



THEMA

théorie économique,  
modélisation et applications

THEMA Working Paper n°2023-10  
CY Cergy Paris Université, France

# **Trade among moral agents with information asymmetries**

José Ignacio Rivero Wildemaue



May 2023

# Trade among moral agents with information asymmetries

José Ignacio Rivero Wildemauwe

May 23, 2023

## Abstract

Two agents trade an item in a simultaneous offer setting, where the exchange takes place if and only if the buyer's bid price weakly exceeds the seller's ask price. Each agent is randomly assigned the buyer or seller role. Both agents are characterized by a certain degree of Kantian morality, whereby they pick their bidding strategy behind a veil of ignorance, taking into account how the outcome would be affected if their trading partner were adopting their strategy. I consider two variants with asymmetric information, respectively allowing buyers to have private information about their valuation or sellers to be privately informed about the item's quality. I show that when all trades are socially desirable, even the slightest degree of morality guarantees that the outcome is fully efficient. In turn, when quality is uncertain and some exchanges are socially undesirable, full efficiency is only achieved with sufficiently high moral standards. Moral concerns also ensure equal ex-ante treatment of the two agents in equilibrium. Finally, I show that if agents are altruistic rather than moral, inefficiencies persist even with a substantial degree of altruism.

**Keywords:** bilateral trade; altruism; *homo moralis*

**JEL Classification:** D03; D82; D91; C78

# 1 Introduction

Asymmetric information constitutes one of the main departures from perfect competition. It may induce coordination breakdowns in trade, leading to mutually beneficial exchanges not taking place. The Myerson and Satterthwaite (1983) theorem is a stark illustration of this issue. While the effects of information asymmetries have been largely studied through models populated by purely self-interested materialistic agents, human motivations are significantly richer.

This paper focuses more specifically on the implications of agents weighing how they themselves would fare if *hypothetically*, other individuals also adopted their own course of action. This kind of behaviour, usually referred to as “Kantian morality”, is potentially very consequential in strategic interactions, and there exists sufficient experimental evidence to justify its serious study.<sup>1</sup> Moreover, the existing literature shows that this sort of moral reasoning helps to remedy market failures in a host of different settings (see Alger and Weibull (2016)).

Here I study the extent to which Kantian moral motivations (or *homo moralis* preferences) *à la* Alger and Weibull (2013) can solve inefficiencies originated in an unequal distribution of information in bilateral trade.<sup>2</sup> In this context, an agent who finds herself in the seller role will take into account her own payoff if she were a buyer and happened to meet a seller who employs her own strategy. Likewise, an individual in the buyer role applies the same reasoning with the roles reversed. Exactly how much agents care about this counter-factual scenario is measured by a parameter henceforth referred to as the *degree of morality*.

In order to capture this precise setting, I assume that agents play a bilateral trading game where one is cast into the buyer role and the other one into the seller role, but they make their decisions behind a *Veil of Ignorance*. That is, without knowing what the final role distribution will be. This in turns implies that each participant needs to choose their seller strategy as well as their buyer strategy before learning their role.

After roles have been assigned, agents play a game where they trade an item in a simultaneous offer setting. The exchange takes place if and only if the buyer’s bid price weakly exceeds the seller’s ask price. Due to the simultaneous-move nature of the game, my framework does not assume asymmetries in bargaining power. It also incorporates information asymmetries

---

<sup>1</sup>Experimental evidence consistent with the presence of Kantian morality is reported in Levine et al. (2020) and Alger and Rivero-Wildemaue (2023). Kantian preferences’ large out-of-sample predictive power is documented in Miettinen et al. (2020) and Alger and Van Leeuwen (2021). For evidence of deontological thinking, see for example Capraro and Rand (2018), Tappin and Capraro (2018), Bursztyrn et al. (2019), Capraro and Vanzo (2019), Bilancini et al. (2020) and Sutter et al. (2020). For the theory grounding these particular moral motivations in evolutionary processes, refer to Alger and Weibull (2013) and Alger et al. (2020).

<sup>2</sup>Theirs is not the only way to consider morality in such a setting. A very relevant alternative is the approach described in Roemer (2010), where the notion of a “Kantian Equilibrium” is proposed. This is defined as a strategy profile where no player would like all players to deviate in the same way.

between buyer and seller. More precisely, I consider two kinds of canonical information asymmetries that have been well studied under the assumption that agents only strive to maximise their own material welfare. The first one is the case where the buyer's valuation for the good is her own private information and the seller can only resort to uniform pricing. The second one is adverse selection. In this scenario, the information asymmetry favours the seller, who knows the quality of her product (which may be high or low), while the buyer can only form an expectation about it.

In the case where the seller is uncertain about the buyer's valuation, the model always produces two kinds of equilibria. The first one is fully efficient as no consumers are excluded from the market. This result is a feature of the one-shot nature of the game I consider. The second one features low-valuation consumers being excluded from the market (as in the standard monopoly model) even though costs are below even the lowest consumer valuation. Introducing moral concerns completely mutes this source of inefficiency, with only the most efficient profiles (those where all consumers are served) surviving as equilibria.

In the adverse selection setting I consider two variants. Firstly, I examine the case where the trade of both qualities produces a net positive surplus. Secondly I tackle the case where trade of the low quality item is socially undesirable, as its cost is higher than the consumer's valuation for it. When trade of both qualities produces a net positive surplus, a sufficiently high expected quality leads to equilibria that feature either no trade, only trade of the low quality or trade of both qualities (the latter being the efficient result). In turn, if expected quality is low, only equilibria with no trade or where high quality sellers are driven out of the market are possible. Introducing even the slightest degree of morality (as long as agents are not completely Kantian) fully restores efficiency, as only equilibria with trade of both qualities are possible in this case.

So far, both the case where the seller is uncertain about the buyer's valuation and the adverse selection setting with desirable trade lead to the same conclusion: going from a game populated by agents only concerned with their own material payoff (*homo oeconomicus*) to the same one but with agents holding moral concerns eliminates all but the efficient equilibria. These surviving equilibria are also symmetric in strategies, producing an *ex-ante* egalitarian result. It can be stated then that introducing *homo moralis* preferences in these contexts (starting from the benchmark self-interested materialistic agents) narrows down the equilibrium set, rather than introducing equilibrium behaviour that would not arise with *homo oeconomicus*. It is worth underscoring that for this narrowing down to occur, all it takes is for agents to be even slightly morally concerned. In other words, the model does not feature a positive threshold degree of morality for the efficient equilibria to be the only outcome possible, as I instead find a stark discontinuity at zero, which implies that even the smallest degree of morality suffices to eliminate all inefficient equilibria.

The situation where trade of the low quality item is undesirable presents important differences with respect to the previous ones. Firstly, the most efficient result, which is having only trade of the high quality item, is not an equilibrium when agents are *homo oeconomicus*.

However, a sufficiently high degree of morality leads to these kind of profiles becoming equilibria and moreover, the only kind possible. Thus, in this case morality does not act as a mere equilibrium selection device. Rather, it allows for the existence of a fully efficient equilibrium and, given a high enough degree of morality, it also selects on it.

Finally, I return to all the model variants described above and compare my results with the case where players are endowed with altruistic preferences *à la* Becker (1976). The general conclusion is that it takes a relatively high degree of altruism to fully restore efficiency even in the settings where all trade is desirable. This result contrasts sharply with the effects of morality, as in most cases partially Kantian agents manage to coordinate on the efficient equilibria regardless of their degree of morality.

This paper is broadly related to the theoretical literature that studies the effects of pro-social preferences and moral motivations on the equilibrium outcomes in a host of strategic interactions (see e.g. Arrow (1973), Becker (1974), Andreoni (1988, 1990), Bernheim (1994), Levine (1998), Fehr and Schmidt (1999), Akerlof and Kranton (2000), Bénabou and Tirole (2006), Alger and Renault (2007), Ellingsen and Johannesson (2008), Englmaier and Wambach (2010), Dufwenberg et al. (2011)). However, the closest references within this literature are those that analyse the effects of moral concerns *à la* Alger and Weibull (2013).

In this line, Alger and Weibull (2017) study several canonical games and compare the outcomes when players have standard, altruistic and moral preferences. Meanwhile, Sarkisian (2017) analyses the optimal contract in a moral hazard problem where a risk-neutral principal hires two risk-averse agents to perform a joint task and may be endowed with either standard, altruistic and moral preferences. Ayoubi and Thurm (2020) rely on Kantian moral concerns to explain why agents may engage in costly pro-environmental behaviour. Muñoz-Sobrado (2022) tackles optimal linear and non-linear taxation problems with moral agents. Finally, Juan-Bartroli and Karagozoglu (2022) study a standard Nash bargaining game where individuals have moral concerns.

The aforementioned references have in common the fact that they deal with symmetric games where information is equally distributed. In contrast, this paper focuses on contexts with different information structures and importantly, where information asymmetries are central. This in turn implies that the nature of the interaction is by definition asymmetric. The *Veil of Ignorance* approach that I propose here allows me to carry out the analysis of this *a priori* asymmetric situation through the lens of *homo moralis* and obtain the first results on how moral concerns interact with asymmetric information.

The rest of the paper is organised as follows. Section 2 presents the general framework and the benchmark bilateral trade game under complete information. Section 3 extends the model to consider the case where buyers are heterogeneous in their valuations but the seller can only set one price. Section 4 analyses the adverse selection case, where the buyer does not know the product's quality. Section 5 revisits the previous models now populated by altruistic agents. Finally, Section 6 provides an overall discussion of my results.

