emptyset \end{cases}$$

Set a matching market such that $\#M = \#F = n$ and number the individuals $(m_i)_{i \in \{1, \dots, n\}}$, $(f_j)_{j \in \{1, \dots, n\}}$. I give necessary and sufficient conditions for the modified LPUV-stability:

Proposition 6.2. Set a matching $\mu = [(m_k, f_k)]_{k \in \{1, \dots, n\}}$.

$$\mu \text{ is LPUV-stable} \iff \forall 1 \leq i < j \leq n : \begin{cases} s_{ij} \leq \max\left(\frac{r_{ii}}{r_{jj}} \cdot s_{ii}, \frac{1 - r_{jj}}{1 - r_{ii}} \cdot s_{jj}\right) \\ s_{ji} \leq \max\left(\frac{1 - r_{ii}}{1 - r_{jj}} \cdot s_{ii}, \frac{r_{jj}}{r_{ii}} \cdot s_{jj}\right) \end{cases}$$

What is interesting to note is that here LPUV-stability embeds some notion of welfare maximisation: realised surpluses are required to be high with respects to unrealised surpluses.

Proof. 6.2.

$$\begin{aligned} (\alpha_i, \beta_j) \text{ is a LPUV-blocking pair} &\iff \begin{cases} r_{ii} \cdot s_{ii} < r_{jj} \cdot s_{ij} \\ (1 - r_{jj}) \cdot s_{jj} < (1 - r_{ii}) \cdot s_{ij} \end{cases} \\ &\iff s_{ij} > \max\left(\frac{r_{ii}}{r_{jj}} \cdot s_{ii}, \frac{1 - r_{jj}}{1 - r_{ii}} \cdot s_{jj}\right) \\ \mu \text{ is LPUV-stable} &\iff \forall 1 \leq i < j \leq n : (m_i, f_j) \text{ and } (m_j, f_i) \text{ are not LPUV-blocking pairs} \\ &\iff \begin{cases} s_{ij} \leq \max\left(\frac{r_{ii}}{r_{jj}} \cdot s_{ii}, \frac{1 - r_{jj}}{1 - r_{ii}} \cdot s_{jj}\right) \\ s_{ji} \leq \max\left(\frac{1 - r_{ii}}{1 - r_{jj}} \cdot s_{ii}, \frac{r_{jj}}{r_{ii}} \cdot s_{jj}\right) \end{cases} \end{aligned}$$

□

6.3 Link with Valuation Equilibrium

6.3.1 Valuation-Best matchings

In the body of the text, I proposed another equilibrium concept for matching markets involving LPUV-rational agents, based on optimality rather than stability:

Definition 6.1. μ is **optimal** for side M under the family of valuation $(v_m)_{m \in M}$ if:
 $\forall m \in M, \mu(m) \in \arg \max_{f \in F \cup \emptyset} v_m(f)$

Definition 6.2.

μ is a **LPUV-best** matching for side M if there exists a family of valuations $(v_m)_{m \in M}$ s.t.:

- (i) μ is optimal for side M under $(v_m)_{m \in M}$.
- (ii) $\forall m \in M, v_m$ is μ -LPU.

Following Jehiel and Samet (2007) [17], let me remind the standard definition of Valuation Equilibrium. I use the following notations for extensive form games:

- $I = \{1, \dots, i\}$: set of players
- Z : set of terminal nodes
- N : set of non-terminal nodes

- A : set of arcs (n, m)
- N_i : set of nodes where i must play
- $M_i(n) = \{m \mid (n, m) \in A\}$: possible moves of i at $n \in N_i$
- M_i : Set of all possible moves
- $(\sigma_i(n) \in \Delta M_i(n))_{n \in N_i}$: behavioral strategy of player i
- $f_i : Z \rightarrow \mathbb{R}$: i 's payoff at terminal nodes
- Λ_i : partition of M_i , henceforth player i 's similarity partition
- $\lambda(m)$: similarity class to which node m belongs
- $Z(\lambda)$: set of terminal nodes that are descendants of some node in λ
- $P^\sigma(z)$: probability that $z \in Z$ is reached if σ is played.

