# Losing Face

Thomas Gall[a] and David Reinstein[b]

a Department of Economics, University of Southampton
b. Department of Economics, University of Exeter, Streatham Ct. EX4 4PU, daaronr@gmail.com

**Abstract**

When Al makes an offer to Betty that Betty observes and rejects, Al may suffer a painful and costly 'loss of face' (LoF). LoF can be avoided by letting the vulnerable side move second, or by setting up 'Conditionally Anonymous Environments' that only reveal when both parties say yes. This can impact bilateral matching problems, e.g., marriage markets, research partnering, and international negotiations. We model this assuming asymmetric information, continuous signals of individuals' binary types, linear marriage production functions, and a primitive LoF term component to utility. LoF makes rejecting one's match strictly preferred to being rejected, making the 'high types always reject' equilibrium stable. LoF may have non-monotonic effects on stable interior equilibria. A small LoF makes high types more selective, making marriage less common and more assortative. A greater LoF (for males only) makes low-type-males reverse snobs, which makes high-females less choosy, with ambiguous effects on the marriage rate.

JEL codes: D83, D03, C78

## 1 Introduction

In a market that involves two-sided matching (as surveyed in Burdett and Coles, 1999), the fear of rejection can lead to inefficiency. A proposer may not ask someone out on a date, ask for a study partner, apply for a job, make a business proposition, propose a paper co-authorship, or suggest a peace treaty, because she does not want the other party to *know* of her interest and then turn her down. This may have consequences for reputation and future play, or it may have a direct psychological cost. In general, we call the disutility from this outcome 'loss of face' (LoF). Consider a game where each player can choose 'accept' or 'reject' and there is asymmetric information about players' types. Assume that the outcome of the game (actions and payoffs) becomes common knowledge after all actions have been taken. Here LoF may worsen the set of Nash equilibria. There may be a set of mutually beneficial transactions that would occur without LoF, but do not occur with LoF because:

1. the proposer does not know for sure whether the other party will accept or reject *and*

2. a high enough probability of rejection can outweigh the expected gains to a successful transaction.

This goes *beyond* the standard problems of asymmetric information. Even where Al perceives that his *expected* utility from actually *marrying* Betty would be positive, his expected utility from making an *offer* may be negative. Thus he may still reject Betty — to avoid LoF — if he anticipates a high enough chance she will reject him. It is also distinct from the 'self-image preservation' motive, discussed in Köszegi (2006), that may lead to over- or under-confident task choice. In that model Al ‘hates to *learn* that Betty deems him to be low quality’, as it reduces his self image. In contrast, our LoF comes from ‘conditional on her rejecting me, I dislike her knowing that (i) I perceive her as being good enough for me and (ii) that I have made an offer to her.’

We believe this will be intuitive for most readers. Consider: which scenario below would be more painful? (Suppose you are romantically interested in women.)

1. A friend or colleague, in whom you have an unexpressed romantic interest, while discussing her tastes, informs you that she wouldn’t date you because you are not ‘her type.’ You have no reason to believe that she knows of your interest in her, and you are certain that she is telling the truth.

2. Without having the conversation in scenario 1, you ask this same person out on a date and she refuses because you are not ‘her type.’

We speculate that the second scenario would be more painful: now both you and she know that you have asked her out and she has refused. Although she may have tried to soften the blow by posing this as a matter of idiosyncratic preference rather than quality, you have lost face, and you are established as her inferior in one sense. In the first case, although you can presume she is not interested in you, and this may hurt your self esteem, she doesn’t know you like her, and you have not lost face, as we define it.

We mainly take this as a primitive (but also present a reputation model in appendix B; *all* appendices are online ‘supplementary material’); future work could unpack this in a more extensive model. E.g., her knowing I chose ‘accept’:

1. informs her about my quality, affecting my reputation, which I may care about directly, or through its impact on my payoffs in future interactions (see online appendix B);

2. may be undesirable through a reciprocity motive (Falk et al., 2006): I want to harm someone who harms me, and playing ‘reject’ may be seen as harmful and ‘accept’ beneficial.

Simple institutional changes can eliminate this risk: if only mutual ‘accept’ choices are revealed, the rejected party’s choice is thus hidden from the rejector. Al will never have to worry that Betty will both reject him and learn that he accepted her. (In contrast, whenever Al accepts Betty *he* will always learn *her choice*; his self-image cannot be easily protected.)

As we discuss in section 2, there is evidence that a desire not to lose face is a *primal* human concern, perhaps a product of evolutionary factors, or perhaps an automatic internalization of a reputation motive. (If reputation concerns are long-term, anticipating the additional short-run pain of losing face may help counteract present-bias as well as overconfidence.) Thus the LoF may enter into an individual’s utility function *directly.* There is a special loss from the combined knowledge that you accepted somebody, but they rejected you.[1]

When (e.g.) a woman accepts a man and he rejects her, her material payoffs from this one-shot game are the same no matter what beliefs or information either party has. However, with LoF, her psychic payoffs are lower when *she knows that he knows that* she accepted him and he rejected her. In other words, what the other player knows for sure – the other player’s information – is a component of a player’s utility function (as in Battigalli and Dufwenberg, 2007). Thus, as long as we know the (terminal) information structure, LoF transforms material payoffs into psychological payoffs in a

[1] This assumption puts our model into the category of a *psychological game*, as modelled by Battigalli and Dufwenberg (2007), in which my payoffs may depend on another player’s *beliefs* about my action. However, in our model, for a given (exogenous) information structure, the relevant beliefs (are a one-to-one mapping from the players’ actions; thus our analysis is standard. As described in section 4), we assume the structure of these ‘terminal information sets’ is common knowledge; we use this terminology to avoid confusion with the standard setup in which information sets are only defined in connection with decision nodes. The simplicity of our game means we do not have to worry about, e.g., actions responding to equilibrium beliefs responding to actions.

straightforward way.[2]

We focus on the primal LoF interpretation: this is particularly relevant to one-shot games where no outside parties observe the results. However, we suspect that many of our results will carry over to a case where the LoF concern can be justified instrumentally. With asymmetric information, as in our model, in many types of dynamic matching and sorting/screening games, an individual's willingness to 'accept' another person be taken by others as a negative signal of her type, reducing her utility and/or her continuation value. To reinterpret Groucho Marx 'if I am willing to be part of this club, how good can *I* be?' We give a simple formalization of this in a two-period model in online appendix B, where we derive conditions under which a players' previous 'accept' choice hurts her continuation value. (However, a complete characterization of equilibria for this model is left for future work).

As noted above, if loss of face depends on the terminal information sets in this way, i.e., on the information each player has at the end of the game over the game's history, then it can be avoided by changing the information structure so that players *only learn about each other's behaviour if they both play Accept.* For example, speed dating agencies often ask men and women to mark the partners who they are interested in, and then inform only those couples who both marked each other. Now, after playing accept, *you* will still be able to infer if you have been rejected, but the *other person* will not know that you accepted them; knowing this, you will not suffer a loss of face. Thus, while your ego-utility can not be preserved, your *face* can be. We call such setups *Conditionally Anonymous Environments* or CAE's.

Our paper proceeds as follows. In section 2 we discuss our concept in more detail and offer intuitive, anecdotal, and academic support for it, motivating the *assumptions* of our model. We also give a short survey of the related economic literature. In section 3 we describe our baseline setup (similar to a single stage of Chade, 2006), and formally define LoF. This environment yields only monotonic equilibria, following the theory of games with strategic complementarities (summarized in Vives, 2005). In section 5 we characterize the best response strategies and equilibria, considering both a symmetric case and a case where only males suffer LoF. The latter allows us to consider both direct and indirect effects. We demonstrate that LoF can make a coordination failure equilibrium tatonnement-stable, and present monotone comparative statics as LoF is introduced or increased (applying Milgrom and Shannon, 1994). We show that while a small amount of LoF makes the low types 'reverse snobs' and generally reduces the efficiency of the marriage market, a greater LoF may actually *increase* the marriage rate. We conclude in section 6, considering extensions and discussing policy implications. Our appendices (all online) provide longer proofs, details, and numerical examples, a comparison to a 'rejection hurts' model, and our model of reputation concerns in a two-stage game.

## 2 Background

There is abundant psychological evidence that 'rejection hurts' (Eisenberger and Lieberman, 2004) and that social ostracism can cause a neurochemical effect that resembles physical pain (Williams, 2007). However, these studies do not distinguish between cases where it is common knowledge that the rejected party has expressed an interest from cases where this is private information. We claim that people fear proposing, and they fear it more when proposals are known.

[2]Furthermore, LoF itself has no obvious interpretation in terms of fairness/reciprocity (Rabin, 1993). Since revelation of 'who proposed to whom' or 'who was kind to whom' occurs after these decisions were made it should have no impact on beliefs about whether a player knew his play was 'fair' in the sense of being congruent with the other player's kindness or unkindness.

Our speculation in the introduction is consistent with a plausible interpretation of much previous work. While some of the examples below admit alternative explanations (e.g., self-image preservation), we believe that the overall picture offers support for our model's assumptions over LoF. Bredow et al. (2008) represent previous research through the formula $V = f(A \times P)$ for the 'strength of the valence of making an overture' to a romantic partner, where $A$ represents attraction and $P$ is the estimated probability that an overture will be accepted. Shanteau and Nagy's (1979) experimental work finds that 'when the probability of acceptance is low, people's interest in pursuing a relationship is nil, or nearly nil, regardless of how attracted they are to the person.' One reason for these attitudes and preferences might be the fear having one's overtures known in the event of being rejected. Such a cost may be intrinsic or reputation-driven, psychological or material.[3]

The fear of losing face or reputation may motivate people to put in effort and incur costs in order to learn whether a potential partner is likely to respond positively. Baxter and Wilmot (1984) described six types of secret tests used in the delicate dance of 'becoming more then friends', e.g., 'third-party tests' (Hitsch et al., 2010). Douglas (1987) 'reports eight strategies that individuals reported using to gain affinity-related information from opposite sex others in initial interactions.'

The fear of LoF is closely related to what psychologists call 'rejection sensitivity.' For example, London et al (2007) provide evidence from a longitudinal study of middle school students that, for boys, 'peer rejection at Time one predicted an increase in anxious and angry expectations of rejection at Time 2.' They also find that anxious and angry expectations of rejection are positively correlated to later social anxiety, social withdrawal, and loneliness. In explaining the connection to loneliness, they posit that the rejection sensitive may exhibit 'behavioural overreactions' such as 'flight' (social anxiety/withdrawal) or 'fight' (aggression).' It is easy to interpret either of these as a way to choose 'reject' in our matching game in order to avoid further loss of face.

Erving Goffman (2005) has written extensively about losing and preserving *face:*

> The term *face* may be defined as the positive social value a person effectively claims for himself by the line others assume he has taken during a particular contact ...The surest way for a person to prevent threats to his face is to avoid contact in which these threats are likely to occur. In all societies one can observe this in the avoidance relationship and in the tendency for certain delicate transactions to be conducted by go-betweens ...

In the context of our paper, Goffman's 'avoidance' is essentially pre-emptive rejection: you cannot be matched with a partner if you don't show up.

In the USA over a recent ten-year period, 17% of heterosexual and 41% of same-sex couples met online (Rosenfeld and Thomas, 2012), and the dating industry has been reported to constitute 'a $2.1 billion business in the U.S., with online dating services ... representing 53% of the market's value' (MarketData Enterprises, 2012). Internet dating *itself* can be seen as an institution designed to minimize the LoF that comes with face-to-face transactions, allowing people to access a network of potential partners who they are not likely to run into again at the office or on the street. However, going online may not eliminate the LoF; as noted in Hitsch et al. (2010): 'If ... the psychological cost of being rejected is high, the man may not send an e-mail, thinking that the woman is 'beyond his reach,' even though he would ideally like to match with her.' (Here, this psychological cost could include both the loss of face we consider and self-esteem concerns outside our model.)

[3]For example, in the 2005 'Northwestern Speed-Dating Study' on 163 undergraduate students, 'participants who desired everyone were perceived as likely to say yes to a large percentage of their speed-dates, and this in turn negatively predicted their desirability' (Eastwick and Finkel, 2008).

