Online Appendix (not for publication): Cooperation and Mistrust in Relational Contracts

July 26, 2017

1 Infinitely Repeated Games

We now provide theoretical support for the main hypotheses of the paper. In this section, we use a standard repeated games approach. In the next section, we will consider a finite horizon game where we allow for committed fairness types.

Here we consider an infinite horizon version of the game.[1] The common discount factor is $\delta \in (0,1)$. $S(x) \equiv 10x - 10$ is the surplus increase that the buyer receives from a seller who chooses x rather than the minimal quality 1. To simplify, we allow for continuum choice sets $[0,100]$ and $[1,10]$ for p and x, respectively.[2]

1.1 Complete Information

We first consider the complete information game. We focus on equilibria in *trigger strategy profiles* (TSP) such that the buyer (seller) chooses some fixed p^* (x^*) as long as neither player has deviated from his assigned choice; after any deviation, both players choose their respective minimal action. A *trigger strategy equilibrium* (TSE) is a subgame-perfect equilibrium in trigger strategy profiles.[3]

Proposition 1 *Consider the repeated trust contract game with complete information for $\theta \in \{L, H\}$. (p^*, x^*) is sustainable as a TSE for (δ, θ) if and only if*

$$p^* \geq c(x^*, \theta)/\delta, \tag{1}$$

$$S(x^*) \geq p^*. \tag{2}$$

[1] We ignore the first, cheap-talk, stage at the beginning of the game, as the analysis of the repeated game does not depend on the signal.

[2] We define the cost function at non-integer values by linear interpolation.

[3] The results are very similar if we use forgiving strategies where both players only punish once or cut-off strategies where players only punish below a certain level of the other player's action. Details are avalable upon request.

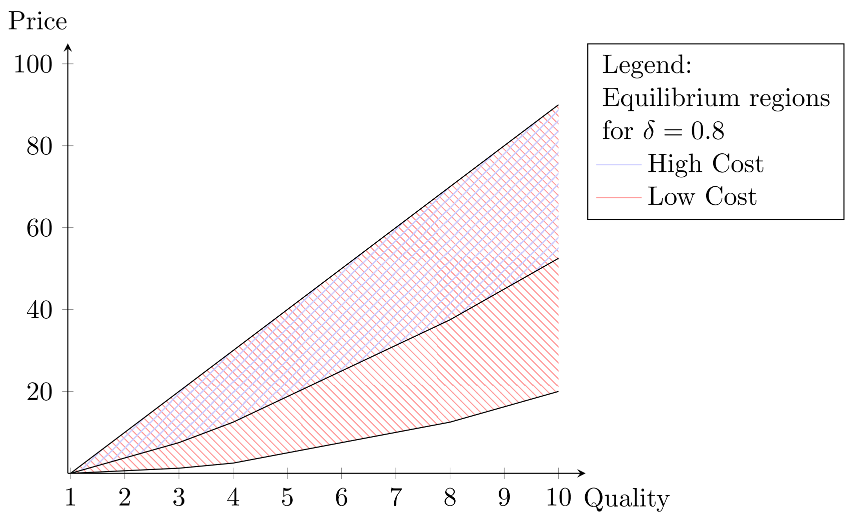

Figure 1: Trigger Strategy Equilibrium

The proof is in the appendix. Condition (1) guarantees that the seller who expects a discounted future equilibrium payment δp^* is willing to incur the equilibrium costs of $c(x^*, \theta)$ rather than deviate to the minimal quality. Condition (2) guarantees that the buyer is willing to pay the price p^* rather than zero: She must expect an equilibrium quality x^* generating benefits $S(x^*)$ that are high enough to compensate for the prices p^*.

For $\theta \in \{H, L\}$, all quality levels in $[1, 10]$ are sustainable for suitable prices and sufficiently high discount factors: This follows from (1) and (2) because $S(x) > c(x, H) \geq c(x, L)$ for all $x \in [1, 10]$. Figure 1 illustrates the set of sustainable price-quality vectors for a given parameter choice in the trust contract game.[4] The following corollary is immediately intuitive from this figure.