## 2 Framework

In this Section I describe the general setting used throughout the paper and also present the benchmark complete information case. As stated in the Introduction, a simple and tractable way to consider agents capable of universalising their choices in an asymmetric setting is to take an *ex-ante* approach, analysing decisions made behind a “Veil of Ignorance” with respect to the role distribution. In this line, I set up an environment where two identical agents play one of two games, with Nature deciding which game is to be played: either the one that has Player 1 in the first role (say, “seller”) and therefore Player 2 in the second role (“buyer”) or vice-versa.

More formally, consider  $G^1$ , a “contingent game” between players 1 and 2. Player 1 chooses strategy  $s_1 \in S$  and Player 2 chooses strategy  $b_2 \in B$ . They respectively obtain the payoffs  $\pi^s(s_1, b_2)$  and  $\pi^b(s_1, b_2)$ , where  $\pi^s, \pi^b : S \times B \rightarrow \mathbb{R}$ . In addition, take the contingent game  $G^2$ , a game between players 1 and 2 where Player 1 chooses strategy  $b_1 \in B$  and Player 2 chooses strategy  $s_2 \in S$ . They respectively obtain the payoffs  $\pi^b(s_2, b_1)$  and  $\pi^s(s_2, b_1)$ .

Next, define  $G$  as the game where Nature randomly chooses whether  $G^1$  or  $G^2$  is played, with equal probabilities. The *ex-ante* material payoff obtained by Player  $i \in \{1, 2\}$  when using strategy  $(s_i, b_i) \in S \times B$  against Player  $j$ 's strategy  $(s_j, b_j) \in S \times B$  (with  $i \neq j$ ) is:

$$\pi((s_i, b_i), (s_j, b_j)) = \frac{\pi^s(s_i, b_j) + \pi^b(s_j, b_i)}{2}. \quad (1)$$

Finally, define the *utility* that Player  $i$  derives from choosing  $(s_i, b_i) \in S \times B$  against  $(s_j, b_j) \in S \times B$  as:

$$U((s_i, b_i), (s_j, b_j)) = (1 - \kappa)\pi((s_i, b_i), (s_j, b_j)) + \kappa\pi((s_i, b_i), (s_i, b_i)), \quad (2)$$

where  $\kappa \in [0, 1]$  is a parameter representing the player’s “degree of morality”. When choosing a strategy in the game  $G$ , each agent maximizes a convex combination of her own payoff and the one she would obtain if the other player chose the same strategy as her. The term  $\pi((s_i, b_i), (s_i, b_i))$  thus reflects the moral element in the agent’s decision-making process, as it provides an answer to the question “what would my payoff be if the other player behaved as I do?”, while  $\kappa \in [0, 1]$  weighs how much the agent cares about this moral component. The second term in (2) can thus be interpreted as a way to reflect Kant’s categorical imperative, “to act only on the maxim that you would at the same time will to be a universal law” (Kant (1785)), mediated by  $\kappa$ .

At one end of the morality spectrum are the agents with  $\kappa = 0$ , to whom I refer as *homo oeconomicus*. This is the “standard” case of individuals who are merely trying to maximise their own material payoff. On the other end are players with  $\kappa = 1$ , whom (again following Alger and Weibull (2013)) I denominate *homo kantiansis*. These purely Kantian agents only care about their (hypothetical) payoff when the opposing party employs the same strategy as theirs. That is, they only think about what would happen if both players in the game behaved in the same way as they do and decide irrespectively of the actual strategy chosen by the other

player in their game. Finally,  $\kappa \in (0, 1)$  gives way to partially moral agents or *homo moralis*, who put some weight on the morality of their actions while at the same time considering their own material payoff.

The setting I am proposing can be used for the Veil-of-Ignorance analysis of any two-player game and is not limited to bilateral trade or bargaining processes. In order to specifically tackle a given two-player interaction, it suffices to appropriately define  $S, B, \pi^s, \pi^b$  and specify  $\kappa$ . In this paper, contingent game  $G^i$  is a game where a seller is in possession of an object and posts a price for it, while the buyer simultaneously decides which prices she is willing to accept. In this line, it can also be viewed as a *contingent market* where Player  $i$  is the seller and Player  $j$  is the buyer. If the seller's price belongs to the buyer's set of acceptable prices, trade takes place at that price. If it does not, trade does not occur.

## 2.1 Benchmark: Complete information

I now present the complete information game that I will use as a benchmark. The agent assigned to the seller role (say, Player  $i$ ) sets a price  $p_i \in \mathbb{R}_+$ . Simultaneously, the player cast as a buyer (say, Player  $j$ ) chooses a threshold  $\bar{p}_j \in \mathbb{R}_+$  that defines her set of acceptable prices as all those that belong to the interval  $[0, \bar{p}_j]$ .<sup>3</sup> The agent assigned to the seller role faces a cost of  $r \in \mathbb{R}_+$  if they sell the object, while the player in the buyer role derives some consumption utility  $v \in \mathbb{R}_+$  from acquiring the item. If no trade happens, the payoff obtained by each agent is nil.

In terms of the framework laid out at the beginning of the Section, the above describes the contingent game  $G^i$ , where  $s_i$  is  $p_i \in \mathbb{R}_+$  and  $b_j$  is  $\bar{p}_j \in \mathbb{R}_+$ . This means that  $S \times B = \mathbb{R}_+^2$ . The contingent payoffs  $\pi^s(p_i, \bar{p}_j)$  and  $\pi^b(p_i, \bar{p}_j)$  are then defined as:

$$\begin{aligned}\pi^s(p_i, \bar{p}_j) &=: \mathbb{1}\{p_i \leq \bar{p}_j\}(p_i - r), \\ \pi^b(p_i, \bar{p}_j) &=: \mathbb{1}\{p_i \leq \bar{p}_j\}(v - p_i).\end{aligned}$$

Substituting  $\pi^s(p_i, \bar{p}_j)$  and  $\pi^b(p_i, \bar{p}_j)$  in expressions (1) and (2) we obtain the *ex-ante* payoff and utility, respectively:

$$\pi((p_i, \bar{p}_i); (p_j, \bar{p}_j)) = \frac{1}{2} [\mathbb{1}\{p_i \leq \bar{p}_j\}(p_i - r) + \mathbb{1}\{p_j \leq \bar{p}_i\}(v - p_j)], \quad (3)$$

$$\begin{aligned}U((p_i, \bar{p}_i); (p_j, \bar{p}_j)) &= (1 - \kappa) \cdot \frac{1}{2} [\mathbb{1}\{p_i \leq \bar{p}_j\}(p_i - r) + \mathbb{1}\{p_j \leq \bar{p}_i\}(v - p_j)] + \\ &\quad \kappa \cdot \frac{1}{2} \cdot \mathbb{1}\{p_i \leq \bar{p}_i\}(v - r)\end{aligned} \quad (4)$$

---

<sup>3</sup>Throughout the paper, I restrict attention to strategy spaces where buyers' sets of acceptable prices are defined by thresholds. I provide a generalisation in the Appendix.

Notice that the term inside the parenthesis in the last line of Expression (4) is the total net trade surplus. I consider a materially efficient outcome (or, simply put, an efficient outcome) as the one that maximises this surplus. As a result, the efficient equilibria in the complete information game are those that induce trade in both contingent markets.

My main concern throughout the paper is efficiency (although I do study the equity properties of the equilibrium outcomes whenever relevant). Notice that when considering a given strategy profile of the bilateral trade game, in each contingent market it can either generate trade (if the price in that market is no larger than the threshold) or not. As a consequence, we can classify each profile into a type  $A/B$  where  $A$  indicates whether there is trade in the first contingent market and  $B$  in the second one. For the complete information game studied in this Section, profiles can be of four types: Trade/Trade, Trade/No Trade, No Trade/Trade or No Trade/No Trade.

I begin with the version of the model where  $\kappa = 0$  (so players only care about their own material payoff), which I use as a benchmark. In order to find the game's Nash equilibria in pure strategies for this case, I start by noticing that, when deciding on her price if in the seller role, the only way this action affects her payoff is through its interaction with the other player's threshold. Likewise, her action as buyer only influences her payoff through its interaction with the other player's price. It then follows that the equilibrium set of the bilateral trade game is comprised of all the ordered pairs formed by an equilibrium in  $G^1$  and an equilibrium in  $G^2$ . Proposition 2.1 summarises them.<sup>4</sup>

**Proposition 2.1.** *Consider the complete information bilateral trade game with  $\kappa = 0$ . The Nash equilibria in pure strategies are:*

- *Trade/Trade:  $\bar{p}_i^* = p_j^* \in [r, v]$  for all  $i \neq j \in \{1, 2\}$*
- *Trade/No Trade or No Trade/Trade:  $\bar{p}_i^* = p_j^* \in [r, v], \bar{p}_j^* < r, p_i^* > v$  for all  $i \neq j \in \{1, 2\}$*
- *No Trade/No Trade:  $\bar{p}_i^* < r, p_j^* > v$  for all  $i \neq j \in \{1, 2\}$*

The complete information game between two *homo oeconomicus* has three types of equilibrium profiles. The first one is the Trade/Trade result, where exchanges take place in both contingent markets at prices between seller cost and consumer valuation. The second kind are No Trade/No Trade equilibria, where there is no trade in either contingent market. Here, thresholds are set below seller costs and prices are above consumer valuation. The third sort of equilibria are those with contingent trade, where there is trade in one of the contingent markets (at a price between cost and valuation) but not in the other (where the threshold is set below cost and the price is above consumer valuation).

---

<sup>4</sup>All proofs are relegated to the Appendix.



Recall that I assume throughout that  $r < v$ , so a materially efficient outcome in this model is for trade to actually happen. We see then that the second and third types of equilibrium profiles (those with no trade or with contingent trade) constitute a sort of coordination failure, a negotiation breakdown that is brought about by the simultaneity in the agents' decisions. In this regard, we could say that the particular set-up of the game is "stacked" against the occurrence of socially beneficial exchanges. The following sections focus on additional inefficiencies stemming from asymmetric information and explore whether they are mitigated in the presence of moral concerns.