Perhaps in response to this, numerous dating sites and applications have introduced some form of the CAE environment, where member A can express interest in member B and member B only finds out about this if B also expresses an interest in A.[4] However, there is a trade-off between preserving face and getting noticed: with thousands of members, each member may only view a fraction of eligible dates, and if A expresses anonymous interest there is no guarantee that B will even see A's profile. This has been applied to the internet context at least since Sudai and Blumberg (1999), who were granted a patent for such a 'computer system', noting 'often, even when two people want to initiate first steps in a relationship, neither person takes action because of shyness, fear of rejection, or other societal pressures or constraints.'[5]

Perhaps the most widely used dating platform is the smartphone app Tinder. The site claims it has lead to to 1.6 billion swipes per day, 1 million dates per week, and over 20 billion matches (2018, 2018). Here, users are presented a sequence of profiles (with pictures and bios), and can 'swipe right' to indicate their interest. Only mutual right-swipers are informed 'It's a match', are then able to chat directly. Swiping right on someone does *not* imply they will see your profile; the order in which Tinder's optimization algorithm presents profiles does not reveal who liked you. Thus, this app resembles our Conditionally Anonymous Environment. Sean Rad, a founder and CEO, noted his motivation for this 'double opt-in' systems as the app's impetus. 'No matter who you are, you feel more comfortable approaching somebody if you know they want you to approach them' (Witt, 2014).

'Speed dating' was an earlier innovation in the singles scene. These events usually attract an equal number of customers of each gender; men rotate from one woman to another, spending a few minutes in conversation with each. Here there is also an effort to minimize the possibility of public rejection (and perhaps LoF). In fact, speed dating agencies often promote themselves on these grounds, e.g., in 2012 Xpress dating advertised 'rejection free dating in a non-pressurized environment'.[6]Typically, participants are asked to select whom they would like to go on 'real dates' with only *after* the event is over. In most cases the agency will only reveal these 'proposals' where there is a mutual match, i.e., where both participants have selected each other. 'Speed dating' institutions have been extended outside the realm of romance and marriage, into forming study groups, 'speed networking' and business partnering; these may have been established (in part) to minimize LoF (Collins and Goyder, 2013; CNN International, 2005).

LoF may not be limited to the dating world. Both psychological LoF and material losses from publicly observed 'ances and rejections can be seen in many spheres. These concerns may be present on both sides of the job market. A job-seeker may lose face when she makes a special appeal and is rejected, and an employee may lose face when rejected for a promotion or a special firm project. Akerlof and Kranton's (2000) model of social exclusion is also relevant. If being seen 'acting white' involves sacrificing Black identity, a Black person may choose not to attempt 'admission to the dominant culture' because she is uncertain about the 'level of social exclusion' she will face; e.g., whether she will be accepted by a school, employer, or White social group. On the other hand, if she can attempt this anonymously, she can avoid the risk of a public threat to her identity, and also avoid the potential *material* costs of social exclusion. In fact, race-based rejection sensitivity has been found

[4]We recognize that other models, e.g., derived from the 'rejection hurts' idea stated above, might justify such policies. However, we argue in appendix C that the LoF model is the most plausible justification.

[5]Online dating has been portrayed as a modern analogue to the traditional 'matchmaker', who was able to separately interview prospective mates and their families about their likes and preferences, helping arrange marriages while preserving anonymity (see Ariely and Jones, 2010, chapter 8, forgiving the misuse of the term *Yenta*). However, the internet and social media cuts both ways. Although the internet affords the opportunity to make connections outside one's usual network, the 'gossip network' may grow, increasing reputation concerns.

[6]Accessed 10 Dec 2018 via `https://web.archive.org`, search term 'http://www.xpressdating.co.uk:80/speed_intro.htm'.

to negatively correlate with measures of African-American students' success at predominantly White universities (Mendoza-Denton et al., 2002). This concern might help justify outreach programs for under-represented minorities; in effect 'asking them first' or letting them know when they will have a high probability of succeeding.

The employer too may be vulnerable to LoF. Cawley (2012), in his guide for economists on the junior job market, writes that he has 'heard faculty darkly muttering about job candidates from years ago who led them on for a month before turning them down.' This aggravation may involve LoF in addition to the loss of time and opportunity costs. This LoF is recognized by professional recruiters as well: 'recruiters lose face when candidates pull out of accepted engagements at the last minute' (Direct Search Allowance, 2007). Concerns on both sides of the job market may have inspired companies like Switch ('the Tinder for job apps') to develop conditionally anonymous (double opt-in) employment platforms.[7]

For the rejection-sensitive, any economic transaction that involves an 'ask' may risk a LoF. This may explain the prevalence of posted prices, aversion to bargaining in certain countries, and the relative absence of neighbourhood cooperation, social interaction, consumption and task-sharing in many modern societies (Putnam, 2000). Rejection sensitivity is particularly disabling for sales personnel, who may suffer from 'call reluctance'.[8]

Our model may also be important in an archetypal situation where preserving face is valued – the resolution of personal and political disagreements. Neither side may want to make a peaceful overture unilaterally – this can be seen as evidence of admission of guilt or weakness, and may be psychologically painful in itself. Again, where a double-blind mechanism is available, it can resolve this dilemma; if not, our model offers insight into why negotiations often fail. Often, peace talks are made in secret, and only announced if a successful agreement has been reached. This contradicts one of Woodrow Wilson's famous '14 points': 'Open covenants openly arrived at', became a principle, according to Eban (1983). However, Eban claims that 'the hard truth is that the total denial of privacy even in the early stages ... has made international agreements harder to obtain than ever'. Armstrong (1993) analysed three key cases of international negotiations finding a high degree of secrecy and few participants. 'In these secret and private negotiations, assurances and commitments were provided, which were essential for the parties to negotiate 'in good faith' (Jönsson and Aggestam, 2008).

While economists have previously studied related concepts, to our knowledge none have considered the difference between 'mutually-observed acceptance and rejection' and 'rejection where only one side knows he was rejected' (and the other side does not know whether or not she was proposed to). Becker (1973) introduced a model of equilibrium matching in his 'Theory of Marriage.' He considers the surplus generated from marriage through a household production function, and allows the division of output between spouses to be divided ex-ante according to each party's outside option in an efficient 'marriage market.' Anderson and Smith (2010) brought reputation into this context, noting 'matches yield not only output but also information about types' (but note that *offers* are not observed in their model). Chade (2006) explored a search and matching environment where participants observe 'a noisy signal of the true type of any potential mate.' He noted 'as in the winner's curse in auction theory – information about a partner's type [is] contained in his or her acceptance decision.' However, in Chade's model there is only a single interaction between the same man and woman, and outside

[7] Switch has claimed more than 400,000 job applications and 2 million 'swipes' as of 2015 (Crook, 2015).

[8] 'Call reluctance, which strikes both individuals and teams, develops in many forms. Representatives may be 'gun shy' from an onslaught of rejection or actively avoid certain calling situations such as calling high-level decision makers or asking for the order. Call reluctance is the product of fear; fear of failure, fear of losing face, fear of rejection or fear of making a mistake. If the fear perpetuates, productivity suffers' (Geery, 1996).

parties do not observe the results; thus there is no scope for either party's actions to affect their future reputations, nor any direct cost of being rejected.

Simundza (2015) embeds a two-stage matching game in a marriage model. He finds an equilibrium where saying 'no' in a first round can increase the continuation value in the next round. Our focus and modelling choices differ in important ways. Simundza considers a binary signal, leading to a focus on stationary strategies. Simundza's model isolates the strategic value of reputation (which is endogenous), and thus has no comparable 'Loss of Face' parameter. Thus, unlike Simundza, we can consider the welfare and distributional implications of changes in the information environment (CAE to FRE to ARE), changes that are relevant to many real-world contexts, as we note. Our differences lead to distinct results. For example, Simundza's high-types have no incentive to 'play hard to get,' i.e. no incentive to reject when observing high signals in the first round. In contrast, our high types are affected by their *own* potential LoF, raising their own thresholds or even shutting down completely. Simundza can compare the co-existing 'Nonstrategic' and 'Socially Strategic' equilibria; the latter yield greater assortative matching (sorting); as he assumes productive complementarity, this implies 'mating is more efficient'. In contrast, we compare the marriage rate and assortativeness of stable interior equilibria (as well as stability of corner equilibria) as the *cost* of LoF increases.[9]

# 3 Model Setup

## 3.1 Agents

The economy is populated by a continuum of individuals on market sides $M$ and $F$ ('male and female genders') endowed with measure 1 each. An individual $m \in M$ or $f \in F$ is characterized by a binary type $x_g \in \{\ell, h\}$; the type—'low' or 'high'—is an agent's private information (and $g \in \{m, f\}$).[10] For brevity, we will sometimes refer to 'an $h$' or 'an $\ell$', depicting an individual's type, and to 'an $m$' or 'an $f$', reflecting an individual's gender, and to 'male $\ell$-types', or 'a low $f$', etc. We will also refer to a generic individual as 'she/her', except where this would cause confusion. Let the share of high types be the same on both sides of the market, and denote it by $p$. (This assumption, for notational simplicity, does not affect our results qualitatively.)

## 3.2 Matching

Each individual in $M$ is randomly matched to an individual in $F$; all matches are chosen by nature with equal probability. Individual $i$ obtains a noisy signal $s_j$ about the type $x_j$ of her match $j$, but does not observe $s_i$, the signal of her *own* type that $j$ received. After observing the signals, individuals accept ($A$) or reject ($R$) the match. We distinguish three informational settings (depicted in figure 1):

1. a *Full Revelation Environment* (FRE) where both observe each others' actions (proposals) after they have both been made,

[9] Simundza's model somewhat resembles our two-stage 'reputation' model (appendix B). However, Simundza considers the *same* individuals playing the accept/reject game twice, with no new signals, unlike in our motivating examples. Our reputation model assumes a new second-round match who observes a new signal; we consider the impact of this match *also* observing her match's *previous-round game play.*

[10] While our discussion in the above sections also encompasses one-sided matching, we exclusively model a two-sided market (labelled 'male' and 'female' with apologies for political incorrectness). This choice is relevant to many examples and also allows us to isolate direct and indirect effects of (a fear of) LoF on one side. Note that our prior mimeo (Hugh-Jones and Reinstein, 2010) derived related results with *continuous* types under specific functional restrictions.

2. an *Asymmetric Revelation Environment* (ARE) where females observe the action $A$ or $R$ taken by a male but males do not observe females' proposals (but can infer them ex-post in some contingencies), and

3. a *Conditionally Anonymous Environment* (CAE) where neither side directly observes the other side's action (proposal), but each player only observes whether or not *both parties* have played accept.

I.e., the FRE captures a setting where both males and females are informed of the action of their match and know that their match will be informed of their own action. By contrast, in the CAE males and females can infer their match's actions (proposals) if and only if they themselves play accept. In an ARE only one market side (here females) is informed about the action of their match; the (male) player on other side is informed only if the female accepts. Hence, a female will never be observed accepting a male who rejects her. The ARE is strategically equivalent to a sequential game where both are vulnerable to LoF, but the male moves first—a second-mover can always avoid losing face by always playing $R$ after observing $R$. We discuss this further below (page 11).

### 3.3 Signals

Individuals in a matched pair each obtain a signal $s \in [\underline{s}, \overline{s}]$ of the other agent's type. Signals are drawn independently and their distribution depends on the type of the sender: type $x$'s ($x \in \{\ell, h\}$) signal is distributed according to $F_x(s)$ with continuously differentiable density $f_x(s)$. The densities must be bounded, i.e., $f'_x(s) < \infty$ (by Weierstrass Theorem, this holds for continuously differentiable densities where the support is a compact interval). Suppose that the signal is informative in the sense that $f_\ell(s)$ and $f_h(s)$ satisfy the monotone likelihood ratio property (henceforth mlrp), i.e.,

**Assumption 1.** *$f_h(s)/f_\ell(s) > f_h(s')/f_\ell(s')$ for all $s > s'$ where defined.*

We assume that the signals are fully revealing at their limits, i.e., observing the best (worst) signal implies that the type is $h$ ($\ell$), i.e.,

**Assumption 2.** *$f_h(\underline{s}) = 0$, $f_\ell(\underline{s}) > 0$, $f_\ell(\overline{s}) = 0$, and $f_h(\overline{s}) > 0$.*

Assuming that the probability of a high (low) type converges to one (zero) is needed to ensure that the game has an interior equilibrium (i.e., signal thresholds for accepting a match will be interior).[11]

### 3.4 Payoffs

If both individuals in a matched pair accept they become 'married' and each individual's payoff depends positively on the *pizazz* (see Burdett and Coles, 2006) of their partner: $x \in \{l, h\}$. Low types have pizazz $\ell$ and high types have pizazz $h$, where $0 < \ell < h$.[12] Types (and thus pizazz) become fully observable during the marriage. Marriage payoffs for a match $(m, f)$ are given by $u_m(x_f) = x_f$ and $u_f(x_m) = x_m$. That is, payoffs are linear in the match's type, which implies that total surplus does not depend on the precise assignment of types, but only on the number of marriages formed.