Corollary 2 *In the repeated trust contract game, fix* $\delta > c(10, H) / S(10)$.[5]
(i) For any $\theta \in \{L, H\}$*, the lower and upper bound of the set of sustainable* x^* *are both increasing in* p^*.
(ii) For any $(\theta, p^*) \in \{L, H\} \times [0, 100]$*, the maximal sustainable* x^* *is weakly higher for* $\theta = L$ *than for* $\theta = H$.
(iii) For any $(\theta, x^*) \in \{L, H\} \times (1, 10]$*, the minimal* p^* *for which* x^* *is sustainable is higher for* $\theta = H$ *than for* $\theta = L$.

Part (i) of Corollary 2 provides support for Hypothesis 2 which states that we should expect higher quality and higher price going hand in hand. Parts (ii) and (iii) of Corollary 2 both state in slightly different ways that, consistent

[4] We fixed $\delta = 0.8$.

[5] This condition guarantees that (1) and (2) both hold for some prices. In our parametrization, it corresponds to $\delta \geq 0.42$.

with Hypothesis 3(i), high quality is harder to sustain with high costs than with low costs: For any given price, the maximal equilibrium quality is higher in the latter case than in the former; similarly, for any given quality, the minimal required price is lower. Finally, the second part of Hypothesis 3 reflects the fact that, according to Condition (1), the lower bound of the equilibrium set is proportional to the costs of producing the respective output. Thus, for high marginal costs this bound is steeper than for low marginal costs. Intuitively, for any given quality increase, the seller requires a greater compensation with high costs than with low costs.

1.2 Incomplete Information

We now proceed to the incomplete information case. We confine ourselves to the subgame starting after the signal choice of the seller. We suppose that the seller has type L with probability $\mu \in (0,1)$; and this probability is common knowledge. We think of this probability as the exogenous probability corresponding to the type distribution. We consider pooling strategy profiles in the incomplete information game that correspond to those used in the complete information case. A *pooling trigger strategy equilibrium (PTSE)* is a weakly perfect Bayesian equilibrium in which the players use trigger strategies.

Proposition 3 *In the trust contract game with incomplete information,* (p^*, x^*) *is sustainable as a PTSE given* δ *if and only if (2) and (1) with* $\theta = H$ *hold.*

Intuitively, in a pooling equilibrium, the binding incentive constraint is that the high-cost types are willing to provide the desired quality. The observed average quality is the same with incomplete information and with complete information. This provides a first way to support Hypothesis 1.

2 Fairness

In the following, we discuss Hypothesis 1 in a setting with finite horizon when some players are commitment types à la Kreps et al. (1982). Specifically, the analysis relies on the idea that some players have social preferences. We assume that, for both buyers and sellers, a fraction $\phi \in (0,1)$ are fair types (F), whereas the remaining $1-\phi$ are selfish types (S) who maximize monetary payoffs. A fair type responds to the previous action of the other player by choosing what he considers a fair response, as long as he does not put probability 1 onto the other player being a selfish type, in which case he takes the minimal action. There is (two-sided) asymmetric information about fairness types. Thus even the game with complete cost information becomes a game of two-sided asymmetric information, and the game with one-sided incomplete cost information becomes a game with two-sided asymmetric information with two-dimensional seller types: Buyers have types S (selfish) and F (fair). Sellers have types SH, SL, FH and FL, where the first letter corresponds to the preference type (selfish or fair) and the second letter corresponds to the cost type (high or low).

2.1 Complete Cost Information

First, consider the game with complete cost information with cost type $\theta \in \{H, L\}$. In the following, let $p_\theta(x)$ be the price that splits the surplus $10x-10-C(x,\theta)$ evenly for the quality level x, so that $10x-10-p_\theta(x) = p_\theta(x) - C(x,\theta)$. Similarly, we write $x_\theta(p)$ for the inverse function, that is, the quality that splits the surplus equally given the price p.[6] We assume that fair sellers behave as follows:

Assumption A1: *When fair types are sure they are facing a selfish type, they choose the minimal action. Otherwise they behave as follows:*
(i) Fair sellers with cost type θ respond to a price p by choosing $x_\theta(p)$.
(ii) Fair buyers start by choosing $p_\theta(10)$. Thereafter, they respond to any quality level x by choosing $p_\theta(x)$ in the following period.