Having established this benchmark, I explore whether considering partially moral agents (for whom  $\kappa \in (0, 1)$ ) modifies this result.<sup>5</sup> The following Proposition lays out the bilateral trade game's equilibrium set with *homo moralis* agents.

**Proposition 2.2.** *Consider the complete information bilateral trade game with  $\kappa \in (0, 1)$ . The Nash equilibria in pure strategies are  $r \leq p_i^* = \bar{p}_j^* = p_j^* = \bar{p}_i^* \leq v, \forall i, j \in \{1, 2\}, i \neq j$ .*

Proposition 2.2 states that only symmetric, full-trade profiles with prices in the  $[r, v]$  interval are Nash equilibria in pure strategies of the bilateral trade game. This equilibrium set is a subset of the equilibria resulting from the game between *homo oeconomicus* agents.

Partial morality fully restores efficiency in the complete information set-up, as it effectively precludes any no-trade or contingent-trade outcome from taking place. Even more strikingly, in order for the conclusion to go through, it is sufficient for only one of the players,  $i$  to be endowed with *homo moralis* preferences. In fact, the attainment of full efficiency does not depend on the intensity of players' moral concerns as long as they are not fully Kantian.

It is worth further reflecting on the implications of Proposition 2.2 by contrasting the equilibria to the version of the model with *homo oeconomicus*. Recall that those profiles that constitute equilibria for the latter but not for *homo moralis* are those that feature no trade in at least one contingent market, and those with trade in both but at different prices. The intuition is fairly simple. Notice that these profiles necessarily imply that at least one of the players is setting a threshold higher than her own price. While this is completely irrelevant in the *homo oeconomicus* case, *homo moralis* agents will find deviations from these strategies morally profitable, as they care about setting a price that they themselves would be willing to accept. Moreover, they can do this either without facing material losses or with only infinites-

---

<sup>5</sup>The approach of solving the contingent game to find the game's equilibrium set is not applicable in this case. This is due to the fact that when setting her price, each agent needs to take not only her opponent's threshold into consideration, but also her own. As a consequence, I take the alternative and admittedly more cumbersome route of analysing each equilibrium type (no-trade, full trade and contingent trade), and discarding those profiles for which profitable deviations are available (see Lemmas A.1 to A.6 in the Appendix). Following this, I show that the remaining candidate equilibria are indeed mutual best responses and characterise them (see Lemma A.7).

imally large ones (thus trading off a strictly positive moral gain against an arbitrarily small material loss in this latter case).

A final interesting remark is that due to the symmetry of the bilateral trade game under analysis, a symmetric profile will result in the same *ex-ante* utility. So in this way, partial moral concerns lead to complete *ex-ante* equality. Before the role distribution is revealed, each player can expect to obtain the same level of utility.

It is reasonable to wonder whether moral motivations produce the same effects on bilateral trade's efficiency when inefficiencies result from information asymmetries between the buyer and the seller. This is a pervasive issue that ranks amongst the most studied explanations for why markets may fail to deliver fully efficient results. In particular, the case where the seller is not completely certain of the buyer's willingness to pay for the item is formalised as early as in Pigou (1920)'s classic analysis of monopoly with uniform pricing. In turn, the situation where buyers do not know the quality of the object being bargained over, which could either be low or high was brought forth notably by Akerlof (1970). The following sections extend the complete information model described here to embed those two cases in the bargaining framework.

### 3 Asymmetric information: Heterogeneous buyer valuations

In order to analyse the valuation uncertainty case, I re-define the contingent game  $G^i$ , introducing a move by Nature where it randomly decides the buyer's valuation. This can be high ( $v_h \in \mathbb{R}_+$ ) with probability  $\lambda$  or low ( $v_l \in \mathbb{R}_+$ ) with probability  $1 - \lambda$ . The seller's cost is  $r \in \mathbb{R}_+$ , as in Section 2.1, and I assume that  $0 < r < v_l < v_h$  (so the efficient outcome is for both types of consumers to be served). The agent in the buyer role is informed of her valuation, while the seller only knows the probability  $\lambda$ . As a consequence, the former now sets two thresholds, while the latter sets only one price (as in Section 2.1). Contingent game  $G^i$  thus represents the standard monopoly problem with uniform pricing, with the difference that price and thresholds are decided simultaneously.

Strategy  $s_i$  is defined as  $p_i \in \mathbb{R}_+$  (as in Section 2.1), while  $b_j$  is defined as  $(\bar{p}_{jh}, \bar{p}_{jl}) \in \mathbb{R}_+^2$ . Meanwhile, the contingent game's payoffs are re-defined as:

$$\begin{aligned}\pi^s(p_i, \bar{p}_{jh}, \bar{p}_{jl}) &=: \lambda \mathbb{1}\{p_i \leq \bar{p}_{jh}\}(p_i - r) + (1 - \lambda) \mathbb{1}\{p_i \leq \bar{p}_{jl}\}(p_i - r) \\ \pi^b(p_i, \bar{p}_{jh}, \bar{p}_{jl}) &=: \lambda \mathbb{1}\{p_i \leq \bar{p}_{jh}\}(v_h - p_i) + (1 - \lambda) \mathbb{1}\{p_i \leq \bar{p}_{jl}\}(v_l - p_i).\end{aligned}\tag{5}$$

Substituting the contingent payoffs in Expression (1), we obtain the *ex-ante* utility:

$$\begin{aligned}U((p_i, \bar{p}_{ih}, \bar{p}_{il}); (p_j, \bar{p}_{jh}, \bar{p}_{jl})) &= (1 - \kappa) \cdot \frac{1}{2} [\lambda \mathbb{1}\{p_i \leq \bar{p}_{jh}\}(p_i - r) + (1 - \lambda) \mathbb{1}\{p_i \leq \bar{p}_{jl}\}(p_i - r) + \\ &\quad \lambda \mathbb{1}\{p_j \leq \bar{p}_{ih}\}(v_h - p_j) + (1 - \lambda) \mathbb{1}\{p_j \leq \bar{p}_{il}\}(v_l - p_j)] + \\ &\quad \kappa \cdot \frac{1}{2} [\lambda \mathbb{1}\{p_i \leq \bar{p}_{ih}\}(v_h - r) + (1 - \lambda) \mathbb{1}\{p_i \leq \bar{p}_{il}\}(v_l - r)]\end{aligned}\tag{6}$$

The first two lines of Expression (6) state that the seller's payoff in the contingent game is the weighted average of the surplus obtained by trading with either type of buyer, which

is either the difference between price and cost or zero (if that particular consumer type is not served). The weights are each consumer type's probability. In turn, the buyer's payoff is the weighted average of the surplus that she would get if she was a high valuation buyer and that corresponding to being a low valuation one. Meanwhile, the third line (which represents the moral term) is just half the expected total surplus obtained in game  $G$ .

Having two types of consumers (high and low-valuation) requires more cumbersome expressions to characterise equilibrium prices and thresholds. As a consequence, in Proposition 3.2 I only state whether trade takes place and with which type of consumer, relegating the full characterisation to Proposition 3.2\* in the Appendix. Definition 3.1 describes every kind of outcome that can arise in either contingent market. I then use this classification to state the game's equilibrium types.

**Definition 3.1.** *Consider the contingent market where  $i$  is in the seller role and  $j \neq i$  is in the buyer role. The contingent game's outcomes can be of the following types:*

- *Full trade: both types of buyers are served. If and only if  $p_i \leq \min\{\bar{p}_{jh}, \bar{p}_{jl}\}$*
- *High valuation: only high valuation buyers are served. If and only if  $\bar{p}_{jl} < p_i \leq \bar{p}_{jh}$*
- *Low valuation: only low valuation buyers are served. If and only if  $\bar{p}_{jh} < p_i \leq \bar{p}_{jl}$*
- *No trade: no buyers are served. If and only if  $p_i > \max\{\bar{p}_{jh}, \bar{p}_{jl}\}$*

**Proposition 3.2.** *Consider the heterogeneous buyer valuations bilateral trade game with  $\kappa = 0$ . The Nash equilibria in pure strategies are of the type A/B.  $A, B \in \{\text{Full trade, High val., No trade}\}$ .*

For a given contingent market, two of the three possible outcomes are familiar results in monopoly settings with uniform pricing. The first one is when all consumer types are being served ("Full trade"), while the second one has low-valuation consumers excluded from the market ("High valuation"). Recall that the latter is inefficient, as even the low consumer valuation is above seller cost. Unsurprisingly, the range of prices that support equilibria with exclusion is increasing in the probability of the consumer being of high valuation  $\lambda$ . The third kind is the negotiation failure result ("No trade"), a consequence of the simultaneity of decisions assumed in my framework that we had already encountered in the previous section. The bilateral trade game's equilibria are all the pairs that can be formed from this set.

Introducing buyer heterogeneity in this otherwise unmodified context then naturally brings about a new source of inefficiency: asymmetric information. More specifically, the fact that the seller is unsure about the consumer's willingness to pay for the item makes it possible for exclusion equilibria to exist. They occur when the maximum prices accepted by the low and high valuation consumers are different enough so that, given the share of valuation consumers, the seller is better off excluding.