[11] Otherwise, in the case of overlapping supports, i.e. for all $s \in [\underline{s}, \overline{s}]$ $f_h(s) > 0$ if and only if $f_\ell(s) > 0$, we could not rule out equilibria where high types do not respond to the signal and instead 'always accept' or 'always reject,' and these could be stable. However, even under overlapping supports our remaining results carry over for 'responsive' equilibria where high types have interior thresholds. Details are available by request.

[12] We use the same terms, $h$ and $\ell$, to represent both the type index and the pizazz of this type; this slight abuse of notation should not cause confusion.

Agents who remain solitary obtain a payoff of $u_j(x_j) = \delta x_j$ for $j \in \{m, f\}$ with $\delta < 1$. For intuition, suppose types represent productivity, production is shared by the married couple, and those who are more productive alone are also more productive in a marriage.[13] Therefore low types always prefer a marriage to remaining alone. To make the model non-trivial, we suppose $h$'s prefer to remain unmarried to marrying an $\ell$, i.e., high types have a good outside option:

**Assumption 3.** *$\delta h > \ell$.*

In summary, homogenous marriages benefit both partners and mixed marriages benefit $\ell$'s more than they hurt $h$'s, because $h + \ell > \delta(h + \ell)$.

### 3.5 Loss of Face

Loss of face as described in section 2 is an intrinsic psychological pain, which can only matter if a player's potentially embarrassing action is observed by the other player. Therefore we define loss of face as follows.

**Definition 1.** *A player $j$ who suffers from loss of face experiences a loss $L$ when*

1. *$j$ played accept. $j$ knows that his match, player $k$, played reject, and*
2. *$j$ knows that $k$ knows (for certain) that $j$ played accept.*

The 'j knows that' part of Point 2 of may be necessary for a *primal* LoF, but not for the reputational LoF we model in appendix B; a player's reputation and future payoffs may suffer whether or not she knows that her decision is observed.[14]

Since LoF results from the *common knowledge* (or at least the higher order beliefs described above) of one party accepting and the other rejecting, to model LoF we need to make payoffs depend not only on actions, but also on the information players hold at the end of the game. These *terminal information sets* for players $m$ and $f$ are defined as standard information sets, but they are not at a decision node: they characterize a player's knowledge about the complete history of the game after all actions have been taken.

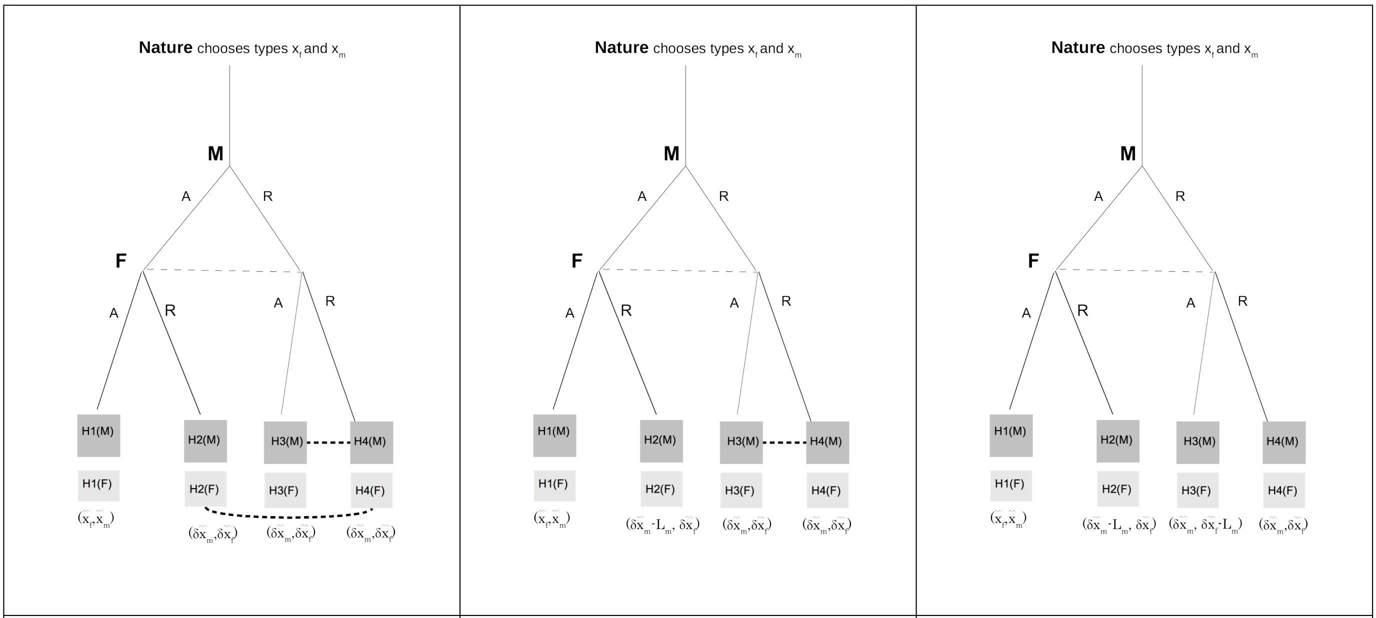
## 4 Terminal information sets and game trees

As shown in figure 1, the set of end nodes of the game, defined by their histories, is $H = \{H1, H2, H3, H4\} = \{AA, AR, RA, RR\}$.[15] Let $\bar{I}_f$ be the collection of $f$'s terminal information sets over these end nodes, and $\bar{I}_m$ be $m$'s information partition. Since neither player 'has the move' at the terminal node, we give each history two boxes to depict each player's terminal information set; $H_j(m)$ and $H_j(f)$ are the same (for $j \in \{1, 2, 3, 4\}$).

[13] An alternative justification: The payoff to no match may represent the continuation value in an indefinitely repeated matching game, as in, e.g., Adachi (2003), or as in our two-period model in appendix B.Simundza's 2015 model finds a similar related result, as does Chade (2006): higher-type players tend to have higher signals and others accept them more often.

[14] We conjecture that making LoF a continuous function of 'the probability $k$ puts on $j$ having played accept' would imply a secular decrease in payoffs for both sides in the CAE, but have no impact on the best-response functions derived below, implying qualitatively identical outcomes; informal proof available by request.

[15] We leave nature's move out of these histories; it does not affect our discussion. For completeness we can assume that players never learn the other players' types. Thus, in our model LoF will only depend on the conditional expectation of the other player's type, not the type itself.




| (i) CAE | (ii) m-Blind (ARE) | (iii) FRE |
|---|---|---|

Figure 1: Terminal Information structures

(i) Conditionally Anonymous (CAE): $\bar{I}_m = \{(AA), (AR), (RA, RR)\}$ and $\bar{I}_f = \{(AA), (AR, RR), (RA)\}$.
(ii) Asymmetric Revelation (ARE): $\bar{I}_m = \{(AA), (AR), (RA, RR)\}$; $\bar{I}_f = \{(AA), (AR), (RA), (RR)\}$.
(iii) Full revelation Environment (FRE): $\bar{I}_m = \bar{I}_f = \{(AA), (AR), (RA), (RR)\}$.

In the games defined above terminal information sets depend on the information environment in place. The three different environments are illustrated in the trees in figure 1, specifying the terminal information partitions for each case. The payoffs shown include LoF terms whenever the terminal information structure implies this may be relevant, by our definition above.

Denote an action tuple by $(a_m a_f) \in \{A, R\}^2$. If both players in a match only observe their own actions and whether or not there is a marriage (the conditionally anonymous environment, CAE), both players' information sets are $(AA)$, $\{(RA), (AR)\}$, or $(AR)$, depicted on the left of figure 1. Note that the $(AA)$ terminal information set is a singleton for both players, while the histories where a player played 'reject' are part of the same terminal information set (for that player). This implies that females cannot distinguish between action profiles $(AR)$ and $(RR)$, and males cannot distinguish between $(RA)$ and $(RR)$. Therefore there is no loss of face under the CAE.

In an asymmetric revelation environment (ARE) one market side observes the actions of the other side, but not vice versa; the other side only learns whether or not a marriage occurred. Here we will assume that females observe males' actions in a match, but not vice versa. I.e., consider 'males' as synonymous with 'the side vulnerable to LoF.' Hence, in the ARE possible terminal information sets are $(RA)$, $(RR)$, $(AA)$, or $(AR)$ for females, and $(AA)$, $(AR)$, and $\{(RA), (RR)\}$ for males. That is, females can distinguish the action profile $(AR)$ from $(RR)$, whereas males cannot distinguish between $(RA)$ and $(RR)$. Here, the males lose face when the action profile $(AR)$ is played, but females cannot lose face under the ARE.

In a full revelation environment (FRE) both genders have four terminal information sets, $(RA)$, $(RR)$, $(AA)$, or $(AR)$, as shown on the right in figure 1; i.e., both players in a match observe the choice of their partner, and know that their partner observes their own choice.

Note that while we depict a game with simultaneous actions (or equivalently, incomplete information – players don't observe each other's actions when they make their choices), in many applications

the game will be sequential, with one side having to make the first offer. If the terminal information sets are complete and the males must move first, this will be is strategically equivalent to the ARE above. The males—moving first— would be vulnerable to LoF. Females, moving second, would only consider playing accept if the male first-mover also did; thus they will never suffer LoF.
Therefore individual payoffs of the game played by a randomly matched pair $(m, f)$ can be summarized by the following payoff matrix:

| $m \backslash f$ | $A$ | $R$ |
|---|---|---|
| $A$ | $x_f, x_m$ | $\delta x_m - L_m, \delta x_f$ |
| $R$ | $\delta x_m,\ \delta x_f - L_f$ | $\delta x_m, \delta x_f$ |

. (1)

Setting $L_f = L_m = 0$ will correspond to the payoffs in the CAE, where no loss of face can occur by design. In an ARE with males moving first $L_f = 0$ and $L_m = L > 0$, and in a FRE with the same LoF on both sides, $L_m = L_f = L > 0$. We limit our attention to the case of symmetric loss of face and to symmetric equilibria of the game in the FRE.

While our setting does not allow for a generic analysis of matching games, it captures a large set of interactions in matching environments where loss of face may be relevant. Our assumptions embody agreed-upon preferences over a partner's type – partners are better or worse along a single dimension, although this may be a reduction of several characteristics.

Individuals' acceptance decisions will depend on the inference they make about their match's type given the signal and given the event of being accepted. We will look for Perfect Bayesian Equilibria and consider tatonnement stability, i.e., stability with respect to the iterative responses to deviations or 'cobweb dynamics' (see Hahn, 1962; Dixit, 1986; and Vives, 2005). In this setting tatonnement stability will require that, if one player slightly deviates from equilibrium play, the other player's best response, and the best response to this, ad-infinitum, will gradually move best responses back to the equilibrium play. (We will also mention when our results hold under the trembling-hand perfection refinement.)

# 5 Solving the Model

We note first that the game always has a trivial coordination failure equilibrium where both players always reject.[16] If $i$'s match rejects with certainty, then for $i$, rejecting yields payoff $\delta x_i$, which is at least as high as $i$'s payoff when accepting, and strictly greater when $i$ is vulnerable to loss of face.