The concept is subsumed in the four coming definitions:

Definition 6.3. A **valuation** for player i is any function: $v_i : \begin{cases} \Lambda_i \rightarrow \mathbb{R} \\ \lambda \mapsto v_i(\lambda) \end{cases}$

Definition 6.4. σ_i is **optimal** for v_i if $\forall n \in N_i, m \in M_i(n)$:

$$\text{supp}(\sigma_i(n)) \subseteq \arg \max_{m \in M_i(n)} v_i(\lambda(m))$$

Definition 6.5. v_i is **consistent** with σ if $\forall \lambda \in \Lambda_i$ s.t. $P^\sigma(Z(\lambda)) > 0$:

$$v_i(\lambda) = \frac{1}{P^\sigma(Z(\lambda))} \cdot \sum_{z \in Z(\lambda)} P^\sigma(z) f_i(z)$$

Definition 6.6. σ is a **Valuation Equilibrium** if $\exists (v_i)_{i \in I}$ s.t. $\forall i \in I$:

- (i) σ_i is optimal for v_i .
- (ii) v_i is consistent with σ .

I associate a strategy profile σ to a matching μ . In a LPUV-matching, everything happens as if the agents would construct similarity classes for inference on the match surplus (only). Abusing notations, I write that each player $m' \in M$ partition the set of possible pairs $M \times F$ into $2 \times \#F$ similarity classes: $\lambda(f) = \bigcup_{m \in M - \{m'\}} (m, f)$, $\lambda'(f) = \{(m', f)\}$. However, several points

prohibit interpretation of LPUV-best matchings as extension of the Valuation Equilibrium to matching environments.

First, the general form of Valuation Equilibrium authorises flexible similarity partitions (although the partition remains exogenous): multiple similarity partitions can be used in the same game. This flexibility allows to derive equivalence with the more standard equilibrium concepts. Indeed, under the coarsest similarity partition (only one similarity class), any strategy profile is a Valuation Equilibrium, whereas under the finest similarity partition (when any similarity class is a singleton), Valuation Equilibria and Subgame Perfect Nash Equilibria exactly coincides. By contrast, the valuation partition in LPUV-best matchings is fixed, (exogenous) and close to symmetric (identical similarity partitions across different agents on a given side of the market, except for valuation of the matching with the current partner). Due to this lack of flexibility, no equivalence result between GS-stability and LPUV-stability can be formulated as depending on how fine is the similarity partition. Second, (and whatever the similarity partition), existence of Valuation Equilibrium is guaranteed. Existence of LPUV-best matchings is also true but it is somehow a voided result, as it relies on LPUV-optimality of the null matching, thus indirectly on the unconstrained specification of the valuation for unmatched agents. This may prove clearer with the coming formal statements and associated proof:

Theorem 6.3. μ_0 is LPUV-best for both sides.

Proof. 6.1. Set an agent $m \in M$. $\mu_0(F) = \emptyset$. So any Valuation $v : F \cup \{\emptyset\} \rightarrow \mathbb{R}_+$ s.t. $v_m(\emptyset) = u_{m\emptyset}$ is μ_0 -LPU. Let us also set $\forall f \in F : v_m(f) = 0$. Then, $\forall f \in F : v_m(\emptyset) = u_{m\emptyset} \geq v_m(f) = 0$, so m prefers to remain unmatched, and optimality is satisfied. So μ_0 is LPUV-best. \square

This directly implies existence:

Corollary 6.4. *For a given side of a matching market, there always exists a LPUV-best matching.*

I now provide a characterisation of LPUV-best matchings that is slightly more convenient than the original definition:

Proposition 6.5. μ is a LPUV-best matching for side M if and only if there exists a family of valuation $(v_m)_{m \in M}$ s.t. $\forall m \in M$:

$$(i) \quad \forall m \in M : \mu(m) \in \arg \max_{f \in F | \mu(f) \neq \emptyset} v_m(f)$$

(ii) v_m is μ -LPU.

Proof. 6.3.