Considering partially moral agents drastically changes the types of equilibria attainable. I state it in the following Proposition, relegating the full characterisation of prices and thresholds to the Appendix, as before.<sup>6</sup>

**Proposition 3.3.** *Consider the heterogeneous buyer valuations bilateral trade game with  $\kappa \in (0, 1)$ . The only kind of Nash equilibria in pure strategies is Full trade/Full trade.*

As in Section 2.1, these profiles are a subset of the equilibria in the bilateral trading game with *homo oeconomicus* agents. Similarly to the case with complete information, partial moral concerns eliminate inefficient equilibria originated in the coordination failure brought about by the one-shot assumption. But the extension analysed here shows that this is not the only effect they have. In fact, they also do away with inefficiencies stemming from the information asymmetries, as they also preclude exclusion. Importantly, the precise degree of morality is still not relevant as long as it remains strictly between 0 and 1. Therefore, we again find a marked discontinuity in the effects of introducing morality in the agents' decision making process. Going from  $\kappa = 0$  to a strictly positive degree of morality, however small (and as long as it does not equal 1) suffices to do away with every inefficient equilibrium in this case as well.

## 4 Asymmetric information: Adverse selection

In contrast to the previous Section, I study here a common values setting where the information asymmetry concerns both costs and valuation. More precisely, I consider a bilateral trading game where the seller may be in possession of either a high quality object or a low quality one. This is echoed in the consumer's valuation, which is  $v_h$  for the high quality and  $v_l$  for the low one. Importantly, I also assume that high-quality sellers face a higher cost than low-quality ones, namely  $r_h$  and  $r_l$ , with  $r_h > r_l$ .

While the seller knows her item's quality, the buyer only knows the probability distribution. To reflect this in the definition of contingent game  $G^i$ , I assume that Nature first randomly chooses whether the object is of high or low quality with probabilities  $\lambda$  and  $1 - \lambda$ . The seller is then informed of the result of this draw, while the buyer only knows  $\lambda$ . In order to capture this new information structure in a tractable way, we must then have the seller posting two prices, one for when she is of the low quality type and one for when she is of the high quality kind. In turn, the agent in the buyer role posts a threshold above which she will not trade (just as in Section 2.1). All of this means that strategy  $s_i$  is now defined as  $(p_{ih}, p_{il}) \in \mathbb{R}_+^2$ , and  $b_j$  as

---

<sup>6</sup>For the sake of completeness, I also present in the Appendix, in Lemma B.11 the equilibria with fully moral agents. The takeaway is the same as in the complete information case.

$(\bar{p}_j) \in \mathbb{R}_+$ . The contingent game's payoffs take the form:

$$\begin{aligned}\pi^s(p_{ih}, p_{il}, \bar{p}_j) &=: \lambda \mathbb{1}\{p_{ih} \leq \bar{p}_j\}(p_{ih} - r_h) + (1 - \lambda) \mathbb{1}\{p_{il} \leq \bar{p}_j\}(p_{il} - r_l) \\ \pi^b(p_{ih}, p_{il}, \bar{p}_j) &=: \lambda \mathbb{1}\{p_{ih} \leq \bar{p}_j\}(v_h - p_{ih}) + (1 - \lambda) \mathbb{1}\{p_{il} \leq \bar{p}_j\}(v_l - p_{il}).\end{aligned}$$

Substituting these contingent payoffs in (2) gives the utility function for the quality uncertainty version of the bilateral trading game:

$$\begin{aligned}U((p_{ih}, p_{il}, \bar{p}_i); (p_{jh}, p_{jl}, \bar{p}_j)) &= (1 - \kappa) \cdot \frac{1}{2} [\lambda \mathbb{1}\{p_i \leq \bar{p}_{jh}\}(p_i - r) + (1 - \lambda) \mathbb{1}\{p_i \leq \bar{p}_{jl}\}(p_i - r) + \\ &\quad \lambda \mathbb{1}\{p_j \leq \bar{p}_{ih}\}(v_h - p_j) + (1 - \lambda) \mathbb{1}\{p_j \leq \bar{p}_{il}\}(v_l - p_j)] + \\ &\quad \kappa \cdot \frac{1}{2} [\lambda \mathbb{1}\{p_i \leq \bar{p}_{ih}\}(v_h - r) + (1 - \lambda) \mathbb{1}\{p_i \leq \bar{p}_{il}\}(v_l - r)]\end{aligned}\quad (7)$$

In line with the previous sections, the first line in Expression (7) represent the expected material payoff that an agent  $i$  choosing  $(p_{ih}, p_{il}, \bar{p}_i) \in \mathbb{R}_+^3$  pitted against an individual who plays  $(p_{jh}, p_{jl}, \bar{p}_j) \in \mathbb{R}_+^3, j \neq i$  would obtain when in the role of a seller. It is the weighted average surplus that she would get if in the possession of a high quality object and the surplus corresponding to having a low quality item. Again, the weights are the probabilities of getting the high quality object or the low quality one. If  $j$  acquires the good (that is, if  $p_{iQ} \leq \bar{p}_j$ , with  $Q \in \{h, l\}$ ), she receives the price she set and loses an amount equivalent to the cost of the good,  $r_Q$ . Her total surplus from selling is then  $p_{iQ} - r_Q$ . In turn, if the aforementioned exchange does not occur, the payoff is zero.

The second line is the expected material surplus that the agent would obtain when in the role of a buyer. It is the weighted average of the surplus that she would get if paired with a high quality seller and that corresponding to being matched with a low quality seller. The weights are the probabilities of the object being of high quality ( $\lambda$ ) or of low quality ( $1 - \lambda$ ). In each case, Player  $i$ 's surplus corresponding to her potential buyer role is equivalent to her valuation  $v_Q$  for the good minus the price  $p_{jQ}$  she paid if it is the case that Player  $j$  sells her the object (which happens only if  $p_{jQ} \leq \bar{p}_i$ ), while she obtains 0 if not.

As in the previous sections, the equilibrium set of the bilateral trading game with quality uncertainty and *homo oeconomicus* agents can be found by studying the contingent game between a buyer and a seller. I next define all the possible types of result that can arise in the contingent game in Definition 4.1.

**Definition 4.1.** *Consider the contingent market where  $i$  is in the seller role and  $j \neq i$  is in the buyer role. The contingent game's outcomes can be of the following types:*

- *Full trade: both qualities are traded. If and only if  $\max\{p_{ih}, p_{il}\} \leq \bar{p}_j$*
- *Low quality: only low quality is traded. If and only if  $p_{il} \leq \bar{p}_j < p_{ih}$*
- *High quality: only high quality is traded. If and only if  $p_{ih} \leq \bar{p}_j < p_{il}$*
- *No trade: no quality is traded. If and only if  $\bar{p}_j < \min\{p_{ih}, p_{il}\}$*

The equilibria are all pairs of profiles that can be formed with the elements belonging to the contingent game's equilibrium set. In order to make the reading more efficient, I first define two relevant quantities that will be used in this and the following sections. Firstly, the expected consumer valuation, which is the average of the high and low quality valuations weighed by their respective probabilities ( $\lambda$  and  $1 - \lambda$ ).

$$v_e \equiv \lambda v_h + (1 - \lambda)v_l. \quad (8)$$

Secondly, the minimum value of  $\lambda$  for which the expected consumer valuation is weakly above the high-quality cost ( $r_h \leq v_e$ ):

$$\lambda_e \equiv \frac{r_h - v_l}{v_h - v_l}. \quad (9)$$

I study here two cases of these adverse selection setting. The first one is when not only the exchange of the high quality object but also of the low quality item produces a positive surplus and thus, is socially desirable. The second one is when the trade of the low quality good results in a negative surplus and it is then efficient from a social standpoint to avoid it exchanging hands.

## 4.1 Desirable trade

I present in Proposition 4.2 the equilibria for the case where all trade is socially desirable and agents are of the *homo oeconomicus* type. The contingent market outcomes listed are those defined in Definition 4.1.

**Proposition 4.2.** *Consider the adverse selection bilateral trade game with  $\kappa = 0$  and assume  $0 < r_l < v_l < r_h < v_h$ . The Nash equilibria in pure strategies are:*

- *If  $\lambda \geq \lambda_e$ ,  $A/B$  where  $A, B \in \{Full\ trade, Low\ quality, No\ trade\}$ .*
- *If  $\lambda < \lambda_e$ ,  $A/B$  where  $A, B \in \{Low\ quality, No\ trade\}$*

When the expected buyer valuation  $v_e$  is (weakly) above the high quality cost  $r_h$  or, equivalently,  $\lambda \geq \lambda_e$ , there are six types of equilibria. The first kind are those where trade takes place no matter the role distribution or quality of the goods. Next, there are equilibria where Player  $i$  sells to Player  $j$  only when in possession of a bad quality item, but the latter sells when her good is either quality. The following class of equilibria has Player  $i$  not selling to Player  $j$  and the latter selling to Player  $i$  no matter the quality. The fourth kind of equilibria are those where only trade of the bad quality item takes place in both contingent markets. The game also features equilibria where Player  $i$  does not sell to Player  $j$  but the latter sells when her good is of low quality. The last class of equilibria are those where trade does not take place at all. Notice that in this case, inefficient outcomes arise because of the coordination failure introduced by the simultaneous-action assumption, mirroring the complete information case.

In turn, when  $\lambda < \lambda_e$ , the equilibria laid out in Proposition 4.2 are of three types. The first one is that where only the bad quality good is traded in both contingent markets. The second sort is the case where the bad quality good is traded in one contingent market but the other one features no trade. Finally, there are equilibria where no trade occurs in either contingent market. The expected consumer valuation  $v_e$  being below high quality cost  $r_h$  implies that the asymmetric information problem is much more harmful for market efficiency, as the high quality item is never traded, even though it is socially desirable to do so. This situation resembles the “lemons” problem, as high quality sellers are driven out of the market and, even when the coordination failure is avoided (and trade takes place) only bad quality sellers remain.