## 5.1 Individual best reply functions

### 5.1.1 High types' best replies

For an individual $i$ of type $h$ and gender $g$ in a match $(i, j)$, playing $R$ yields a payoff $\delta h$, whereas playing $A$ either yields $x_j$ (if $j$ accepts) or $\delta h - L_g$ (if $j$ rejects). High types of both genders find it weakly profitable to accept after observing a signal $s$ if and only if the expected payoff from accepting meets or exceeds the outside option. I.e., for a high type of gender $g \in \{m, f\}$, letting $g' \in \{m, f\} \neq g$,

[16] This is distinct from a case where low types always accept and high types always reject, which we call the C-F Equilibrium; we return to this below.

$A$ is weakly preferred if

$$\underbrace{\frac{pf_h(s)}{(1-p)f_\ell(s)+pf_h(s)}}_{pr(x_j=h|s)}[\underbrace{q_{g'}(h,h)h}_{marry\ h}+\underbrace{(1-q_{g'}(h,h))(\delta h-L_g)}_{rejected\ by\ h}]$$
$$+\underbrace{\frac{(1-p)f_\ell(s)}{(1-p)f_\ell(s)+pf_h(s)}}_{pr(x_j=\ell|s)}[\underbrace{q_{g'}(\ell,h)\ell}_{marry\ \ell}+\underbrace{(1-q_{g'}(\ell,h))(\delta h-L_g)}_{rejected\ by\ \ell}]\geq\underbrace{\delta h}_{solitude}, \tag{2}$$

where $q_g(x_j,x_i)$ is the probability that an agent $j$ of type $x_j$ and gender $g$ accepts an (opposite-gender) agent $i$ of type $x_i$. Rearranging the above: an $h$ considers the 'gains'—relative to solitude—from marrying another $h$ to the losses from marrying an $\ell$, taking into account the probability of rejection and weighting the relative probabilities of each type, and taking into account the information conveyed by the event of being accepted (i.e., the acceptance curse). Hence, an $h$ of gender $g$ accepts if

$$pf_h(s)\underbrace{[q_{g'}(h,h)(h-\delta h)-(1-q_{g'}(h,h))L_g]}_{\text{E(an h's 'gains' if playing A vs. an } h)}\geq(1-p)f_\ell(s)\underbrace{[q_{g'}(\ell,h)(\delta h-\ell)+(1-q_{g'}(\ell,h))L_g]}_{\text{E(an h's 'losses' if playing A vs. an } \ell)}, \tag{3}$$

noting that the first terms on each side express the relative conditional probability the partner is of each type and 'losses' are defined as the negative of gains.

Note that the overall probability of a player $i$ being accepted by some $j$ as a function of types, $q_g(x_j,x_i)$, does not depend on the signal $s$ that player $i$ *observes*, as signals are drawn independently and individuals do not observe the signals of their own type. I.e., in condition 3, only $f_h(s)$ and $f_\ell(s)$ depend on the observed signal $s$. The mlrp then implies that there is a unique $\hat{s}$ such that an $h$ accepts if and only if (s)he observes $s\geq\hat{s}$. That is, high types (of both genders) use 'floor' threshold strategies, accepting only if the signal exceeds their threshold, $\hat{s}_m$ and $\hat{s}_f$ respectively. This implies that an $h$ of gender $g$ accepts an agent of type $x$ with probability

$$q_g(h,x)=1-F_x(\hat{s}_g),\ \text{with}\ g\in\{m,f\}.$$

### 5.1.2 Low types' best replies

The condition for low types of gender $g$ to prefer to accept is similar to (3); using $q_g(h,\ell)=1-F_\ell(\hat{s}_g)$ it is given by

$$pf_h(s)\underbrace{[(1-F_\ell(\hat{s}_{g'}))(h-\delta\ell)-F_\ell(\hat{s}_{g'})L_g]}_{\text{E(an } \ell\text{'s 'net gains' if playing A vs. an } h)}\geq(1-p)f_\ell(s)\underbrace{[q_{g'}(\ell,\ell)(\delta\ell-\ell)+(1-q_{g'}(\ell,\ell))L_g]}_{\text{E(an } \ell\text{'s 'losses' if playing A vs. an } \ell)}, \tag{4}$$

where $g'\neq g$. Again the mlrp implies that the condition is monotone in $s$ and implies there is at most one value of $s$ such that the condition holds with equality. However, it may *never* hold with equality: as low types prefer a marriage to *either* type partner, for $L_g$ close to zero an $\ell$ (of gender $g$) will prefer to play $A$ regardless of the signal received, and *strictly* prefer this unless both types of the opposite gender play 'reject always'.

Rearranging the above, we can characterize the low type's best response. For $L_g$ sufficiently small 'always accept' is a weakly dominant strategy for low types, and a strict best response where any type of the opposite gender sets an interior threshold. More generally, 'low types always accept' strategies (leading to $q_m(\ell,\ell) = q_f(\ell,\ell) = 1$), are mutual best replies if

$$p\frac{f_h(s)}{f_\ell(s)}[(1-F_\ell(\hat{s}_{g'}))(h-\delta\ell) - F_\ell(\hat{s}_{g'})L_g] \geq -(1-p)(\ell-\delta\ell) \text{ for } g \in \{m,f\}, \forall s \in [\underline{s},\bar{s}], \tag{5}$$

which will hold if and only if[17]

$$\underbrace{(1-F_\ell(\hat{s}_{g'}))(h-\delta\ell)}_{\text{an } \ell\text{'s expected gain if plays } A \text{ vs. an } h} \quad \geq \quad \underbrace{F_\ell(\hat{s}_{g'})L_g}_{\ell\text{'s expected LoF if plays } A \text{ vs. an } h} \quad . \tag{6}$$

Condition: $\ell$-types prefer to accept against a *certain h*

This condition also implies that the left-hand side of (4) is non-negative. Intuitively, if low types expect a (weak) gain from accepting against a *certain* $h$, and they know other low types always accept, then no signal will deter them from accepting.

However, even where (6) holds, there may also be an equilibrium where low types do *not* always accept. If low types of the opposite gender are very selective, the expectation of the gain from marrying an $\ell$ may not outweigh the risk of LoF (i.e., the right-hand side of (4) may be positive). Thus, just as the high types do, low types may also use a floor threshold $\hat{s}_{g\ell}$, rejecting after observing signals that are 'too low'. Intuitively, even though other $\ell$'s are less selective, accepting against a certain-$h$ may yield an expected net benefit, while accepting against a certain-$\ell$ may yield an expected loss, because the gain to *marrying* high exceeds the gain to marrying low.

Next consider the case where (6) fails, implying that the left-hand side of (4) is negative, and low types expect a *loss* from accepting against a certain $h$. Here, even if low types of the opposite gender always accept, if there is a large enough chance the match is an $h$, an $\ell$ will prefer to reject. Formally, there is an $\tilde{s} \in (\underline{s}, \overline{s})$ such that a low type (of gender $g$) prefers to reject after observing $s > \tilde{s}$, implying $q_g(\ell,\ell) < 1$. In turn, if $q_{g'}(\ell,\ell) < 1$ (and (6) fails), there is either a unique value of $s$ such that (4) holds with equality, or it never holds.

The former case implies that a male $\ell$ uses a single interior threshold, the latter implies that he never accepts. As low types here seek to *avoid* accepting when matched with a high-type, this threshold must be a *ceiling*, with low types accepting only after observing *lower* signals, i.e., if $s < \check{s}_{g\ell}$, with $\check{s}_{g\ell} = \underline{s}$ for the shut-down response. (We use the inverted hat to distinguish ceiling thresholds.)[18] We call such behaviour 'reverse snobbery'.

### 5.1.3 Summary of best replies

**Lemma 1** (Individual behaviour)**.** *In a Nash equilibrium players use threshold strategies: high types use floors, i.e. 'accept iff $s \geq \hat{s}_g$' for $g = m, f$, and low types may use either floors ('accept iff $s \geq \hat{s}_{g\ell}$') or ceilings ('accept iff $s \leq \check{s}_{g\ell}$'). If low-type males (females) prefer to play A against a certain-h of the opposite sex (i.e., if condition (6) holds for this gender), then low types of this gender use floors. Here, if $L_g$ is sufficiently small, then if females (males) accept with positive probability, then male (female) low-types always accept (i.e., this floor is $\hat{s}_{g\ell} = \underline{s}$) otherwise $\hat{s}_{g\ell} > \underline{s}$. If condition (6) does not hold for males (females), then low types of this gender use ceilings.*

[17] Proof of equivalence: The bracketed term in 5 represents an $\ell' s$ expected net gain, relative to solitude, from accepting when faced with a known $h$. If this is positive, the left-hand side is minimized at $s = \underline{s}$, where it equals zero (zero relative probability of a high-type); this is thus equivalent to 6. If the bracketed term is negative, it is minimized at $s = \overline{s}$, which implies that this condition fails whenever 6 also fails.

[18] Note that all equilibrium strategies will involve only at most a single non-trivial threshold, a floor or a ceiling. Thus, to save notation, where we denote a ceiling threshold $\check{s}$ one can assume a trivial floor threshold $\hat{s} = \underline{s}$ and vice-versa.

The proposition describes the players' best replies: the *mlrp* ensures every type will have a unique optimal threshold value given any behaviour of the other types (see appendix for detailed best-response functions). Summarizing, low types are either *picky*, using floors, they act as *reverse-snobs*, using ceilings, or they are *indiscriminate*, accepting any signal. Note that the responses as characterized in Lemma 1 allow for multiple equilibria. For instance, all types playing 'reject' independently of observed signals is an equilibrium. Moreover, plugging $L_f = L_m = L$ into the above and using symmetry, we see that symmetric LoF implies there is a symmetric equilibrium, where $\hat{s}_m = \hat{s}_f$ and $\hat{s}_{m\ell} = \hat{s}_{f\ell}$, although this may take the form of a coordination failure.

### 5.2 Equilibria and stability

We next derive a sufficient condition for the existence of 'interior equilibria': equilibria where high types of both genders accept with positive probability, i.e., where $\hat{s}_m, \hat{s}_f \in [\underline{s}, \overline{s})$. All work is in online appendix A.0.2.

We consider the best reply functions derived from (3). Note that, independent of $\hat{s}_{f\ell}$ and $\hat{s}_{m\ell}$, a high-type male's best response to $\hat{s}_f = \underline{s}$ is $\hat{s}_m > \underline{s}$, as, even if $h$-type females always accept, a low enough signal implies the match is almost surely an $\ell$. Taking the total differential with respect to $\hat{s}_m$ and $\hat{s}_f$ yields the slope of a high $m$'s best reply function (henceforth, *brf*) in terms of $\hat{s}_f$ (equation A.6 in the appendix). This *brf* has zero slope at $\underline{s}$ and the slope becomes positive for higher values of $\hat{s}_f$ (independent of $\hat{s}_{f\ell}$, $\hat{s}_{m\ell}$, the $\ell$-types' thresholds). Hence, if this slope exceeds unity at $\hat{s}_f = \overline{s}$ (again, independent of $\ell$-types' strategies) then the *brf* crosses the 45° line at least once, implying that an interior equilibrium exists (for graphical intuition, see Figure 2).

Since a player's best reply only depends on his or her *own* LoF parameter $L_g$ this logic applies to all the environments that we consider. The slope of the *brf* at $\hat{s}_f = \overline{s}$ will also determine whether an equilibrium at $\hat{s}_f = \hat{s}_m = \overline{s}$ is tatonnement-stable. This case, where high types 'always reject' (although low types still may accept) has the flavour of a coordination failure; we call this the 'C-F equilibrium'. With large enough $L_g$, this becomes risk-dominant, as increasing LoF decreases the possible loss when unilaterally deviating from an interior equilibrium, and increasing LoF increases the possible loss when deviating from the C-F equilibrium. Proposition 1 states this formally.