- (\Rightarrow) By definition of a LPUV-best matching.
- (\Leftarrow) Suppose there exists a family of valuations $(v_m)_{m \in M}$ such that $\forall m \in M, \mu(m) \in \arg \max_{f \in F | \mu(f) \neq \emptyset} v_m(f)$ and v_m is μ -LPU. Define $\forall m \in M, w_m$ s.t.:

$$w_m(f) = \begin{cases} v_m(f) & \text{if } \mu(f) \neq \emptyset \\ 0 & \text{if } \mu(f) = \emptyset \end{cases}$$

By construction:

$$\begin{aligned} \arg \max_{f \in F | \mu(f) \neq \emptyset} v_m(f) &= \arg \max_{f \in F} w_m(f) \Rightarrow \mu(m) \in \arg \max_{f \in F} w_m(f) \\ v_m \mu\text{-LPU} &\Rightarrow w_m \mu\text{-LPU} \end{aligned}$$

So μ is LPUV-best for side M . \square

I give a more precise characterisation for the specific case when the splitting is constant for one side of the market:

Proposition 6.6.

Suppose the splitting is constant for the man side: $\forall m \in M, \forall f \in F : r_{mf} = r_m$. Then:

- μ is LPUV-best for side $M \Rightarrow \exists c \in \mathbb{R}$ s.t. $\forall m \in M$ s.t. $\mu(m) \neq \emptyset, s_{m\mu(m)} = c$.
- Set $(m, f) \neq (m', f')$ s.t. $s_{mf} \neq s_{m'f'}$. Then there is no LPUV-best matching in which (m, f) and (m', f') are matched.
- Set a matching market s.t. $\forall (m, f), (m', f'), i \neq i', j \neq j', u_{ij} \neq u_{i'j'}$. Then in any LPUV-best matching there is at most one matched pair.

The intuition is direct. Because the splitting is constant for the man side, the male ranking of women only depends on the estimate of the surpluses. Under LPUV-rationality, this estimate is the same for all agents. So the (ordinal) preferences of all men are the same. At most one matched man can satisfy optimality.

The concept thus seems very strong. On a non-trivial matching market (where match surpluses are not equal) and under one side constant splitting, only the null or single-pair matching can be LPUV-best.

Proof. 6.4.

- (i) Set μ a LPUV-best matching (under $(v_m)_{m \in M}$).
 Set m, m', f, f' s.t. $\mu(f) = \{m\}, \mu(f') = \{m'\}$.
 By optimality:

$$\begin{cases} v_m(f) \geq v_m(f') \\ v_{m'}(f') \geq v_{m'}(f) \end{cases} \iff \begin{cases} r_m s_{mf} \geq r_m s_{m'f'} \\ r_{m'} s_{m'f'} \geq r_{m'} s_{mf} \end{cases} \iff s_{mf} = s_{m'f'}$$

(ii) By contraposition of (i).

- (iii) Set a matching market s.t. $\forall (m, f), (m', f'), i \neq i', j \neq j', s_{ij} \neq s_{i'j'}$. By contraposition of (ii), there is at most one matched pair. \square

I finally formalise the relation between LPUV-best and LPUV-stable matchings:

Proposition 6.7. μ is LPUV-best for side $M \Rightarrow \mu$ is LPUV-stable.

Proof. 6.5.

Set μ LPUV-best matching for side M , under the family of valuation $(v_m)_{m \in M}$.

By optimality: $\forall m \in M, v_m(\mu(m)) \geq v_m(\emptyset)$ $(IR)_m$. Set: $\forall f \in F : v_f(\emptyset) = 0$. Then, $\forall f \in F, v_f(\mu(f)) \geq v_f(\emptyset)$ $(IR)_f$. By optimality again, $\forall m \in M, f \in F, v_m(\mu(m)) \leq v_m(f)$ $(IC)_{(m,f)}$. No man engage in a blocking pair. So μ is LPUV-stable. \square

What stands out from this analysis is that the concept of LPUV-best matching is too strong to be really interesting. This ex post validates the decision to focus on the weaker LPUV-stability concept.