While considering valuation uncertainty and *homo oeconomicus* players brought about the possibility of equilibria where low-valuation consumers are not served (like in the traditional monopoly pricing problem), the introduction of quality uncertainty adds a different potential inefficiency, that of adverse selection (where only the bad quality item is traded). In fact, for sufficiently low expectations about product quality, the most efficient profiles (those featuring full trade) do not even belong to the equilibrium set. This renders the adverse selection setting more problematic than the heterogeneous buyers case in terms of efficiency, as with *homo oeconomicus* agents the efficient equilibria are not even attainable for sufficiently low values of the  $\lambda$ .

I now explore whether considering partially moral agents does away with the various inefficient equilibria found in the *homo oeconomicus* version of the model. Proposition 4.3 presents the bilateral trade game’s equilibrium types when agents are *homo moralis*. The detailed characterisation is relegated to Proposition 4.3\* in the Appendix.

**Proposition 4.3.** *Consider the adverse selection bilateral trade game with  $\kappa \in (0, 1)$  and assume  $0 < r_l < v_l < r_h < v_h$ . If  $\lambda \geq \lambda_e$ , the only Nash equilibria in pure strategies are of the type Full trade/Full trade. If  $\lambda < \lambda_e$ , there are no Nash equilibria in pure strategies.*

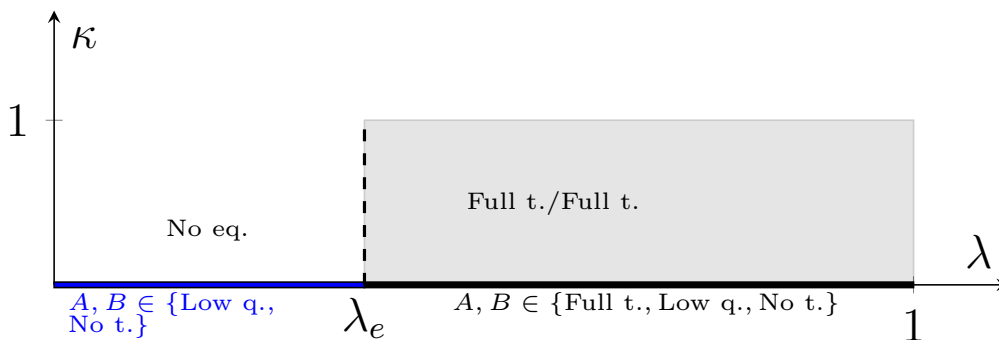


Figure 1: Equilibrium regions in the  $(\lambda, \kappa)$  space, with  $\kappa \in [0, 1)$ .

Proposition 4.3 states that when the probability of the item being of good quality is large enough (or, equivalently, the high-quality cost  $r_h$  is weakly below the expected consumption valuation  $v_e$ ) the only Nash equilibria in pure strategies are symmetric and feature full trade in both contingent markets. Furthermore, trade takes place at prices between the high-quality

cost  $r_h$  and the expected valuation  $v_e$ . Notice again, that these profiles are a subset of the equilibria in the *homo oeconomicus* model.

Hence, in the case where agents are sufficiently “optimistic” about quality (so that  $\lambda \geq \lambda_e$ ), going from a bilateral trade game where agents are endowed with *homo oeconomicus* preferences to one where they are *homo moralis* effectively eliminates all equilibria but the most materially efficient ones. In this context, these are the ones where both qualities are traded in both contingent markets. Thus, (partial) moral concerns only prevent negotiation failures, as was already clear from the benchmark complete information model. Moreover and also in keeping with the findings for the previous variants of the model (complete information and heterogeneous buyer valuations), *homo moralis* preferences result in a completely *ex-ante* egalitarian result, as trade takes place in both contingent markets at the same price.

The intuition behind the result that profiles with no trade of either quality in at least one contingent market are not equilibria when agents are *homo moralis* is the same as for the full information benchmark. Starting from a situation with none or partial trade, agents can adjust their strategies so as not to modify the actual trade taking place but managing to set prices and thresholds such that the latter would not be above the former. Given that trade of both qualities is socially desirable, these deviations produce moral gains that are not offset by material losses and are thus profitable.

Secondly and also in line with the results for the full information case, asymmetric full trade profiles cannot be equilibria either. This is due to the fact that in a asymmetric profile with full trade in both contingent markets, there is at least one agent setting at least one price above her threshold. This provides an opportunity to make a moral gain and avoid material losses by simultaneously adjusting prices and thresholds.

Finally, it is worth noting that when the the probability of having a high quality item is not high enough ( $\lambda < \lambda_e$ ), there exist no Nash equilibria in pure strategies. This derives from the fact that *homo moralis* will always find it convenient to deviate from a profile with full trade. The reason is that such a profile features at least one contingent seller trading at a price below cost or one contingent buyer paying above her expected valuation, so there are material gains to be had from deviating. Agents will also be able to deviate from those featuring contingent or no trade through moral gains.

Moreover, even if an equilibrium in mixed strategies were to exist, a small degree of morality would not restore efficiency, as there would always be some probability that the threshold is below some price in some contingent game.<sup>7</sup> In fact, this assertion holds for  $\kappa$  close to 1 (although in that case we might expect that the probability of no trade becomes very small).

---

<sup>7</sup>The existence of a mixed-strategy equilibrium would be guaranteed by Nash’s theorem if we consider a finite grid of prices and thresholds. Furthermore, if the grid is between and including costs and valuations for each quality and if it is fine enough, then all the pure strategy equilibria described in the paper remain so.



## 4.2 Socially undesirable trade

In all of the above, I have assumed that costs are always lower than consumer valuations and thus that trading always produces a social surplus. In this Section I modify the model laid out in Section 4 in order to reflect a situation where it would be socially inefficient to trade the low quality good. In this case, the sellers' valuation for the good ( $r_l$ ) is actually higher than the buyers' ( $v_l$ ). In line with this, I assume now that  $0 < v_l < r_l < r_h < v_h$ . Proposition 4.4 characterises the equilibria for the benchmark *homo oeconomicus* case. Recall that quantities  $v_e$  and  $\lambda_e$  represent, respectively, the expected consumer valuation and the value of the probability of high quality above which the former is larger than the high quality cost  $r_h$ . They are defined in expressions (8) and (9).

**Proposition 4.4.** *Consider the adverse selection bilateral trade game with  $\kappa = 0$  and assume that  $0 < v_l < r_l < r_h < v_h$ . Then:*

1. *If  $\lambda \geq \lambda_e$ , the Nash equilibria in pure strategies of the bargaining game between two homo oeconomicus are of the kind A/B where  $A, B \in \{\text{Full trade, No trade}\}$*
2. *If  $\lambda < \lambda_e$ , the Nash equilibria in pure strategies of the bargaining game are of the type No trade/No trade*

When the low quality object's cost is higher than its consumer valuation, the most efficient result is that only the high quality item be traded. However, Proposition 4.4 shows that this type of profiles are not equilibria when agents are *homo oeconomicus*. In fact, in a given contingent market, either both qualities are traded or non are, with the former equilibria only existing when  $\lambda \geq \lambda_e$  (that is, the probability of the object being of high quality is large enough). This constitutes a relevant difference with the versions of the model analysed so far, as the most efficient profiles possible were in fact part of the equilibrium set when  $\lambda \geq \lambda_e$ . In addition, the case where  $\lambda < \lambda_e$  leads to equilibria with no trade at all, whereas in the previous case it could also lead to exclusion of the high quality in addition to no-trade. How does the equilibrium set change when agents are of the *homo moralis* type? Proposition 4.5 provides the answer to that question.

In order to facilitate the reading, I define the following thresholds for  $\lambda$  and  $\kappa$ :

$$\lambda_1 := \frac{r_h - r_l}{v_h - r_l}$$

$$\kappa_1 := \frac{r_h - r_l}{r_h - v_l}$$

$$\kappa_2(\lambda) := \min \left\{ \frac{\lambda(v_h - v_l) + v_l - r_l}{\lambda(v_h - r_l) + r_l - v_l}, \frac{\lambda(v_h - v_l) + v_l - r_h}{\lambda(v_h - r_l) + r_l - r_h} \right\}, \text{ for } \lambda \in [\lambda_e, 1]$$

Notice that  $\lambda_1 < \lambda_e$ . In addition,  $\kappa_2(\lambda)$  is continuous and strictly increasing with  $\kappa_2(\lambda_e) = 0$  and  $\kappa_2(1) \in (\kappa_1, 1)$ .

**Proposition 4.5.** Consider the adverse selection bilateral trade game with  $\kappa \in (0, 1)$  and assume that  $0 < v_l < r_l < r_h < v_h$ . The only kind of pure strategy Nash equilibria are:

1. High quality/High quality: if and only if  $\lambda \geq \lambda_1, \kappa \geq \kappa_1$ .
2. Full trade/Full trade: if and only if  $\lambda \geq \lambda_e, \kappa \leq \kappa_2(\lambda)$ .
3. High quality/Full trade or Full trade/High quality: if and only if  $\kappa \in [\kappa_1, \kappa_2(\lambda)]$

Figure 2 summarises propositions 4.4 and 4.5. The first relevant point to be made about Proposition 4.5 is that it now takes a sufficiently high degree of morality to do away with all but the most efficient equilibria. This contrasts sharply with the conclusions for the preceding sections, where only efficient equilibria are attained for *any*  $\kappa \in (0, 1)$ . Even though a requirement on expected quality persists, sufficiently moral agents require a lower probability of good quality to reach these equilibria. In the same line, the existence of pure strategy Nash equilibria with partially moral agents now requires a lower expectation of good quality than in the previous case, provided that agents are moral ( $\kappa \geq \kappa_1$ ) enough, since  $\lambda_1 < \lambda_e$ .