**Proposition 1** (Existence and Stability of Interior and C-F Equilibria)**.**

*(a) If $L_g$ is sufficiently close to* 0 *for both genders and, for both genders $g \in \{m, f\}$*

$$f_h(\overline{s})^2 > -f'_\ell(\overline{s})\frac{1-p}{p}\frac{\delta h - \ell}{h - \delta h + L_g}, \tag{7}$$

*i.e., if $f'_\ell(\overline{s}) \leq 0$ is sufficiently close to* 0*, then a tatonnement-stable interior equilibrium with $\hat{s}_m, \hat{s}_f \in (\underline{s}, \overline{s})$ exists.*

*(b) If condition (7) holds, then all 'C-F equilibria'— i.e., equilibria where $\hat{s}_f = \hat{s}_m = \overline{s}$—are tatonnement-stable and trembling-hand perfect if and only if $L_g > 0$ for some $g = m, f$.*

*(c) For large enough $L_g$, the C-F equilibrium must risk-dominate all other equilibria.*

Condition (7) requires the right tail of $f_\ell(s)$ to be sufficiently flat, or the right tail of $f_h(s)$ sufficiently high (implying that the likelihood of having met a high type still increases even for high signal realizations), or the high type's loss from matching with a low type sufficiently low compared to remaining solitary. It would be implied by $\frac{\partial f_\ell(\overline{s})}{\partial s} = 0$, i.e., if the $\ell$'s signal distribution becomes flat at $\bar{s}$. This is sufficient, but by no means necessary, see the numerical example in section 5.2.4. For the remainder of the paper we focus on the case where (7) holds and thus where an interior equilibrium is guaranteed without LoF.[19]

### 5.2.1 CAE: No Loss of Face

Proposition 1 shows that loss of face has a dramatic effect on equilibrium behaviour: in particular, a strategy profile involving coordination failure among high types becomes a stable equilibrium, and possibly the only one. We thus inspect the case of $L_g = 0$ for both genders (corresponding to the CAE, where players know that other players do not observe their action) and examine the effects of increasing LoF. As the corner equilibria are unstable when $L_g = 0$ for both genders, we consider a (stable) interior equilibrium. As noted above (and implied by Lemma 1), without LoF, and where high types do not shut down, the low type's strict best response is to play 'always accept'. Here, high types of both genders face the same optimization problem. A male $h$ will find accepting profitable if

$$\frac{f_h(\hat{s}_m)}{f_\ell(\hat{s}_m)} \geq \frac{1-p}{p} \frac{\delta h - \ell}{(1 - F_h(\hat{s}_f))(h - \delta h)},$$

and analogously for a female $h$. The resulting best-reply function is shown in Figure 2, where the male $h$'s best reply crosses the 45° line exactly once.

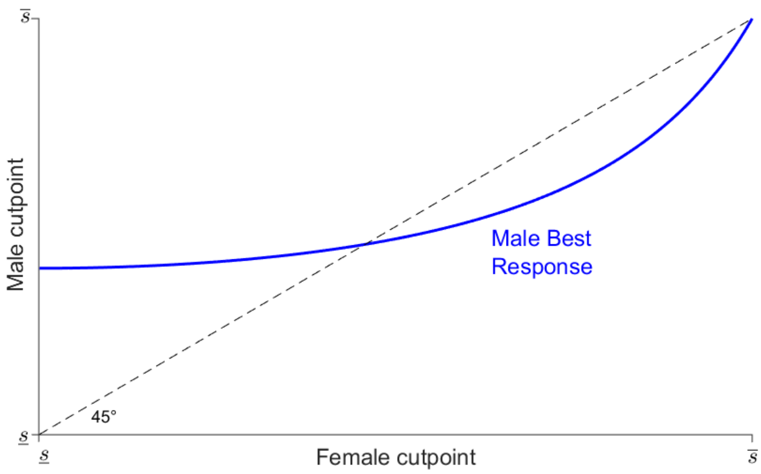



Figure 2: Male high type's best response to high type female cutpoints; the condition (7) in Proposition (1) holds here.

As low types always accept and the game is symmetric by gender, $\hat{s}_m = \hat{s}_f := \hat{s}^*$ must hold in a Nash equilibrium (supposing the contrary leads quickly to a contradiction). As noted above, $\hat{s}^* > \underline{s}$ in any stable equilibrium. Since agents' actions do not affect other agents' information sets, beliefs are always formed according to Bayes' rule, and the issue of out-of-equilibrium beliefs will not arise. This yields the following proposition.

**Proposition 2.** *If $L_g = 0$ for $g \in \{m, f\}$ and Condition (7) holds then at least one interior stable equilibrium exists, and in any* <u>*stable*</u> *equilibrium*

1. *low types always accept,*

2. *high types use symmetric cutoff strategies, accepting if $s > \hat{s}_m = \hat{s}_f := \hat{s}^*$, and*

[19]If (7) does *not* hold for $L_g = 0$, the C-F equilibrium will be stable, and there may or may not also exist stable interior equilibria.

*3.* $\hat{s}^* \in (\underline{s}, \overline{s})$ *defined by*

$$\frac{f_h(\hat{s}^*)}{f_\ell(\hat{s}^*)}\frac{p}{1-p} = \frac{\delta h - \ell}{(1-F_h(\hat{s}^*))(h-\delta h)}. \tag{8}$$

The above also implies that the *trivial* coordination failure (where both types always reject) is unstable without LoF. Simple calculations yield the following results. Where Condition (7) holds, expected payoffs for types $\ell$ and $h$ in a stable equilibrium of the game without LoF (or for any strategy profile where low types always accept) are

$$\begin{aligned} v(\ell) &= \delta\ell + p(1-F_\ell(\hat{s}^*))(h-\delta\ell) + (1-p)(\ell-\delta\ell) \text{ and} \\ v(h) &= \delta h + p(1-F_h(\hat{s}^*))^2(h-\delta h) - (1-p)(1-F_\ell(\hat{s}))(\delta h - l). \end{aligned} \tag{9}$$

Note $v(h) > v(\ell)$. The number of marriages is $(1-p)^2 + 2p(1-p)(1-F_\ell(\hat{s}^*)) + p^2(1-F_h(\hat{s}^*))^2$, which strictly decreases in $\hat{s}^*$.

Intuitively, an $\ell$ will not marry (and will thus get $\delta\ell$) unless he meets another $\ell$ or fools an $h$. An $h$ will marry only if she meets another $h$ and they both send very positive signals, or if she is fooled by an $\ell$ (i.e., she meets an $\ell$ who sends a high enough signal).

As noted, in a stable interior equilibrium without LoF, an $\ell$ always accepts; in fact, even if there is no stable equilibrium,'reject always' is weakly dominated for low types. This implies an $h$ rejects at least against the lowest signals. Thus the acceptance behaviour of players of type $h$ and $\ell$ differs in equilibrium, implying that *being accepted* also conveys some information about the match's type (the 'acceptance curse' in Chade, 2006).

### 5.2.2 Symmetric Full Revelation Environment: (FRE) Positive Cost of Loss of Face

We next consider an environment in which both genders are symmetrically vulnerable to loss of face, implying $L_m = L_f \equiv L$. Considering $L$ increasing from $L = 0$, condition (4) ensures that for small enough $L$, low types still find it optimal to play 'accept' unconditional on the signal in an interior equilibrium. This implies that for a small enough $L$ an $h$ of gender $g \in \{m, f\}$ will have a threshold $\hat{s}_g$ implicitly defined by:

$$\frac{f_h(\hat{s}_g)}{f_\ell(\hat{s}_g)}\frac{p}{1-p} = \frac{\delta h - \ell}{(1-F_h(\hat{s}_{g'}))(h-\delta h) - F_h(\hat{s}_{g'})L}, \tag{10}$$

if $(1-F_h(\hat{s}_{g'}))(h-\delta h) > F_h(\hat{s}_{g'})L$ and $\hat{s}_m = \overline{s}$ otherwise, where $g \neq g'$. Figure 3 shows the male $h$-type's best response to $\hat{s}_f$ with and without positive loss of face.

Since the best replies are symmetric, $\hat{s}_f = \hat{s}_m = \hat{s}_h$ defined by (10), or by $\hat{s}_h^* = \overline{s}$ (the C-F equilibrium).

Proposition 1 states that the C-F equilibrium may arise as a stable equilibrium as one moves from the benchmark setting (the CAE, or in general whenever $L_g = 0$) to an environment with a positive LoF term. Without LoF (where condition 7 holds), only an interior equilibrium is stable; for positive $L$ the C-F equilibrium is always stable. If the C-F equilibrium is plausible, this suggests that LoF may worsen outcomes:

**Remark 1.** *Compared to an interior equilibrium allocation without loss of face, the C-F equilibrium in an environment with LoF induces*

1. *a lower overall marriage rate,*

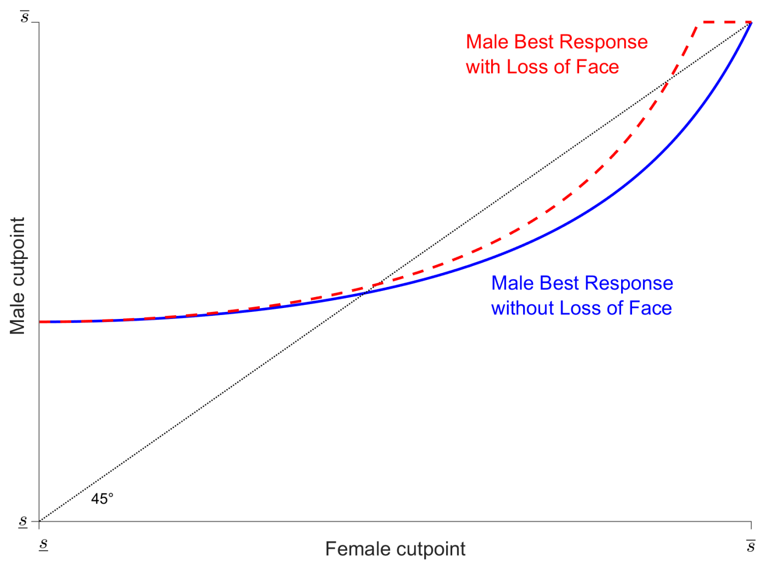



Figure 3: Male high-type's best response to (high type) female cutpoints, with and without loss of face; condition 7 holds here

2. *lower aggregate surplus (even without directly including LoF in the surplus calculation), and*

3. *lower expected surplus (again, even without subtracting the LoF) for both types of both genders.*

*(Details of this remark are in the appendix.)*

There is an important caveat to this remark.[20] In environments where agents must pay a cost to *enter* a matching market the 'C-F outcome' will not arise on the equilibrium path. Under rational expectations an agent will only be willing to bear the entry cost if her expected market equilibrium payoff is strictly positive. Thus, the C-F outcome would not arise in equilibrium on any matching market platform that has customers. However, in the larger game with a participation decision there will still be multiple rational expectations equilibria, including 'all agents expecting a C-F outcome and thus not joining the market'.

We next consider the monotone comparative statics of the equilibrium in $L$. For small enough $L$, low types always accept, implying that equation (10) determines the high types' equilibrium play. We examine the behaviour of (10) (as a system of two equations for $g = m, f$) in the neighbourhood of the equilibrium threshold $\hat{s}^*$ as we increase $L$. This yields the following statement (proof in the appendix).

**Proposition 3.** *Suppose (7) holds where $L_g = L = 0$. Then there is $\bar{L} > 0$ (defined by condition 6 and expression 10) such that for all $L \in [0, \bar{L}]$ there is an interior, stable equilibrium where low types play 'accept' unconditionally, and high types use thresholds $\hat{s}^*$ implicitly defined by (8).*

*Under these conditions, for a small increase in $L$ to $L' \in (L, \bar{L}]$ low types still always accept, i.e. $\hat{s}_\ell = \bar{s}$, and the symmetric equilibrium floor cutoffs for high types will increase in a stable interior equilibrium (and will decrease in an unstable interior equilibrium).*

*For $L > \bar{L}$, in any interior stable equilibrium low types use ceilings, i.e. play 'accept if $s \leq \check{s}_\ell$', and high types use thresholds $\hat{s}_h$ implicitly defined by the best response conditions (given in equations A.4 and A.5 in the online appendix.) If $\bar{L} > \delta h - \ell$, for a small increase in $L$ to $L' \in (L, \bar{L}]$ the symmetric equilibrium ceiling cutoff for low types will decrease, while the symmetric equilibrium floor cutoffs for high types will increase, i.e., both types play 'accept' less often.*

[20]We thank an anonymous referee for this point, which could also be considered a *refinement*.