6.3.2 Tree representation

In order to make the full connection with Valuation Equilibrium, I propose a tree representation of a matching market. In graph theory, a tree is any connected acyclic graph. I recall the associated game theory vocabulary. For a tree Γ :

- Nodes $V \begin{cases} r_0 \text{ root} \\ Z \text{ terminal nodes} \\ D = V - Z \text{ non terminal nodes} \end{cases}$
- Predecessor function: $\phi : V - \{r_0\} \rightarrow D$
- Terminal payoff: $f : Z \rightarrow D$
- Moves: $\{(\phi(n), n), \forall n \in V\}$

Besides, I define a two-sided (one-to-one (OtO) or many-to-one (MtO)) matching market (MM) by:

- 2 sides I, J .
- Match utilities: $u_{kl}, \forall k \in I \cup \{\emptyset\}, l \in J \cup \{\emptyset\}$.

I suggest a formal definition for the tree representation of a matching market immediately below:

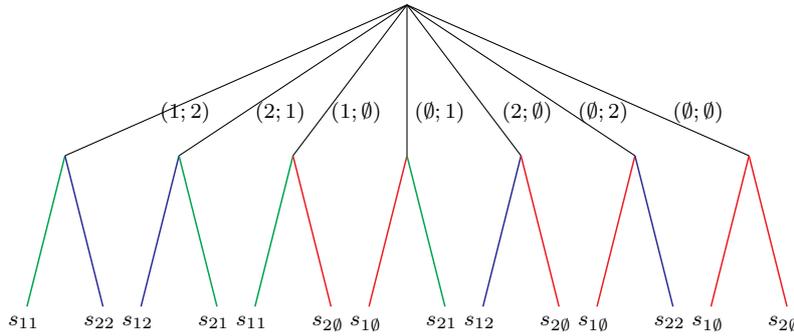
Definition 6.7. Tree Γ **represents** matching market (MM) from the point of view of side I if:

- (i) $\phi(Z) = D - \{r_0\}$
- (ii) $\forall n \in D - \{r_0\} \# \phi^{-1}(n) = \#I$ and $\# \phi^{-1}(r_0) = \begin{cases} (\#J + 1) \#J & \text{OtO} \\ (\#J + 1) \#I & \text{MtO} \end{cases}$

$$(iii) \begin{cases} \forall x \in (J \cup \{\emptyset\})^{\#I} \text{ s.t. } x_i = x_j \iff x_i = \emptyset & \text{OtO} \\ \forall x \in (J \cup \{\emptyset\})^{\#I} & \text{MtO} \\ \exists n \in D \text{ s.t. } \forall i \in I, \exists m \in \phi^{-1}(n) \text{ s.t. } f(m) = u_{ix_i} \end{cases}$$

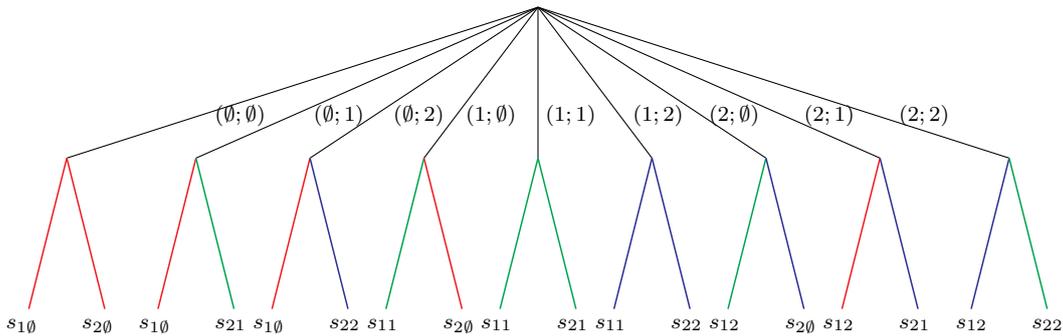
I now giving examples for small-size markets.