Considering partially moral agents brings about additional interesting changes to the equilibrium set of the game under study. In contrast with the previous cases where all trade was socially desirable, under the present parametrisation, *homo moralis* agents reach equilibria that are not attainable by *homo oeconomicus*. In other words, so far the introduction of partially moral concerns has had the effect of “refining” the equilibrium sets in the game between by *homo oeconomicus* players, as every equilibrium profile with *homo moralis* belonged to the former. This is not the case now because any profile with trade of exclusively the good quality in either or both contingent markets does not constitute an equilibrium with *homo oeconomicus*.

This kind of equilibrium is particularly relevant, as only surplus-generating exchanges are undertaken. They exist when the agents’ degree of morality and expectation about product quality are sufficiently high (i.e., the former is above  $\kappa_1$  while the latter is above  $\lambda_1$ ). The simple intuition behind such a result is that, given that trade is occurring at a price above high-quality cost  $r_h$  and thus above low-quality cost  $r_l$ , only sufficiently moral agents will not be tempted to reduce their low quality price to sell their low quality item. It is worth pointing out that the minimum probability for the object to be of good quality required for the existence of this type of equilibria is strictly below that required for any trade to take place between *homo oeconomicus*, or in other words ( $\lambda_1 < \lambda_e$ ).

The second type of equilibrium found in the *homo moralis* model is Full trade/Full trade. This is the only class already present in the *homo oeconomicus* case. It exists when the agents’ degree of morality is not “too high” (below  $\kappa_2(\lambda)$ ) and their expected consumer valuation is above the high quality cost, just as the *homo oeconomicus* case. The fact that they exist only for  $\lambda \geq \lambda_e$  is also line with the result in the previous Section, where I assumed that all trade is socially desirable. The difference is that now there is an upper bound on the degree of morality ( $\kappa_2(\lambda)$ ) for which the equilibria exist. This is because high degrees of morality now cause deviations that impede trade of the low quality to become profitable.

The third type are those with only the good quality being exchanged in one contingent market and full trade in the other one. The minimum degree of morality for which they exist is the same as the High quality/High quality equilibria, but its maximum is also bounded from above by  $\kappa_2(\lambda)$ , exactly like Full trade/Full trade profiles. This type of equilibria exist in the intersection of the sets of  $\lambda$  and  $\kappa$  values where the first two types are possible. That means that for a given  $(\lambda, \kappa)$  vector that supports High quality/Full trade equilibria, High quality/High quality and Full trade/Full trade also exist. As reported in Proposition 4.5\* in the Appendix, the price at which exchanges take place is exactly the same in both contingent markets and moreover, it is unique and increasing in the degree of morality. Intuitively, this is because more moral players who are selling both items will require a larger price not to deviate and cease the trade of the bad quality object. Simultaneously, for those selling only the good quality, they will be enticed to cease bad quality trade if the price is sufficiently small. The maximum price below which these deviations are profitable is increasing in  $\kappa$ .

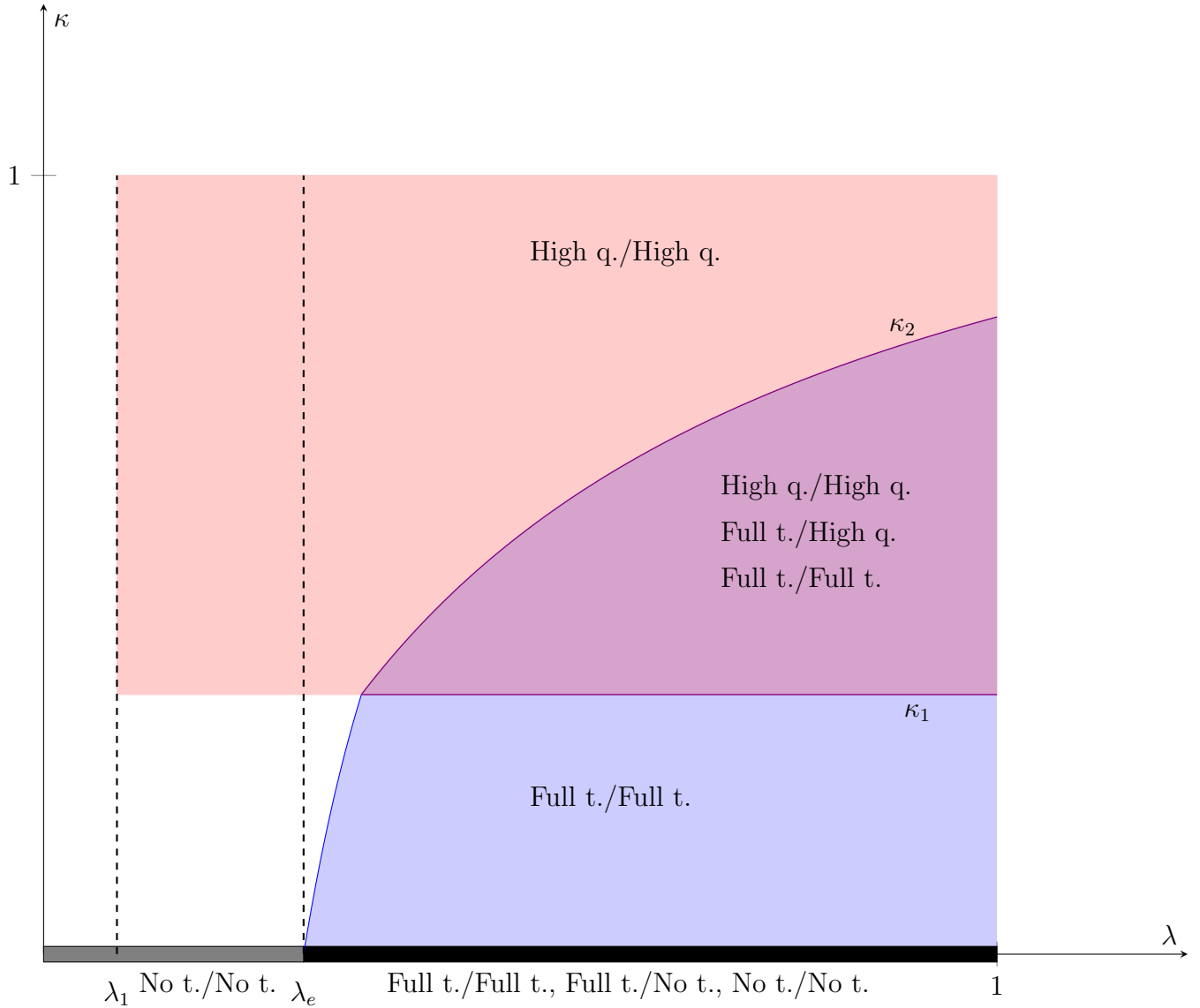


Figure 2: Equilibrium regions in the  $(\lambda, \kappa)$  space. Shaded areas indicate types of equilibria for  $\kappa \in [0, 1)$ .

## 5 A comparison with altruism

I have shown in the previous sections that (partially) Kantian agents manage to prevent negotiation failures and significantly reduce or even eliminate the inefficiencies stemming from information asymmetries. Cases where all potential exchanges produce a positive surplus feature particularly strong effects of morality on market outcomes: it suffices for agents to be partially moral for all but the most efficient profiles to cease being equilibria.

In this section I modify the previous models to consider agents with altruistic preferences. I describe as “altruistic” an individual who cares not only about her own material welfare but also about that of others. Both moralists and altruists deviate from the standard assumption that agents only care for their material payoff. Moreover, they may be behaviourally identical

in certain contexts (see Alger and Weibull (2016)). However, the rationale for their observed choices is radically different. While the former care about the hypothetical material payoff they would get if their strategy is imitated by the other agent, altruists care about the material payoffs induced by their own strategy and their opponent's *actual* strategy. Confusing morally-driven players and altruists could then lead to very different predictions when studying the reaction of agents to changes in the environment.

An altruistic agent is partially concerned by the other player's material payoff. The utility function presented in (2) when the *ex-ante* game was defined needs to be modified accordingly. A standard way of introducing altruism is:

$$U^{alt}((s_i, b_i), (s_j, b_j)) = \pi((s_i, b_i), (s_j, b_j)) + \alpha \cdot \pi((s_j, b_j), (s_i, b_i)), \quad (10)$$

where  $\alpha \in [0, 1]$  represents the degree of altruism. Equation (10) states that the utility of an individual playing strategy  $(s_i, b_i)$  against a player who employs strategy  $(s_j, b_j)$  depends on her material payoff but is also positively affected by the material payoff obtained by her opposing party ( $\pi((s_j, b_j), (s_i, b_i))$  represents the payoff obtained by the player employing strategy  $(s_j, b_j)$  against  $(s_i, b_i)$ ). The more altruistic this player is (that is, the larger  $\alpha$  is), the more utility she derives from her opponent's material well-being.<sup>8</sup>

In the remainder of this Section I revisit the three games studied so far but I assume that agents are altruistic rather than morally concerned. The main conclusion is that in every setting considered, efficient results are attained only when altruism is high enough, and even a very high degree of altruism may not lead to full efficiency in some cases. In addition and echoing the *homo oeconomicus* case, equilibria in the game between altruistic agents can be asymmetric. This implies that situations where one player has a larger *ex-ante* utility

In Proposition 5.1 I present the equilibria of the complete information game introduced in Section 2.1 when agents have altruistic preferences.