Fair types are thus committed to non-strategic surplus-splitting behavior, unless they are sure they are facing a selfish type. Intuitively, selfish types can benefit from imitating the fair types and not revealing their selfishness. To confirm this intuition, we investigate the conditions under which an *efficient pooling equilibrium* with the following strategies and beliefs exists:

1. Fair types choose actions as described in Assumption A1.
2. In all periods except the last one, selfish types choose the same actions as fair types, unless they have previously deviated. In the last period, they choose their minimal action.
3. Beliefs about fairness types are given by priors unless the other player has chosen at least one action that a fair player would not choose. In the latter case, beliefs are that the other player is a selfish type with probability one.

If such an equilibrium exists, it involves efficient quality levels, except in the last period for selfish players. Moreover, the price is always $p_\theta(10)$.

Proposition 4 *Suppose A1 holds. An efficient pooling equilibrium exists if and only if $\phi > \frac{p_\theta(10)}{90}$. As $p_L(10) = 53$ and $p_H(10) = 66$, the required shares of fair players are 0.59 and 0.73 for $\theta = L$ and H, respectively.*

The essential part of the proof (see Appendix below) is to make sure that selfish buyers want to pool, even in the final period. This requires a sufficient share of fair players.

2.2 Asymmetric Cost Information

When there is incomplete information about cost types, it is not immediately clear what "fair prices" are, even when buyers and sellers intend to split the surplus fifty-fifty. In particular, players' fairness assessments could depend on their

[6] If the price is so high that the maximal quality $x = 10$ leaves at least half the surplus to the seller, then $x_\theta(p) = 10$.

informational situation, resulting in disagreements. Moreover, one might argue whether fair behavior implies truth-telling or not. We consider two different settings, with different assumptions in these two dimensions.

2.2.1 Setting 1: Pure quality commitment and agreement about fair behavior

We first consider a setting where (1) fair types are not committed to truth telling and (2) both players agree on what constitutes a fair price response $p^f(x)$ to a given quality level x; with the inverse function $x^f(p)$ capturing the quality that is a fair response to a price p.

With respect to (2), we have two polar cases. First, we consider *credulous* buyers who take a high signal for granted and consider $p^f(x) = p_H(x)$, the fairness price corresponding to high-cost agents under complete information, as an adequate price. Second, we consider *skeptical* buyers who do not think a signal is informative and consider $p^f(x) = \lambda p_L(x_{t-1}) + (1-\lambda) p_H(x_{t-1})$, the expected fairness price corresponding to the share of each cost type in the population, as an adequate price. In each case, we assume that sellers have the same understanding of fairness, as will be detailed in the following assumptions that apply to credulous and skeptical buyers.

Assumption A2: (i) *When fair buyers are sure they are facing a selfish type, they choose the minimal price.*
(ii) When fair buyers assign the prior probability ϕ to the seller being fair and probability 1 to cost type θ, they choose $p_\theta(x)$ as in the complete cost information case θ.
(iii) When fair buyers assign probability ϕ to the seller being fair and λ to the cost type being low, they choose price $p_t(x_{t-1}) = p^f(x_{t-1})$; in period 1, they choose price $p^f(10)$.

The assumptions on fair sellers are similar.