For intuition, consider that for equilibrium dynamics, becoming more selective by increasing one's cutoff has a twofold effect on the expected quality of a marriage partner. First, there is a screening effect, increasing the expected quality of a match *holding constant* the acceptance behaviour of the other gender. Second there is a supply effect in the opposite direction: if one side becomes more selective, then the other side will react by also becoming more selective, implying a greater acceptance curse on both sides.[21] In the above case, while $L$ remains small, the supply effect only stems from the high types on the other market side.[22]

Next consider where $L > \bar{L}$. Here, the condition $\bar{L} > \delta h - \ell$ implies that the LoF from being rejected exceeds a high type's cost of marrying down, so a high-type who plays *accept* prefers that a low type *accepts* her. This implies that as low types become more reverse-snobbish, high types are *less* motivated to play accept against them, thus they become more selective. The above proposition presents a sufficient condition for this intuitive comparative static.

However, a counter-intuitive response also is possible. If $\bar{L} < \delta h - \ell$, then, for $L \in [\bar{L}\delta h - \ell]$, while $L$ is large enough to make low types become reverse snobs, it is small enough that high types prefer that low types play *reject.* Thus, in this range, as $L$ increases, and low types become more reverse-snobbish, high-types find lower signals less risky, and thus may *decrease* their threshold, becoming *less* choosy. We offer a numerical example of this in the appendix (page 5).

### 5.2.3 Asymmetric Revelation Environment (ARE): Positive Cost of Loss of Face

We turn now to the ARE, where only one market side is subject to LoF. For the sake of concreteness suppose that males move first and know they are observed by females, and thus $L_m = L$ but $L_f = 0$. The ARE is of particular interest. Firstly, it allows us to separate out the direct and indirect effects of LoF; for the vulnerable side and for the opposite side. Secondly, it describes a common situation, and one that can be easily engineered by requiring one side to move first. Thirdly, it provides a simpler environment to show the intriguing possibility that the high-types equilibrium cutoffs may be nonmonotonic in $L$, and that LoF may even increase the rate of successful matches beyond the benchmark (CAE) case!

Propositions 1 to 3 carry over to the ARE almost unchanged. Under Condition (7) the CF-equilibrium is again tatonnement-stable even if only males are vulnerable to LoF.

With asymmetric LoF, equilibrium behaviour is also asymmetric. Analogously to Lemma 1 the three nontrivial thresholds (for high-type females, high-type males, and low-type males) are defined by the system of equations defining the observed signal that makes each indifferent between accepting and rejecting:

$$\frac{p}{1-p}\frac{f_h(\hat{s}_f)}{f_\ell(\hat{s}_f)} = \frac{\delta h-\ell}{h-\delta h}\frac{F_h(\check{s}_{m\ell})}{1-F_h(\hat{s}_m)}, \tag{11}$$

$$\frac{p}{1-p}\frac{f_h(\hat{s}_m)}{f_\ell(\hat{s}_m)} = \frac{\delta h-\ell}{h-\delta h}\frac{1}{1-F_h(\hat{s}_f)(1+L/(h-\delta h))}, \text{ and} \tag{12}$$

$$\frac{p}{1-p}\frac{f_h(\check{s}_{m\ell})}{f_\ell(\check{s}_{m\ell})} = \frac{l-\delta\ell}{h-\delta\ell}\frac{1}{(1-F_\ell(\hat{s}_f))(1+L/(h-\delta\ell))-1}. \tag{13}$$

[21] The equilibrium tradeoff between screening and the acceptance curse was present without LoF. However, in the ARE LoF makes accepting less attractive for males, and this effect is stronger the more females reject, implying a steeper reaction function.

[22] Note that we cannot rule out a 'perverse' equilibrium in the FRE where low types use nontrivial floors even though $L < \bar{L}$; if low-types on one side are very selective, low-types on the other side may may prefer to play reject against them to avoid the risk of LoF, (noting that the *marriage gain* is also higher in the latter case).

Once again high types use floors, $\hat{s}_f$ and $\hat{s}_m$.

In contrast to the FRE, since females face no LoF, by (4) low females always accept (in any stable equilibrium); i.e., $\hat{s}_{f\ell} = \bar{s}$, $\check{s}_{f\ell} = \underline{s}$. Thus (again by (4)) low males must always accept for small $L \geq 0$ that satisfies (6).[23] With severe LoF, (i.e., $L > \bar{L}$ as defined in Proposition 3) they attempt to avoid being rejected and use the signal to screen for low females, using a ceiling threshold $\check{s}_{m\ell}$. Thus, in the ARE, for large $L$ the low males act as *reverse snobs*.[24]

As $L$ increases further low males become more reluctant to accept. This reduces high females' risk of being fooled, making them more eager to accept (lowering their thresholds). This in turn drives down the male $h$-types' thresholds. The implications are intriguing: as $L$ increases from 0, high types become first *less* and then *more* inclined to accept. In contrast, as $L$ increases from 0, low-type male's behaviour first remains constant (always accepting) and then becomes more strict (reverse snobbery).[25,26] This implies that each type's surplus is non-monotonic. The aggregate matching frequency may *also* be non-monotonic, first decreasing and then increasing in $L$, and positive LoF in an ARE may even *increase* the number of successful matches (relative to no LoF), as demonstrated in the next section.[27]

#### 5.2.4 ARE: triangular distribution example illustrating non-monotonicity in L

The effect of changes in $L$ on equilibrium behaviour depends on the parameters and the distribution function; it is ambiguous in general. As we were not able to generally characterize all equilibrium comparative statics, we focus on a convenient specification. Suppose the signal distribution is a triangular distribution of the form $F_h(s) = s^2$ and $F_\ell(s) = 2s - s^2$ with $s \in [0, 1]$. Under this assumption an equilibrium without LoF is given by all low types playing 'accept' and high types using the threshold:

$$\hat{s}^* = \frac{1}{2}\left(\sqrt{1 + 4\frac{1-p}{p}\frac{\delta h - \ell}{h - \delta h}} - 1\right). \tag{14}$$

The equilibrium is interior and stable if $\frac{1-p}{p}\frac{\delta h-\ell}{h-\delta h} < 2$, i.e., if condition (7) is satisfied, implying that the slope of the high types' *brf* is greater than 1 as $s$ approaches $\bar{s}$. Proposition 3 carries over (details in appendix) and thus both $\hat{s}_m$ and $\hat{s}_f$ increase in $L$ for $L \in [0, \bar{L}]$ as defined in the proposition.

[23] This rules out the potential 'perverse' equilibrium in the FRE (see previous footnote), making the non-monotonic response below a more general result.

[24] Reverse snobbery (on both sides) was also possible in the symmetric FRE, if low-types of both genders preferred to reject against a known-high-type (see Lemma 1).

[25] Finally, for *very* large $L$ the CF-equilibrium becomes risk-dominant, as noted above, and the male $\ell$-types' ceiling approaches $\underline{s}$, implying that the overall marriage rate converges to zero.

[26] We conjecture that the ARE has an interior equilibrium for any $L$. As $L \to \infty$, low males accept against only the lowest signals, while low females always accept. High males only accept for high enough signals, while high females (who don't face LoF) accept against all but the lowest signals, as low males have nearly dropped out. Thus high males can accept against higher signals without fear of LoF.

[27] We speculate that for stable, interior equilibria in the ARE, an increase in $L$ reduces marriage payoffs for all low types. For small $L$, high-types' cutoffs increase in $L$ (prop. 3). For $L > \bar{L}$, low males become reverse snobs and lower their ceiling in $\bar{L}$, and high types reduce their floors in response. The net effect of this latter increase in $L$ must harm all low types. This is because high types decrease their floors only when this implies a greater probability of matching other high types, so the decrease in the floor is overcompensated by the decrease in the low type's ceiling, making a mixed marriage less likely.

With a larger loss of face $L > \bar{L}$ term the equilibrium thresholds must satisfy:

$$\hat{s}_f : \quad \tfrac{p}{1-p}\tfrac{\hat{s}_f}{1-\hat{s}_f} = \frac{\delta h - \ell}{h - \delta h}\frac{\check{s}_{m\ell}^2}{1 - \hat{s}_m^2}, \tag{15}$$

$$\hat{s}_m : \quad \tfrac{p}{1-p}\tfrac{\hat{s}_m}{1-\hat{s}_m} = \frac{\delta h - \ell}{h - \delta h}\frac{1}{1 - \hat{s}_f^2(1 + L/(h - \delta h))}, \text{ and} \tag{16}$$

$$\check{s}_{m\ell} : \quad \tfrac{p}{1-p}\tfrac{\check{s}_{m\ell}}{1-\check{s}_{m\ell}} = \frac{l - \delta\ell}{h - \delta\ell}\frac{1}{(2\hat{s}_f - \hat{s}_f^2)(1 + L/(h - \delta\ell)) - 1}. \tag{17}$$

We offer a numerical case of this parametric example. Setting $p = 1/2$, $h = 1$, $\ell = 1/4$ and $\delta = 2/3$ (satisfying condition 7), figure 4 shows the equilibrium outcome as $L$ increases. Indeed both high types' cutoffs first increase in $L$ up to $\bar{L}$, and then both decrease as the low male's cutoff starts decreasing. This implies that as $L$ increases, high-types' chance of getting married first decreases and then increases; as does low types' chance of marrying high (both in absolute and relative terms).

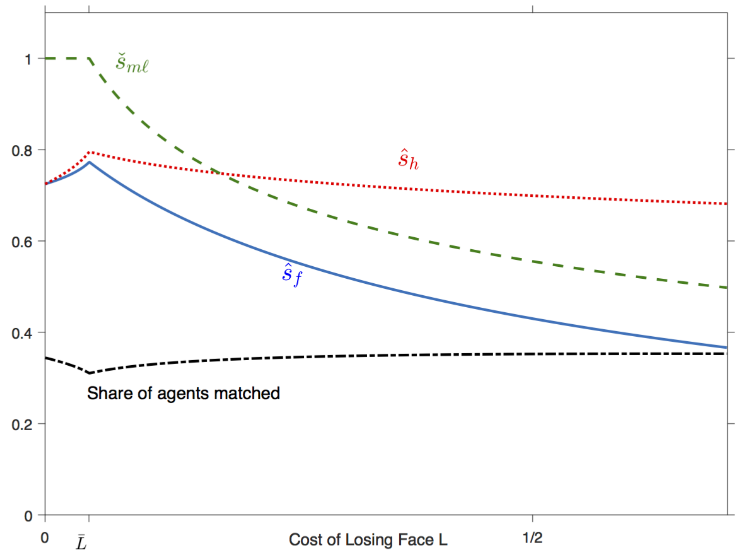



Figure 4: Thresholds and percentage of possible marriages formed as functions of loss of face

Turnover—i.e., the number of marriages formed as a share of possible matches—first declines and then increases in $L$. It may even increase beyond the turnover achieved for $L = 0$; for example no LoF corresponds to a 34.4% turnover, while at $L = 2/3$ turnover is 35.3%.[28] Increases in $L$ are also accompanied by more assortative mating: homogamous ($(h,h)$ or $(\ell,\ell)$) marriages increase as a share of all marriages.

We can extend this to the general parametric example. Suppose that $\frac{1-p}{p}\frac{\delta h - \ell}{\bar{L} + h - \delta h} \geq 1$ holds for the critical value $\bar{L}$ defined by condition 6, and suppose a sufficiently great 'pizazz ratio' $h/\ell$. Then (for this parametric example) $\hat{s}_f$ and $\hat{s}_m$ increase in $L$ for $L < \bar{L}$ and decrease for $L > \bar{L}$, as shown in the appendix (page 5). Let $m(x_f, x_m)$ indicate the measure of marriages between females of type $x_f$ and males of type $x_m$. These responses imply that for $L < \bar{L}$, $m(h,h)$, $m(h,\ell)$ and $m(h,\ell)$ all decrease in $L$, while $m(\ell,\ell)$ remains constant. They also imply that for $L > \bar{L}$, $m(h,h)$ and $m(h,\ell)$, $m(\ell,h)$ increase in $L$, while $m(\ell,\ell)$ decreases in $L$.

This triangular distribution example and the specific case plotted in figure 4 demonstrate the possibility of several non-intuitive outcomes, summarized below (details in the appendix).

[28] We derive overall turnover for the numeric example only, in the appendix.