**Proposition 5.1.** *Consider the complete information game with  $\alpha \in (0, 1)$ . Then:*

- *If  $\alpha \leq \frac{r}{v}$ , the Nash equilibria in pure strategies are of the type A/B where  $A, B \in \{\text{Trade}, \text{No trade}\}$ .*
- *If  $\alpha > \frac{r}{v}$ , the Nash equilibria in pure strategies are of the type Trade/Trade*

When comparing the set of equilibria in the complete information game between altruistic agents with the *homo oeconomicus* version, it is immediately clear that they are qualitatively similar for low levels of altruism ( $\alpha \leq \frac{r}{v}$ ). In turn, for high levels of altruism the only possible

---

<sup>8</sup>Although quite conventional, this representation it is not the only one. See for example Bruhin et al. (2018) for an expression which allows for different degrees of altruism depending on whether the agent has a higher or a lower payoff than her opponent.

equilibria are those of the Trade/Trade type. Therefore, a first conclusion is that only sufficiently high altruism restores efficiency in this game. An additional interesting point is that nothing prevents two equally altruistic agents to trade at different prices in the two contingent markets, a feature that the model with *homo oeconomicus* also shares. What drives this result is the deontological nature of moral decision-making and more precisely, the fact that a player's decisions in a given contingent market are linked to her decisions in the other market through the moral Kantian term. This effect is not present when agents are consequentialistic (be it *homo oeconomicus* or altruistic).

The above effects are a feature of the different versions of the bargaining game. However, considering information asymmetries brings about other interesting differences in the outcomes of the bargaining process between altruists when compared to the benchmark *homo oeconomicus* and to *homo moralis*. I now turn to the buyer heterogeneity game with altruistic agents. Proposition 5.2 lays out the equilibrium types for this version of the model.<sup>9</sup>

**Proposition 5.2.** *Consider the buyer heterogeneity game between two agents with the same degree of altruism  $\alpha \in (0, 1)$ . Then:*

1. *If  $\alpha \leq \frac{r}{v_h}$ , the Nash equilibria in pure strategies are of types A/B where  $A, B \in \{\text{Full trade, High valuation, No trade}\}$ .*
2. *If  $\frac{r}{v_h} > \alpha$ , the Nash equilibria in pure strategies are of types A/B where  $A, B \in \{\text{Full trade, High valuation}\}$ .*

Proposition 5.2 shows that in the heterogeneous buyers setting, altruism does not restore full efficiency, as equilibria with exclusion of low valuation consumers are always possible for any  $\alpha \in (0, 1)$ . Intuitively, this result originates in the fact that the altruistic seller actually perceives a lower cost when trading with a high valuation consumer.

I now turn to the adverse selection game in the case where it is socially desirable to trade both qualities. I define function  $\lambda_{e1}^{alt}(\alpha)$  to be used in the propositions that follow.

$$\lambda_{e1}^{alt}(\alpha) = \begin{cases} \frac{r_h - v_l + \alpha(r_l - v_h)}{v_h - v_l + \alpha(r_l - r_h)} & \text{if } \alpha \leq \frac{r_h - v_l}{v_h - r_l} \\ 0 & \text{otherwise} \end{cases} \quad (11)$$

For the case where  $v_l < r_l$ , I additionally define  $\lambda_{e2}^{alt}(\alpha)$ :

$$\lambda_{e2}^{alt}(\alpha) = \begin{cases} \frac{r_l - v_l + \alpha(r_l - v_l)}{v_h - v_l + \alpha(r_l - r_h)} & \text{if } \alpha \leq \frac{v_h - r_l}{r_h - v_l} \\ 1 & \text{otherwise} \end{cases} \quad (12)$$

---

<sup>9</sup>The full characterisation is relegated to the Appendix because comparing prices and thresholds yields the same conclusion as in the complete information game.

Notice that  $\lambda_{e1}^{alt}(\alpha)$  is continuous and decreasing in  $\alpha$ . Moreover,  $\lambda_{e1}^{alt}(0) = \lambda_e$ , while  $\lambda_{e1}^{alt}\left(\frac{r_h - v_l}{v_h - r_l}\right) = 0$ . In turn, function  $\lambda_{e2}^{alt}(\alpha)$  is continuous and increasing in  $\alpha$  for  $v_l < r_l$ , with  $\lambda_{e2}^{alt}(0) = \frac{r_l - v_l}{v_h - v_l}$ .

Proposition 5.3 presents the game's equilibrium types for different degrees of altruism when low quality trade is socially desirable ( $r_l < v_l$ ), while Proposition 5.4 does the same for the opposite case ( $v_l < r_l$ ).

**Proposition 5.3.** *Consider the adverse selection game between two agents with the same degree of altruism  $\alpha \in (0, 1)$  and assume  $0 < r_l < v_l < r_h < v_h$ . Then, the Nash equilibria are of type  $A/B$ , where  $A, B \in T(\alpha, \lambda)$  with:*

- Full trade  $\in T(\alpha, \lambda)$  if and only if  $\lambda \geq \lambda_{e1}^{alt}(\alpha)$
- Low quality  $\in T(\alpha, \lambda)$  if and only if  $\alpha < \min\left\{\frac{r_h}{v_h}, \frac{r_h - r_l}{v_h - v_l}\right\}$
- High quality  $\in T(\alpha, \lambda)$  if and only if  $\frac{r_h - r_l}{v_h - v_l} < \alpha < \frac{r_l}{v_l}$
- No trade  $\in T(\alpha, \lambda)$  if and only if  $\alpha < \min\left\{\frac{r_h}{v_h}, \frac{r_l}{v_l}\right\}$

**Proposition 5.4.** *Consider the adverse selection game between two agents with the same degree of altruism  $\alpha \in (0, 1)$  and assume  $0 < v_l < r_l < r_h < v_h$ . Then:*

- Full trade  $\in T(\alpha, \lambda) \iff \lambda \geq \max\{\lambda_{e1}^{alt}(\alpha), \lambda_{e2}^{alt}(\alpha)\}$
- High quality  $\in T(\alpha, \lambda) \iff \alpha > \frac{r_h - r_l}{v_h - v_l}$
- No trade  $\in T(\alpha, \lambda) \iff \alpha < \frac{r_h}{v_h}$

When the exchange of all qualities is socially desirable, a sufficiently high degree of altruism does achieve full efficiency, eliminating all equilibria but those featuring full trade in both contingent markets. In contrast, the fully efficient result when there is some socially undesirable trade is not guaranteed by a very high degree of altruism. This is due to the fact that even though a sufficiently large  $\alpha$  is indeed a necessary and sufficient condition for the existence of equilibria featuring only trade of the high quality good, it is not sufficient to eliminate equilibria with full trade in some contingent market.

## 6 Discussion

In this paper I explore a simple bilateral trading process between two agents who may have moral concerns of a Kantian nature. In it, the agent in possession of the good being bargained over proposes a price, while the potential buyer simultaneously sets a threshold. The good only gets exchanged (at the proposed price) if the threshold is at least as high as the price. Importantly, I take an *ex-ante* approach, analysing the game between two agents who face the same odds of finding themselves in either role. This framework is well suited to study the interaction between two agents who in addition to wondering “what is the strategy that would

leave me materially better-off?”, may also pose themselves the counter-factual question: “how would I fare if my counter-party used the same strategy I am considering?”

My set-up does not rule out coordination failures and is flexible enough to allow for the study of the interaction between different kinds of information structures and moral preferences. I focus on two kinds of canonical information asymmetries. The first one originates when the seller is uncertain about the buyer’s valuation for the product, while the second one is the well known adverse selection problem where the buyer is not informed about the product’s quality but the seller is. I tackle two versions of the latter: one where trade is always socially beneficial irrespective of the quality, and another where only trade of the highest quality is desirable.

I show that considering moral preferences has important consequences on the outcome of the market under study. In situations where all trade is desirable (and irrespective of whether there is complete information, valuation uncertainty or adverse selection) and the version of the game populated by agents who are only concerned with their material payoff has efficient equilibria in addition to inefficient ones, partial morality eliminates the latter, leaving only the most efficient equilibria. It also results in a completely *ex-ante* egalitarian outcome, in the sense that expected utility levels are the same.

In the case with adverse selection and undesirable trade of the low quality items, when agents are of the *homo oeconomicus* type, the most efficient equilibria (trade of only high quality goods) are not part of the equilibrium set. It takes a sufficiently high degree of morality for these profiles to become equilibria and an even larger one for them to be the sole kind of equilibrium possible.

The above contrasts sharply with another kind of pro-social preferences that have received a lot of attention in the literature: altruistic preferences. Indeed, I show that altruism is much less capable of eliminating inefficiencies stemming from information asymmetries. In particular, with altruistic agents, efficient equilibria remain as the sole equilibrium type only when information is equally distributed or in the case of adverse selection with desirable trade and, moreover, this is true only when the degree of altruism is high enough.

My results show that although moral preferences may go a long way in making the outcome of the bilateral trade process more efficient, they are not all-powerful. Notably, the most damaging incarnation of the adverse selection problem in this context (that in which trading low quality items actually reduces total surplus) is completely solved only for very high degrees of morality.

The latter notwithstanding, the inefficiencies introduced by the coordination problem as well as those originating in the information problem are strongly mitigated by considering moral preferences. This suggests that these kind of constraints might be less harmful than what standard models predict, especially in contexts well suited for the kind of partially deontological approach to decision-making that characterises *homo moralis*.