Assumption A3: (i) *When fair sellers are sure they are facing a selfish type, they choose the minimal action. They are indifferent about the type they state.*
(ii) When fair sellers assign probability ϕ to the buyer being fair and expect that the buyer assigns probability 1 to cost type θ, they choose $x_t(p_t) = x_\theta(p_t)$ as in the complete cost information case with cost type θ.
(*iii*) *When fair sellers assign probability ϕ to the buyer being fair and expect that the buyer assigns probability $\lambda \in (0,1)$ to the cost type being low and probability $\phi \in (0,1]$ to the seller being fair, they choose actions $x_t(p_t) = x^f(p_t)$.*

We are interested in the existence of an *efficient pooling equilibrium* as defined in Section 2.1 (except that Assumption A1 is replaced with Assumption A2 and A3). In such an equilibrium, selfish players pool with fair players (who play A2 and A3), except that selfish sellers choose the minimal quality in the final period; beliefs are given by priors. Beliefs correspond to prior as long as only equilibrium play is observed; otherwise players assume the other player is selfish for sure.

Proposition 5 *Suppose A2 and A3 hold. There exists an efficient pooling equilibrium if and only if* $\phi > \frac{p^f(10)}{90}$*, no matter whether buyers are credulous or skeptical.*

While Proposition 5 applies to credulous as well as skeptical buyers, the implications are very different in the two cases. For credulous buyers, $p^f(10) = p_H(10) = 66$. Thus, the condition in the proposition is identical to the one in Proposition 4. Intuitively, as the buyer is prepared to accept the high signal as if it were verifiable, everything is as under complete information. In this case, Proposition 5 thus provides no reason to expect that incomplete information has an effect on quality.

With skeptical buyers, $p^f(10) \in (p_L(10), p_H(10)) = (53, 66)$. Thus, the condition for the efficient pooling equilibrium in Proposition 5 is easier to satisfy than the condition in Proposition 4: Sellers accept a lower price, because they accept the buyer's uncertainty about costs. In this sense, one might conjecture that incomplete information increases quality.

2.2.2 Disagreement on fair behavior; truth-telling commitment

We now modify the set-up of the previous section. First, we modify Assumption A3 by assuming that the fair seller is also committed to signalling the true type (Assumption A3'). Second, we allow for disagreements on what constitutes fair behavior that reflect the informational situation of the players. We will be particularly interested in an equilibrium which separates between type FL and the three remaining types.

Suppose that a buyer is sure he is facing FH, SL or SH, with relative probabilities given by priors. Denote the share of fair types FH in the pool consisting of FH, SL and SH as σ. Moreover, we denote the share of low-cost types in this pool as λ. For our purposes, it suffices to modify Assumption A2 on fair buyers slightly.

Assumption A2': (i), (ii): as in A2.
(iii) When fair buyers assign probability σ *to the seller being fair and* λ *to the cost type being low, they choose price* $p^B(x) = \gamma p_H(x) + (1-\gamma)p_L(x)$ *for some* $\gamma \in (0,1)$*; in period 1, they choose price* $p_1 = p^B(10) = \gamma p_H(10) + (1-\gamma)p_L(10)$.

Part (iii) of the assumption states that the buyer considers a price as fair that is lower than if she was sure that the cost type was low, but higher than if she was sure that the cost type was high. The assumptions on fair sellers are similar.

Assumption A3': (i),(ii):as *in A3.*
(iii) *Fair sellers always truthfully reveal their type.*
(iv) *When fair sellers assign probability* σ *to the buyer being fair and expect that the buyer assigns probability* $\lambda \in (0,1)$ *to the cost type being low and* $\phi \in (0,1]$ *to the seller being fair, they choose actions* $x(p)$ *such that the inverse function satisfies* $p^S(x) = \delta p_H(x) + (1-\delta)p_L(x)$ *for some* $\delta \in (0,1)$ *such that* $\delta \geq \gamma$.