Suppose I locate on a two-sided, one-to-one matching market, with two individuals on each of the two sides M and F . I order the individuals on each side: $M = \{m_1, m_2\}$, $F = \{f_1, f_2\}$. I denote $\mu \equiv (i; j) \in \{1, 2\}^2$ the matching μ such that: $\mu(m_1) = f_i, \mu(m_2) = f_j$. I replace i (j) by \emptyset if m_1 (m_2) is unmatched. There are $3 \times (3 - 1) + 1 = 7$ possible matchings. The tree representation of the described market from the point of view of side M is drawn below. I represent in the same color the vertexes corresponding to match surpluses that a LPUV-rational agent bundles together:¹⁸



Looking at the family of green vertexes, for instance, one observes that a LPUV-rational agent on side M would consider equally s_{11} under $(1; 2)$, s_{21} under $(2; 1)$, s_{11} under $(1; \emptyset)$, s_{21} under $(\emptyset; 1)$ for social learning. The agent is right to bundle s_{11} under $(1; 2)$, s_{11} under $(1; \emptyset)$ together, or s_{21} under $(2; 1)$ and s_{21} under $(\emptyset; 1)$, because of the independence of surpluses from the other realised matches. The coarse partitioning he makes consists in bundling s_{11} and s_{21} in any state.

I now locate on a many-to-one matching market, with two agents on the many side of the market F and two agents of the one side of the market E . There are $3^2 = 9$ possible different matchings. The tree representation of this market from the point of view of the one side E is drawn below, with unchanged color code:



6.4 Tool box

6.4.1 Matching belief matrices

I would like to represent agents belief on match utilities in a synthetic easily readable way. I give formal definitions before proceeding to an example.

¹⁸The red color is an exception. Under the chosen information structure, agents perfectly know the reservation utilities, so do not bundle across different $s_{m\emptyset}, m \in M$.

Set a two-sided one-to-one matching market $(M, F; \psi; n)$, where $\#M = \#F = n$ and the match utilities are Beckerian: $\forall (m, f) \in M \times F : \begin{cases} u_{mf}^m = r_{mf} \cdot s_{mf} \\ u_{mf}^f = (1 - r_{mf}) \cdot s_{mf} \end{cases}$.

Definition 6.8. Set a matching on market $(M, F; \psi; n)$, where $\forall (m, f) \in M \times F : u_{m\emptyset}^m = u_{\emptyset f}^f = 0$. I define $\mu_l = [(m_k, f_k)_{k \in \{1, \dots, n\}}]$, $l \in \{1, \dots, n!\}$ as one full matching on $(M, F; \psi; n)$. A matrix $A_l^{\psi, \mathcal{R}} \in \mathcal{M}_{n \times n}((\mathbb{R}^2)^2)$ represents beliefs in matching μ_l under info structure ψ and cognitive rationality type \mathcal{R} if: $\forall (i, j) \in \{1, \dots, n\}^2 : a_{ij} = (u_{ij}^i, u_{ij}^j)$. For instance, if \mathcal{R} is:

- Full information: $a_{ij} = (u_{ij}^i, u_{ij}^j)$ and $\forall \mu_l, A_l$ is the same.
- Limited information ψ and LPUV-rationality: $a_{ij} = (v_i(j), v_j(i)) = (r_{ij} \cdot s_{\mu(j)j}, (1 - r_{ij}) \cdot s_{i\mu(i)})$

Proposition 6.8. A matching $\mu_l = [(m_k, f_k)_{k \in \{1, \dots, n\}}]$ on market $(M, F; \psi; n)$ is stable under information structure ψ and cognitive rationality type \mathcal{R} if and only if:

$$\forall (k, k') \in \{1, \dots, n\}^2 : \begin{cases} a_{kk}^1 \geq a_{k'k}^1 \\ \text{or } a_{k'k'}^2 \geq a_{kk'}^2 \end{cases}$$

I process example a. in subsection §3.1.2 thanks to this matrix representation. The figures were: $(\alpha_1, \alpha_2, \beta_1, \beta_2) = (7, 1, 1, 1)$, $(s_{11}, s_{22}, s_{12}, s_{21}) = (5, 1, 3, 2)$. I consider $\mu = [(m_1, f_1); (m_2, f_2)]$. Beliefs in matching μ under full information can be represented by the following matrix:

		Side F	
		f_1	f_2
Side M	m_1	$\left(\frac{35}{8}, \frac{5}{8}\right)$	$\left(\frac{21}{8}, \frac{3}{8}\right)$
	m_2	$(1, 1)$	$\left(\frac{1}{2}, \frac{1}{2}\right)$

(m_2, f_1) is a GS-blocking pair, and μ is GS-unstable.