**Remark 2.** *In an ARE under equilibrium behaviour, several crucial outcomes may be non-monotonic in L, both increasing and decreasing as LoF increases. These outcomes include (i) the high types' probability of getting married, (ii) the low types' probability of marrying a high type, and (iii) the overall marriage rate (turnover).*

#### 5.2.5 Which side is affected more?

Only the market side that *proposes* may incur loss of face, suggesting a contrast from Gale and Shapley (1962), where the *proposers* in their deferred-acceptance algorithm secure better matches in equilibrium. For instance, if men propose to women the men-optimal matching outcome will attain. This method is often used in practice, e.g. the student-optimal algorithm in school choice. (If LoF is relevant here, our setup suggests a potential cost to students, which a CAE can avoid.) However, as noted below, the side that *doesn't* face direct LoF (here, females) may still suffer indirect harm, and this may even exceed the direct cost (to the males).

We consider, for the ARE in general: Is the side that bears the loss of face, (here, males) more affected than the other side? Note first that low males are always at least as selective as low females, as the latter always accept. For high types, the possibility of losing face may make males more reluctant to accept than females. On the other hand, this effect will increase the females' acceptance curse: it will decrease the probability that, given a female is accepted, her match was high; thus making high females more cautious. The first effect dominates:

**Proposition 4.** *In any equilibrium in an ARE with Loss of Face, high males are more selective than high females, i.e.,* $\hat{s}_f \leq \hat{s}_m$*; this holds strictly if* $\hat{s}_m < \overline{s}$*, i.e., if we rule out the C-F equilibrium.*

*Proof.* Let $L > 0$. Suppose that $\hat{s}_f \geq \hat{s}_m$. Then the monotone likelihood property and equations (12, 11 and 13) imply that

$$F_h(\hat{s}_\ell) \geq \frac{1 - F_h(\hat{s}_m)}{1 - F_h(\hat{s}_f)(1 + L/((1-\delta)h))} > 1,$$

a contradiction. □

Thus, considering high types of both genders, unless LoF induces a coordination failure, the gender facing direct LoF will be more 'snobbish' than the gender sheltered from it. This has a surprising extension: under certain conditions the side *not* facing direct LoF (females) may suffer *more* from it!

Note that when $L$ is small enough that $\hat{s}_\ell = \overline{s}$, the probability that a high male marries 'below his station': $1 - F_\ell(\hat{s}^*_m)$, is less than $1 - F_\ell(\hat{s}^*_f)$, the probability that a high female does so.

**Remark 3.** *In any stable equilibrium in an ARE with a small amount of LoF: (i) high males marry less often than high females but get better spouses on average, and (ii) low males marry more often than low females and get better spouses on average; thus for low types, a small amount of LoF on one side reduces the marriage payoffs on the* other *side more.*

The vulnerable side may suffer less even including the direct LoF costs:

**Remark 4.** *In an ARE, for,* $\delta h$ *sufficiently close to* $\ell$*, a small LoF term causes low males' expected total payoffs to decrease less than those of low females even including the direct cost of losing face.*

*Proof.* In the ARE, the change in low types' *total* payoffs as $L$ increases from zero is, for low males and females, respectively:

$$\frac{\partial v(\ell,m)}{\partial L} = -pF_\ell(\hat{s}_f^*) - p(h+L-\delta\ell)\frac{\partial F_\ell(\hat{s}_f^*)}{\partial s}\frac{\partial \hat{s}_f^*}{\partial L} \text{ and}$$
$$\frac{\partial v(\ell,f)}{\partial L} = -p(h-\delta\ell)\frac{\partial F_\ell(\hat{s}_m^*)}{\partial s}\frac{\partial \hat{s}_m^*}{\partial L}.$$

For $L = 0$ the equilibrium is symmetric, so that we know that $\hat{s}_f^* = \hat{s}_m^*$. Moreover, by Proposition 4, in a $\frac{\partial \hat{s}_f^*}{\partial L} < \frac{\partial \hat{s}_m^*}{\partial L}$. Hence (as noted in the previous remark), starting at $L = 0$ a marginal increase in $L$ will decrease male $\ell$ types' expected *marriage* payoffs less than those of female $\ell$ types. Suppose $\delta h$ is arbitrarily close to $\ell$, so high types only slightly prefer solitude to marrying low. This leads high types to become very permissive in the no-LoF equilibrium, i.e., $\hat{s}^*$ will approach $\underline{s}$ (as clearly seen in equation (14) for the parametric example), implying that $F_\ell(\hat{f}^*)$ will be arbitrarily close to 0 for $\delta h$ close enough to $\ell$. Then $\frac{\partial v(\ell,m)}{\partial L} > \frac{\partial v(\ell,f)}{\partial L}$ for $L$ in a neighbourhood of $L = 0$. □

## 6 Conclusions and suggestions for future work

Our simple models illustrate how the presence and level of loss of face may worsen (or improve) outcomes, providing conditions and intuition for each. There are clear real-world applications. Some mechanisms and policies may be more efficient than others in the presence of LoF concerns, and firms and policymakers should take this into account. Although setting up a *Conditionally Anonymous Environment* may take some administrative effort, and may require a third-party monitor, we imagine many cases in which it will lead to more and better matches and improve outcomes. Consider, for example, the matching of advisors and students in a Ph.D. program. A 'tick box system' might work, although some might be reluctant to participate in such an impersonal system. More generally, the use of a knowledgable, reliable, and discrete intermediary, might be more effective. Our paper motivates the use of such 'matchmakers' in many contexts.[29] We further note (considering the *ARE*) that if only one side is vulnerable to LoF costly intermediaries may not be necessary; it would be sufficient to let the other side choose first ('propose').

However, in considering implementing a *CAE*, designers should look closely at the extent to which LoF seems to be shutting down markets and how it is affecting participants' strategies. As seen in the parametric example, LoF may also *improve* outcomes if it induces low types to become reverse snobs, and this leads high types to become less selective. However, such gains come at the expense of low types and, at least in the parametric example, lead to increased assortative mating and perhaps greater inequality.

Our modelling can be expanded and generalized. For example, while we assume linear payoffs in the match's type, future work could consider super- or sub-modularities in the marriage production function.

In a model allowing both inherent LoF and reputation, the effects of revealing offers on match efficiency may be complex. If a player is known to be vulnerable to LoF, his making an offer might actually be interpreted as a signal of his *confidence* that he will be accepted, thus a positive signal about his own type. Whenever a player rejects another, there is some possibility that he did so merely

[29] Merely encouraging face-to-face meetings may allow colleagues to reveal their potential interest slowly and conditionally, lessening the risk of LoF from a 'desperate bid'. This may help explain Boudreau et al (2012); who exogenously facilitated brief meetings between local scientists, and found significant increases in their probability of collaborating.

to avoid losing face; noting this possibility should presumably 'soften the blow' to a player's reputation when he is rejected.

Relaxing the assumptions further, preferences over types may be heterogeneous or involve a horizontal component, this may change the equilibrium reputation effects of revealing offers. We might also consider the effects of a player who is either altruistic—suffering when the other player loses face—or spiteful, relishing in making others lose face. Consider a sequential game where only the first mover is vulnerable to LoF and the second mover is a known altruist. Here the first-mover might manipulate this altruism, playing 'accept' and in effect guilting the second-mover into marrying her; this could lead to inefficient matching.

Empirically, our anecdotal and referential evidence for LoF should be supplemented by experimental evidence. Field experiments (or contextual lab work) in the mold of Lee and Niederle (2015) will help identify preferences and beliefs. Abstract 'induced values' experiments may also shed light on strategic play and coordination in our simple environment. While a variety of experimental papers, (see footnote) consider such environments, these do not (i) rely on homegrown preferences and beliefs over social interactions or partnerships, (ii) have face-to-face interaction, (iii) test the single-shot matching of our model, (iii) compare environments such as our CAE and ARE, nor (iv) have a subjects' previous choices and history reported to later matches.[30] By varying whether choices are revealed on one side, both sides, or neither side, we can identify how fear of LoF affects strategic play independent of self-image concerns and curiosity motives. However, distinguishing inherent LoF from reputation concerns may be more challenging; this will require an environment where LoF seems likely to be psychologically meaningful, but full anonymity is common knowledge.

As well as strengthening the evidence for the existence of the LoF motivation, these experiments should examine the causes and correlates of LoF, and its efficiency consequences in various environments. Do people act strategically to minimize their own risk of LoF? Will they be willing to pay to preserve the anonymity of their offers? Who is most affected by loss of face and when (considering sex, race, popularity, status, psychometric measures, etc.)? How can these issues be addressed to improve matching efficiency in real-world environments?

Our results, supplemented by empirical work, will have important implications for government and managerial policy. Search and matching models examining the workings of labour market policies may need to adjust for the presence of LoF. Our research suggest that policies that subsidize or encourage sending applications will appear more advantageous. Organizations may want to closely consider when offers, payments, proposals, and attempts should be made transparent, and when they should be obscured. Matchmakers and middlemen in many areas, from actual marriage brokers to career 'headhunters' to venture capital intermediaries may want to guarantee that unrequited offers will be kept secret. As previously noted, secrecy may be helpful for the success of both international negotiations and negotiating over business mergers. Both parties may want a mutual guarantee that no offers or proposals will be leaked. Finally, we note a hub of 'sharing economy' organizations promoting forms of cooperation and sharing that appear efficient but are not yet widely practiced. The fear of LoF may have served as a barrier to these activities in the past; setting up a 'risk-free partnering exchange' may be helpful.

[30]Recent work considers symmetric horizontal matching preferences in an anonymous laboratory setting. E.g., Echenique et al. (2016) and Pais et al. (2012) allow subjects to make and reject/accept offers sequentially over a certain duration, in small groups. The former offers evidence that stability is a good predictor of market outcomes with complete information over preferences. The latter finds that making offers costly leads to fewer and slightly less ambitious offers, less efficiency and less stable matchings. Incomplete information boosts both stability and efficiency.

## Supplementary material

Supplementary material is available on the OUP website. This is the online appendix.

## Acknowledgments

We woud like to thank George Akerlof, Alejandro Atatsuo Moreno, Lori Cantwell, David Hugh-Jones, Amrish Patel, Ted Hampson, Ron Harstad, Joon Song, Justin Tumlinson, Jason Quinley, Charles M. Schultz, Jeroen van de Ven, Jason Quinley, and audiences at the ESI Pre-Spring Workshop, the University of Essex, Loughborough University, and the 1st workshop on psychological game theory (Gothenburg), for helpful feedback and ideas.

## References

H. Adachi. A search model of two-sided matching under nontransferable utility. *Journal of Economic Theory*, 113(2):182–198, 2003.

G. A. Akerlof and R. E. Kranton. Economics and identity. *The Quarterly Journal of Economics*, 115 (3):715–753, 2000.

A. Anderson and L. Smith. Dynamic matching and evolving reputations. *The Review of Economic Studies*, 77(1):3–29, 2010.

D. Ariely and S. Jones. *The upside of irrationality*. HarperCollins New York, 2010.

T. Armstrong. *Breaking the ice: rapprochement between East and West Germany, the United States and China, and Israel and Egypt*. United States Inst of Peace Press, 1993.

P. Battigalli and M. Dufwenberg. Guilt in games. *American Economic Review*, 97(2):170–176, 2007.

L. A. Baxter and W. W. Wilmot. Secret tests. *Human Communication Research*, 11(2):171–201, 1984.

G. S. Becker. A Theory of Marriage: Part I. *Journal of Political Economy*, 81(4):813, 1973.

K. Boudreau, I. Ganguli Prokopovych, P. Gaule, E. C. Guinan, and K. R. Lakhani. Colocation and scientific collaboration: evidence from a field experiment. *Harvard Business School Working Paper, No. 13-023*, pages 1–30, 2012.

C. A. Bredow, R. M. Cate, and T. L. Huston. Have we met before? a conceptual model of first romantic encounters. In S. Sprecher, A. Wenzel, and J. Harvey, editors, *Handbook of relationship initiation*, pages 3–28. Psychology Press, London, 2008.

K. Burdett and M. Coles. Marriage and class. *Journal of Economic Theory*, 129:81–113, 2006.

K. Burdett and M. G. Coles. Long-Term Partnership Formation: Marriage and Employment. *The Economic Journal*, 109(456):307–334, 1999.