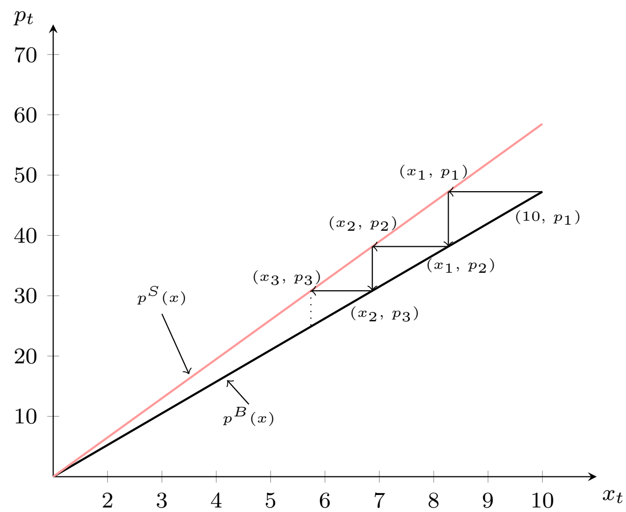

Figure 2: Breakdown of cooperation

Part (iii) extends the commitment of the fair player to truth-telling. Part (iv) states that, like the buyer, the seller considers a price as fair that is lower than if the buyer was sure that the cost type was low, but higher than if the buyer was sure that the cost type was high. Crucially, however, whenever $\delta > \gamma$, he thinks that the price should be closer to $p_H(x)$ than the buyer does. As a result, her response to a given price is lower than what the seller perceives to be a fair quality.

Thus, our assumptions reflect the idea that under asymmetric information about cost types both parties have different assessments about fairness. The uninformed party (the buyer) thinks that the adequate price lies somewhere between those corresponding to the low cost and high cost cases with complete information. In principle, the informed party acknowledges this (she understands the skepticism of the buyer to some extent), but, when she has high costs, she considers a higher price as adequate than the seller does.

We now show that the asymmetry in the fairness notions can yield a gradual unravelling of cooperation.

For any combination of γ and δ, we can recursively define a sequence $(p, q)_{\gamma,\delta}$ of prices and qualities as follows. Figure 2 illustrates the construction.[7]

p_1 is given as in Assumption A2'(iii). For all $t \in \{2, ..., 15\}$, p_t is given as $p_t = p^B(x_{t-1})$, the fair response from the perspective of the uninformed buyer. For all $t \in \{1, ..., 15\}$, x_t is given by the requirement that $p_t = p^S(x_t)$, the fair response from the perspective of the informed seller. By construction, this sequence has the following properties.

[7]For clarity of presentation, the figure corresponds to the extreme case that $\gamma = 0$ and $\delta = 1$; otherwise the two price lines would be closer together.

1. In the boundary case that $\gamma = \delta = 1$, then $(p, q)_{\gamma,\delta}$ is constant (except in the last period), given by a price of 66 and a quality of 10.

2. If $\gamma < \delta < 1$, then $(p, q)_{\gamma,\delta}$ is such that both the price and the quality are decreasing over time.

3. $(p, q)_{\gamma,\delta}$ is continuous in γ, δ.

Note that $\gamma = \delta = 1$ corresponds to the case of a credulous buyer without disagreement on fair prices. Thus, in this case the only difference to the previous section is the truth-telling commitment of the fair seller.

We define a *semi-separating equilibrium with decreasing cooperation* as follows.
(i) Social players choose their actions as described in Assumptions A2' and A3'.
(ii) Seller types SL and SH both pool with FH, except in the last period.
(iii) The selfish buyer pools with the fair buyer.
(iv) After any history with an initial signal of L, constant prices $p_L(10) = 53$ and qualities 10, the buyer believes that the seller is FL with probability 1.
After any sequence with an initial signal of H such that every buyer price satisfies $p_t = p^B(x_{t-1})$ and every subsequent quality choice satisfies $p^S(x_t) = p_t$, the buyer believes that the seller is one of the three types SH, SL or FH, with relative probabilities given by priors.
In all other cases, the buyer thinks the seller is either type SH or SL for sure.
(vi) After any history where the seller signaled L and the buyer chooses $p_L(10) = 53$ (and the seller chooses 10) in all periods, the seller maintains prior beliefs about the fairness type of the buyer.
After any history where the seller signaled H and the buyer chooses p_1 defined as in Assumption 2'(iii) and $p_t = p^B(x_{t-1})$ for $t \in \{2, ..., 14\}$ the seller maintains prior beliefs about the fairness type of the buyer.
After all other histories, the seller assigns probability 1 to the buyer being selfish.