Beliefs in matching μ under limited information ψ and LPUV-rationality can be represented by the following matrix:

		Side F	
		f_1	f_2
Side M	m_1	$\left(\frac{35}{8}, \frac{5}{8}\right)$	$\left(\frac{7}{8}, \frac{5}{8}\right)$
	m_2	$\left(\frac{5}{2}, \frac{1}{2}\right)$	$\left(\frac{1}{2}, \frac{1}{2}\right)$

There is no LPUV-blocking pair and μ is LPUV-stable.

6.4.2 Aggregate comparison of stabilities

The idea in this subsection is to design a systematic tool to assess which of the two stability notions (GS-stability or LPUV-stability) will (1) more often consider a matching (eg.: null matching, full positively assortative matching) as stable (2) overall produce more stability (3) on which types of markets (with assortative, aligned, idiosyncratic preferences).

The general methodology consists in modelling match utilities as random variables, following a distribution close to empirical distributions of preferences on matching markets. I develop an

example below, for idiosyncratic, preferences.

Set a 2×2 matching market. Suppose the match utilities are uniformly distributed:

$$\begin{aligned} \forall (i, j) \in \{1, \dots, n\} : s_{ij} \text{ i.i.d. } &\sim \mathcal{U}([0, \bar{u}]) \\ r_{ij} \text{ i.d.d. } &\sim \mathcal{U}([0, 1]) \\ u_{i\emptyset}^i, u_{\emptyset j}^j \text{ i.d.d. } &\sim \mathcal{U}([0, \bar{f}]) \end{aligned}$$

To ensure $u_{i\emptyset}^i$ and $u_{i j}^i$ ($u_{\emptyset j}^j$ and $u_{i j}^j$) have the same supports and distributions, I pose: $\bar{f} = 2\bar{u}$, for instance $\bar{u} = 1$, $\bar{f} = 2$.

There are 7 possible matches, using notations defined above in subsection §6.3.2:

- None matched pair (null matching): (\emptyset, \emptyset) .
- One matched pair: $(1, \emptyset)$, $(2, \emptyset)$, $(\emptyset, 1)$.
- Two matched pairs (full matchings): $(1, 2)$, $(2, 1)$.

Because the match utilities are i.i.d. distributed, the probabilities to be stable for two matching belonging to the same categories are equal. I compute each of the three probabilities. For instance, the probability for the null matching to be LPUV-stable is one (by definition), and the probability that it is GS-stable writes:

$$\begin{aligned} \mathbb{P}[\mu_0 \text{ GS-stable}] = \mathbb{P}[&[(u_{1\emptyset}^1 \geq r_{11} \cdot u_{11}) \cup (u_{\emptyset 1}^1 \geq (1 - r_{11}) \cdot u_{11})] \\ &\cap [(u_{1\emptyset}^1 \geq r_{12} \cdot u_{12}) \cup (u_{\emptyset 2}^2 \geq (1 - r_{12}) \cdot u_{12})] \\ &\cap [(u_{2\emptyset}^2 \geq r_{21} \cdot u_{21}) \cup (u_{\emptyset 1}^1 \geq (1 - r_{21}) \cdot u_{21})] \\ &\cap [(u_{2\emptyset}^2 \geq r_{22} \cdot u_{22}) \cup (u_{\emptyset 2}^2 \geq (1 - r_{22}) \cdot u_{22})]] \end{aligned}$$

Analytical computation is virtually impossible due to intrinsically complex dependence structure. The computation software such as *Wolfram Mathematica* could not deliver in reasonable time either. I have therefore performed simulations on the statistic software R, using 500,000 Monte Carlo draws of match utilities. I thus recover very precise estimates of the probabilities of interest. They are reported in the following table:

	GS-stability	LPUV-stability
$\mathbb{P}[\mu_0 \text{ stable}]$	0.444	1
$\mathbb{P}[(1, \emptyset) \text{ stable}]$	0.123	0.199
$\mathbb{P}[(1, 2) \text{ stable}]$	0.033	0.031

I do not claim I can answer to questions (1), (2) and (3) based on those figures. For instance, the fact that full matchings are more often GS-stable than LPUV-stable needs to be confirmed with different distributions. Yet, it is clear that similar numerical results for a wide array of preferences could provide guidance in the future.