J. Cawley. A Guide and Advice for Economists on the U . S . Junior Academic Job Market ( 2011-2012 Edition ). *Job Openings for Economists*, (5984), 2012. URL `https://papers.ssrn.com/Sol3/papers.cfm?abstract_id=1933341`. Accessed: 10 Dec 2018.

H. Chade. Matching with noise and the acceptance curse. *Journal of Economic Theory*, 129(1):81–113, 2006.

CNN International. Speed dating's executive makeover, 2005. URL `http://edition.cnn.com/2005/BUSINESS/01/27/speed.networking/index.html`. Accessed: 10 Dec 2018.

N. Collins and J. Goyder. Speed Dating: a process of forming undergraduate student groups. *eCULTURE*, 1(1):9, 2013.

J. Crook. Switch, With $2 Million In Seed Funding, Is Tinder For Jobs — TechCrunch, 2015. URL `https://techcrunch.com/2015/05/22/switch-with-2-million-in-seed-funding-is-tinder-for-jobs/`. Accessed: 10 Dec 2018.

Direct Search Allowance. Jobs in the Staffing Industry: How to Avoid the Counter-Offer Trap, 2007. URL `http://directsearchalliance.blogspot.co.uk/2007/08/how-to-avoid-counter-offer-trap.html`. Accessed: 10 Dec 2018.

A. Dixit. Comparative statics for oligopoly. *International Economic Review*, 27(1):107–122, 1986.

W. Douglas. Affinity-testing in initial interactions. *Journal of Social and Personal Relationships*, 4 (1):3–15, 1987.

P. W. Eastwick and E. J. Finkel. Speed-dating: A powerful and flexible paradigm for studying romantic relationship initiation. In S. Sprecher, A. Wenzel, and J. Harvey, editors, *Handbook of relationship initiation*, pages 217–234. Psychology Press, London, 2008.

A. S. Eban. *The New Diplomacy: International Affairs in the Modern Age*. Random House NY, 1983.

F. Echenique, A. J. Wilson, and L. Yariv. Clearinghouses for two-sided matching: An experimental study. *Quantitative Economics*, 7(2):449–482, 2016.

N. I. Eisenberger and M. D. Lieberman. Why rejection hurts: a common neural alarm system for physical and social pain. *Trends in Cognitive Sciences*, 8(7):294–300, 2004.

A. Falk, D. Huffman, and U. Sunde. Do I Have What It Takes ? Equilibrium Search with Type Uncertainty and Non-Participation. Accessed: 10 Dec 2018, 2006. URL `http://ftp.iza.org/dp2531.pdf`.

D. Gale and L. S. Shapley. College Admissions and the Stability of Marriage. *The American Mathematical Monthly*, 69(1):9–15, jan 1962.

B. J. Geery. Contests Combat Call Reluctance (Sidebar). *Telemarketing and Call Center Solutions*, 15:112–115, 1996.

E. Goffman. *Interaction Ritual: Essays in Face to Face Behavior*. Aldine Transaction, New Brunswick, USA, 2005.

F. H. Hahn. The stability of the Cournot oligopoly solution. *The Review of Economic Studies*, 29(4): 329–331, 1962.

J. C. Harsanyi and R. Selten. A General Theory of Equilibrium Selection in Games. *Behavioral Science*, 1:378, 1988. doi: 10.2307/2073314.

G. J. Hitsch, A. Hortacsu, and D. Ariely. Matching and Sorting in Online Dating. *The American Economic Review*, 100(1):130–163, 2010.

D. Hugh-Jones and D. Reinstein. Losing face. *University of Essex, Department of Economics, Economics Discussion Papers 691*, 2010.

C. Jönsson and K. Aggestam. Diplomacy and Conflict Resolution. *The SAGE handbook of conflict resolution*, page 33, 2008.

B. Köszegi. Ego utility, overconfidence, and task choice. *Journal of the European Economic Association*, 4(4):673–707, 2006.

S. Lee and M. Niederle. Propose with a rose? signaling in internet dating markets. *Experimental Economics*, 18(4):731–755, 2015.

B. London, G. Downey, C. Bonica, and I. Paltin. 15.7.7._Social causes and consequences of rejection sensitivity. *Journal of Research on Adolescence*, 17(3):481–506, 2007. ISSN 10508392. doi: 10.1111/j.1532-7795.2007.00531.x.

I. MarketData Enterprises. *The US Dating Services Markets*. Tampa, Fla.: Marketdata Enterprises, 2012.

R. Mendoza-Denton, G. Downey, V. J. Purdie, A. Davis, and J. Pietrzak. Sensitivity to status-based rejection: Implications for African American students' college experience. *Journal of Personality and Social Psychology*, 83(4):896–918, 2002. URL `http://doi.apa.org/getdoi.cfm?doi=10.1037/0022-3514.83.4.896`.

P. Milgrom and C. Shannon. Monotone Comparative Statics. *Econometrica*, 62(1):157, 1994. doi: 10.2307/2951479.

J. Pais, Á. Pintér, and R. F. Vesztef. Decentralized Matching Markets: A Laboratory Experiment. Technical report, ISEG-Departamento de Economia, 2012. URL `http://ideas.repec.org/p/ise/isegwp/wp082012.html`. Accessed: 10 Dec 2018.

R. D. Putnam. *Bowling Alone: The Collapse and Revival of American Community*. Simon & Schuster New York, 2000.

M. Rabin. Incorporating Fairness into Game Theory and Economics. *American Economic Review*, 83(5):1281–1302, 1993.

M. J. Rosenfeld and R. J. Thomas. Searching for a mate: The rise of the Internet as a social intermediary. *American Sociological Review*, 77(4):523–547, 2012.

J. Shanteau and G. F. Nagy. Probability of acceptance in dating choice. *Journal of Personality and Social Psychology*, 37(4):522, 1979.

D. Simundza. Playing Hard to Get: Strategic Misrepresentation in Marriage Markets. Accessed: 10 Dec 2018, 2015. URL `http://economics.ucdavis.edu/events/papers/1013Simundza.pdf`.

G. S. Sudai and D. J. Blumberg. Method and apparatus for detection of reciprocal interests or feelings and subsequent notification, Sept. 7 1999. URL `http://www.freepatentsonline.com/5950200.html`. US Patent 5,950,200, Accessed: 10 Dec 2018.

Tinder. Tinder — Match. Chat. Meet. Modern Dating., 2018. URL `https://www.gotinder.com/press?locale=en`. Accessed: 10 Dec 2018.

X. Vives. Complementarities and Games: New Developments. *Journal of Economic Literature*, 43(2): 437–479, 2005.

K. D. Williams. Ostracism. *Annual Review of Psychology*, 58:425–452, 2007.

E. Witt. How the Tinder App Became a Success — GQ, 2014. URL `https://www.gq.com/story/tinder-online-dating-sex-app`. Accessed: 10 Dec 2018.

# A Proofs and characterizations

### A.0.1 Best responses (Lemma 1)

The following conditions describe male best responses *given female strategies*, and expresses these for the case of symmetric behaviour. Responses for females are analogous (switching $m$ and $f$).

1. If a male $\ell$ prefers to play $A$ against a certain-$h$ female, i.e., if condition (6) holds for males, i.e., $(1-F_\ell(\hat{s}_{g'}))(h-\delta\ell)-F_\ell(\hat{s}_{g'})L_m \geq 0$, then male $\ell$-types use floors, i.e. 'accept iff $s \geq \hat{s}_{m\ell}$'. If the comparable condition holds for female $\ell$-types, they also use floors, and $\hat{s}_{m\ell} = \max\{\underline{s}, \min\{\overline{s}, \hat{s}\}\}$, where $\hat{s}$ satisfies:
$$\frac{f_h(\hat{s})}{f_\ell(\hat{s})}\frac{p}{1-p} = \frac{F_\ell(\hat{s}_{f\ell})L_m - (1-F_\ell(\hat{s}_{f\ell}))(\ell-\delta\ell)}{(1-F_\ell(\hat{s}_f))(h-\delta\ell)-F_\ell(\hat{s}_f)L_m}. \tag{A.2}$$
Under these conditions, male $h$-types play 'accept iff $s \geq \hat{s}_m$'. The floor threshold $\hat{s}_m = \max\{\underline{s}, \min\{\overline{s}, \hat{s}\}\}$, where $\hat{s}$ satisfies:
$$\frac{f_h(\hat{s})}{f_\ell(\hat{s})}\frac{p}{1-p} = \frac{(1-F_h(\hat{s}_{f\ell}))(\delta h-\ell)+F_h(\hat{s}_{f\ell})L_m}{(1-F_h(\hat{s}_f))(h-\delta h)-F_h(\hat{s}_f)L_m}. \tag{A.3}$$

2. Otherwise $\ell$-type males use ceilings, i.e., 'accept iff $s \leq \check{s}_{m\ell}$'. Again, in a symmetric equilibrium, where the comparable condition holds for females, $\check{s}_{m\ell} = \max\{\underline{s}, \min\{\overline{s}, \check{s}\}\}$, where $\check{s}$ satisfies:
$$\frac{f_h(\check{s})}{f_\ell(\check{s})}\frac{p}{1-p} = \frac{(1-F_\ell(\check{s}_{f\ell}))L_m - F_\ell(\check{s}_{f\ell})(\ell-\delta\ell)}{(1-F_\ell(\hat{s}_f))(h-\delta\ell)-F_\ell(\hat{s}_f)L_m}. \tag{A.4}$$
Under these conditions, male $h$-types use floor thresholds $\hat{s}_m = \max\{\underline{s}, \min\{\overline{s}, \hat{s}\}\}$, where $\hat{s}$ satisfies:
$$\frac{f_h(\hat{s})}{f_\ell(\hat{s})}\frac{p}{1-p} = \frac{F_h(\check{s}_{f\ell})(\delta h-\ell)+(1-F_h(\check{s}_{f\ell}))L_m}{(1-F_h(\hat{s}_f))(h-\delta h)-F_h(\hat{s}_f)L_m}. \tag{A.5}$$

### A.0.2 Proof of Proposition 1

Totally differentiating (3), using $q_g(h,h) = F_h(\hat{s}_g)$, and rearranging, the slope of a high male's (the female case is analogous) best reply function (henceforth '*brf*') must follow:

$$\frac{\partial \hat{s}_m}{\partial \hat{s}_f} = \frac{f_h(\hat{s}_m)f_h(\hat{s}_f)(h-\delta h+L_m)}{f_h'(\hat{s}_m)[(1-F_h(\hat{s}_f))(h-\delta h+L_m)-L_m] - \frac{1-p}{p}f_\ell'(\hat{s}_m)[q_f(\ell,h)(\delta h-\ell)+(1-q_f(\ell,h))L_m]}. \tag{A.6}$$

Therefore $\frac{\partial \hat{s}_m}{\partial \hat{s}_f} = 0$ if $\hat{s}_f = \underline{s}$. Inspecting (3) we see that $\hat{s}_m > \underline{s}$ if $L_m > 0$ or if $q_f(\ell,h) > 0$. Recall that if $L_m = 0$, low males always accept (other than in a trivial coordination failure equilibrium). This implies low females prefer to accept at least against the lowest signals, implying $q_f(\ell,h) > 0$. Thus $\hat{s}_m > \underline{s}$ for any high female brf, implying a high male's brf is strictly positive at $\underline{s}$. Using (3), the best-reply $\hat{s}^*_m(\hat{s}_f)$ strictly increases in $q_f(h,h)$ and thus in $\hat{s}_f$ and $\hat{s}^*_m(\overline{s}) = \overline{s}$.

Given this, for an interior equilibrium to exist $\hat{s}^*_m(\hat{s}_f)$ must cross the 45° line at least once for some interior $\hat{s}_f$ (and similarly for $\hat{s}^*_f(\hat{s}_m)$). The male brf is flat at its origin, i.e., $\frac{\partial \hat{s}_m}{\partial \hat{s}_f} = 0$ must hold at $\hat{s}_f = \underline{s}$. It follows that if $\hat{s}_m$ intersects the 45° line, there must be at least one intersection such that $\frac{\partial \hat{s}_m}{\partial \hat{s}_f} < 1$. Hence, if a symmetric interior equilibrium exists, there is one that is tatonnement stable. Existence is guaranteed if the slope of $\hat{s}^*_m(\hat{s}_f)$ is greater than one at $\hat{s}_f = \overline{s}$ (and analogously