We now provide sufficient conditions for such an equilibrium to exist.

Proposition 6 *Suppose that Assumptions A2' and A3' hold with $\gamma < \delta < 1$, ϕ sufficiently large and γ sufficiently close to δ. Then the game has a semi-separating equilibrium with decreasing cooperation.*

The proof (see below) captures the following intuition. First consider the behavior of fair types after a high signal. Fair buyers start with the price that she deems adequate for the efficient for the efficient effort level. Fair sellers do not consider this price as sufficient (because $\gamma < \delta$); they thus respond by choosing lower effort levels than efficient. In turn, the buyers do not consider these qualities as adequate to justify the original price. They thus respond by reducing the price even more, thereby inducing a further quality reduction. Thus, a further quality reduction takes place. Arguing this way, quality and price decrease over time. To investigate the pooling incentives of selfish sellers, it is most important to compare their payoffs in the equilibrium with the alternative where they imitate the fair L type. In the latter case, they carry out the efficient action, but obtain only the compensation designed for the L type (53). In the former case, that is, in equilibrium, they start out with a compensation that is higher (between 53 and 66) for an effort that is lower than 10. Thus they are initially better off. Over time, the compensation decreases. However, if γ is sufficiently close to δ, however, the decrease is sufficiently slow that deviation is never worthwhile.

The analysis shows that for a wide range of fairness types, including those where the efficient pooling equilibrium exists under complete information, the semi-separating equilibrium with breakdown of cooperation exists. This provides support for the idea that incomplete information may lead to the breakdown of cooperation.

3 Proofs

3.1 Proof of Proposition 1

Fix a pooling trigger equilibrium corresponding to (p^*, x^*). The expected payoffs in periods after histories that trigger p^* are given by $S(x^*) + 10 - p^*$ for the buyer. The optimal deviation after any such history is to give a zero price, yielding payoffs 10. Condition (2) guarantees that this deviation is not profitable. For the seller, the optimal downward deviation is to choose zero quality levels, resulting in cost savings of $c(x^*, H)$ and a payoff reduction of p^*, discounted by δ. Condition (1) with $\theta = H$ guarantees that the deviation is not profitable for seller H. As $c(x^*, H) > c(x^*, L)$, the corresponding condition for seller L is implied. Upward deviations are clearly not profitable either.
After histories triggering the minimal price, the buyer will respond by choosing 0 in the future independent of the seller's current behavior. Thus exerting a non-minimal quality is costly for the buyer, without any corresponding benefits. The argument after histories triggering minimal quality is analogous.

3.2 Proof of Proposition 4

A selfish player who has seen an action of a player that a fair player would not choose is sure that this player is selfish and thus best responds by minimal actions. Thus, we only consider behavior after actions that are consistent with fair types.

Given the proposed beliefs and actions of the buyers, selfish sellers who imitate the fair types earn payoffs $p_\theta(10) - C(10, \theta)$ in all periods except the last one and $p_\theta(10)$ in the last one. Deviation would lead to payoff 0 in every period. As $p_\theta(10) - C(10, \theta) > 0$, imitation is profitable in all periods.
Given the proposed beliefs and actions of the sellers, selfish buyers who imitate the fair types earn payoffs $100 - p_\theta(10)$ in all periods except the last period. In the last period, they obtain a payoff of $100 - p_\theta(10)$ if the seller is fair and $10 - p_\theta(10)$ if the seller is not. Thus the expected payoff is $90\phi + 10 - p_\theta(10)$ in the last period if there has been no previous deviation. Any deviation of a buyer would reveal her type; thus the best deviation is to a price of zero. This would lead to payoff 10 in every period. Thus, imitation is profitable even in the last period if and only if $\phi > \frac{p_\theta(10)}{90}$.
According to the suggested strategies, all players choose the fair actions (except the selfish sellers in the last period). Thus, Bayes' Law implies that beliefs should correspond to priors after fair actions. Also, any deviation from such actions is off equilibrium, so that any posteriors are consistent with Bayes' Law.[8]

[8]Also, if one argues that selfish sellers literally *cannot* choose anything except their commitment type, then a deviation is indeed fully revealing of the selfish type.

3.3 Proof of Proposition 5

Note first that signals in the quality commitment are pure cheap talk and do not reveal anything about the type.
Given the beliefs of the buyers, selfish sellers who imitate the fair types earn payoffs $p^f(10) - C(10, \theta)$ in all periods except the last one and $p^f(10)$ in the last period. Deviation would lead to payoff 0 in every one. As $p^f(10) - C(10, \theta) > 0$, imitation is profitable in all periods.
Given the beliefs of the sellers, selfish buyers who imitate the fair types earn payoffs $100 - p^f(10)$ in all periods except the last period. In the last period, they obtain a payoff of $100 - p^f(10)$ if the seller is fair and $10 - p^f(10)$ if the seller is not. Thus the expected payoff is $90\phi + 10 - p^f(10)$ in the last period. Any deviation would reveal the type; thus the best deviation is to a price of zero. This would lead to payoff 10 in every period. Imitation is thus profitable even in the last period if and only if $\phi > \frac{p^f(10)}{90}$. A player who has seen an action of a player that a fair player would not choose is sure that this player is selfish and thus best responds by minimal actions.
According to the suggested strategies, all players choose the fair actions (except the selfish sellers in the last period). Thus, Bayes' Law implies that beliefs should correspond to priors after fair actions. Also, any deviation from such actions is off equilibrium, so that any posteriors are consistent with Bayes' Law.

3.4 Proof of Proposition 6

First, consider buyer beliefs. Only seller type FL signals L and chooses 10 thereafter. Thus, it is consistent with Bayes' Law that the buyer believes that the seller is FL with probability 1 after such histories. Similarly, only seller types FH, SH and SL signal H and choose according to $p^S(x_t) = p_t$ thereafter. Thus, it is consistent with Bayes' Law that the buyer believes that the seller is FH, SH or SL with probability 1, with relative probabilities given by priors. All other histories at which the buyer moves are off equilibrium, so that beliefs are not restricted.
Second, consider seller beliefs. As selfish buyers pool with fair buyers, it is Bayes-consistent for them to maintain priors if they have signaled high and then choose according to $p^S(x_t) = p_t$ thereafter and buyers chose prices p_1 as in Assumption 2(iii) and $p_t = p^B(x_t)$ for $t \geq 2$.
Third, consider the behavior of selfish buyers. They earn a positive payoff by pooling, which they cannot by deviating (and thereby revealing their selfish type).
Fourth, consider the behavior of selfish sellers. In period $t = 15$, they obviously earn the maximum possible payoff by choosing the minimal action. Consider any pooling history up to period $t \leq 14$. By continuing to pool thereafter, these sellers attain payoffs of $\sum_{\tau=t+1}^{14} (p_t - C(x_t, \theta)) + p_{15}$ thereafter. By revealing their type through deviation in period $t + 1$, they obtain payoffs of 0 in each future period. As $p_t > C(x_t, \theta)$ by construction, a deviation that reveals the type

is not profitable in any period. However, the seller could instead deviate by signaling L and then imitating type FL. This yields a payoff of $53 - C(10, \theta)$ in each period (except the last one, where the payoff would be 53). As $p_1 > 53$ by construction, the net profit in period 1 after this deviation is never higher than with the proposed equilibrium strategy. If γ converges to δ, disagreements in fairness assessments are small. Hence, p_t falls only very slowly and the deviation profit in any of the remaining periods is not higher than under the proposed equilibrium strategy. Therefore deviation is not profitable. [9]

[9] A more general sufficient (but still not necessary) condition for this equilibrium to emerge is that $p_{15} \geq 53$. It is possible to calculate under which conditions this holds.