## References

- Akerlof, G. A. (1970). The market for “lemons”: Quality uncertainty and the market mechanism. *The Quarterly Journal of Economics* 84(3), 488–500.
- Alger, I. and J. Weibull (2016). Morality: evolutionary foundations and policy implications. TSE Working Papers 16-702, Toulouse School of Economics (TSE).
- Alger, I. and J. W. Weibull (2013). Homo moralis: Preference evolution under incomplete information and assortative matching. *Econometrica* 81(6), 2269–2302.
- Alger, I. and J. W. Weibull (2017). Strategic behavior of moralists and altruists. *Games* 8(3).
- Alger, I., J. W. Weibull, and L. Lehmann (2020). Evolution of preferences in structured populations: Genes, guns, and culture. *Journal of Economic Theory* 185, 104951.
- Becker, G. S. (1976). Altruism, egoism, and genetic fitness: Economics and sociobiology. *Journal of Economic Literature* 14(3), 817–826.
- Bilancini, E., L. Boncinelli, V. Capraro, T. Celadin, and R. Di Paolo (2020). ” do the right thing” for whom? an experiment on ingroup favouritism, group assorting and moral suasion. *arXiv preprint arXiv:2002.12464*.
- Bruhin, A., E. Fehr, and D. Schunk (2018, 05). The many Faces of Human Sociality: Uncovering the Distribution and Stability of Social Preferences. *Journal of the European Economic Association* 17(4), 1025–1069.
- Bursztyn, L., S. Fiorin, D. Gottlieb, and M. Kanz (2019). Moral incentives in credit card debt repayment: Evidence from a field experiment. *Journal of Political Economy* 127(4), 1641–1683.
- Capraro, V. and D. G. Rand (2018). Do the right thing: Experimental evidence that preferences for moral behavior, rather than equity or efficiency per se, drive human prosociality. *Judgment and Decision Making* 13(1), 99–111.
- Capraro, V. and A. Vanzo (2019). The power of moral words: Loaded language generates framing effects in the extreme dictator game. *Judgment and Decision Making* 14(3), 309–317.
- Kant, I. (1785). *Immanuel Kant: Groundwork of the Metaphysics of Morals: A German–English edition*. The Cambridge Kant German-English Edition. Cambridge University Press.
- Levine, S., M. Kleiman-Weiner, L. Schulz, J. Tenenbaum, and F. Cushman (2020). The logic of universalization guides moral judgment. *Proceedings of the National Academy of Sciences* 117(42), 26158–26169.

- Miettinen, T., M. Kosfeld, E. Fehr, and J. Weibull (2020). Revealed preferences in a sequential prisoners' dilemma: A horse-race between six utility functions. *Journal of Economic Behavior & Organization* 173, 1–25.
- Myerson, R. B. and M. A. Satterthwaite (1983). Efficient mechanisms for bilateral trading. *Journal of Economic Theory* 29(2), 265 – 281.
- Pigou, A. C. (1920). *The economics of welfare / by A. C. Pigou*. Macmillan London.
- Roemer, J. E. (2010). Kantian equilibrium. *The Scandinavian Journal of Economics* 112(1), 1–24.
- Sarkisian, R. (2017). Team incentives under moral and altruistic preferences: Which team to choose? *Games* 8(3).
- Smith, A. (1759). *Adam Smith: The Theory of Moral Sentiments*. Cambridge Texts in the History of Philosophy. Cambridge University Press. 2002.
- Sutter, M., J. Huber, M. Kirchler, M. Stefan, and M. Walzl (2020). Where to look for the morals in markets? *Experimental Economics* 23(1), 30–52.
- Tappin, B. M. and V. Capraro (2018). Doing good vs. avoiding bad in prosocial choice: A refined test and extension of the morality preference hypothesis. *Journal of Experimental Social Psychology* 79, 64 – 70.

## Appendix A Appendix to Section 2.1

### Proof of Proposition 2.1

The price set by a given player is completely independent of her threshold and only affects her utility through its interaction with her rival's threshold. Thus, the bargaining game's equilibrium set is all the pairs of profiles that can be formed from the equilibria of the contingent game.

Consider now the contingent game. The seller's best reply correspondences are:

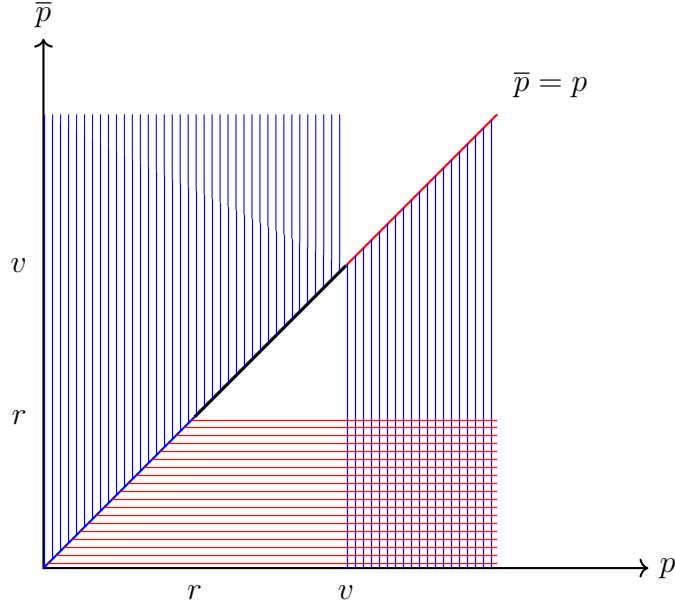
$$\bar{p}^{BR} = \operatorname{argmax}_{\bar{p}} U^B(\bar{p}; p) = \begin{cases} \bar{p} \geq p & \text{if } p \leq v \\ \bar{p} < p & \text{if } p > v \end{cases}$$

$$p^{BR} = \operatorname{argmax}_p U^S(p; \bar{p}) = \begin{cases} p = \bar{p} & \text{if } \bar{p} \geq r \\ p > \bar{p} & \text{if } \bar{p} < r \end{cases}$$

These correspondences intersect whenever  $\bar{p}^* = p^* \in [r, v]$  or  $\bar{p}^* < r, p^* > v$ . □

The best-reply correspondences in the contingent game can be presented graphically in the  $(p, \bar{p}) \in \mathbb{R}_+^2$  plane. I do so in Figure 3. Notice that the contingent game has two types of distinct equilibrium profiles. The first are equilibria with trade, where the buyer sets a threshold in the  $[r, v]$  interval and the seller proposes a price equal to the buyer's threshold. The second are no-trade equilibria, with the seller choosing a price above the the buyer's valuation  $v$ , who in turn sets a threshold below the seller's cost  $r$ .

Figure 3: Buyer and seller best-responses



**Proof of Proposition 2.2** The proof follows directly from Lemmas A.1 to A.7, presented below.

**Lemma A.1.** *Consider the bilateral trade game with complete information and  $\kappa \in (0, 1)$ . Profiles where at least one contingent market features no trade and its price is set below consumer valuation or the threshold above seller cost are not Nash equilibria in pure strategies.*

*Proof.* Assume, without loss of generality, that  $\bar{p}_1^* < p_2^* \leq v$ . Player 1 can profitably deviate by increasing her threshold to  $\bar{p}'_1 = p_2^*$ . The first term of her utility function increases because  $\bar{p}'_1 = p_2^* \leq v$ . The second term either increases (if  $p_2^* \geq p_1^*$ ) or remains the same (if  $p_1^* \leq \bar{p}_1^*$  or  $p_2^* < \bar{p}_1^*$ ).

In turn, suppose  $r \leq \bar{p}_2^* < p_1^*$ . Player 1 can profitably deviate by reducing  $p_1^*$  down to  $p'_1 = \bar{p}_2^*$ . The first term of her utility function increases because  $p'_1 = \bar{p}_2^* \geq r$ . The second term either increases (if  $\bar{p}_1^* \in [\bar{p}_2^*, p_1^*)$ ) or remains the same (if  $\bar{p}_1^* < \bar{p}_2^*$  or  $\bar{p}_1^* \geq p_1^*$ ).  $\square$

**Lemma A.2.** *Consider the bilateral trade game with complete information and  $\kappa \in (0, 1)$ . No-trade/No-trade profiles with prices above valuation and thresholds below costs are not Nash equilibria in pure strategies.*

*Proof.* These profiles are described by  $p_i^* > v, \bar{p}_j^* < r$ ; for all  $i, j \in \{1, 2\}, i \neq j$ . Consider, without loss of generality, Player 1. She can profitably deviate by setting  $\bar{p}'_1 \in (\bar{p}_2^*, p_2^*)$  and

$p'_1 \in (\bar{p}_2^*, p'_1]$ . The first term of her utility function remains unchanged while the second one increases.  $\square$

**Lemma A.3.** *Consider the bilateral trade game with complete information and  $\kappa \in (0, 1)$ . Profiles featuring trade in at least one contingent market at a price below cost or above valuation are not Nash equilibria in pure strategies.*

*Proof.* The first kind of profiles have (1)  $p_i^* \leq \bar{p}_j^* < r$  for some  $i, j \in \{1, 2\}, i \neq j$ . The second, (2)  $v < p_j^* \leq \bar{p}_i^*$  for some  $i, j \in \{1, 2\}, i \neq j$ .

(1): Assume without loss of generality that  $p_1^* \leq \bar{p}_2^* < r$ . Then, Player 1 can profitably deviate to  $p'_1 = r, \bar{p}'_1 = \max\{r, \bar{p}_1^*\}$ . The first term of her utility function increases while the second one does not decrease.

(2): Assume without loss of generality that  $v < p_2^* \leq \bar{p}_1^*$ . Then, Player 1 can profitably deviate to  $\bar{p}'_1 = p'_1 = p_2^* - \epsilon$ . For sufficiently small  $\epsilon$ , the first term of her utility function strictly increases while the second one does not decrease.  $\square$

**Lemma A.4.** *Consider the bilateral trade game with complete information and  $\kappa \in (0, 1)$ . Profiles featuring trade at a price between cost and valuation in one contingent market and no trade with price above valuation and threshold below cost in the other one, are not Nash equilibria in pure strategies.*

*Proof.* These profiles are described by  $\bar{p}_i^* < r, p_i^* \leq \bar{p}_j^* \in [r, v]$  and  $p_j^* > v; i, j \in \{1, 2\}, i \neq j$ . Assume without loss of generality that  $\bar{p}_1^* < r, p_1^* \leq \bar{p}_2^*, p_1^* \in [r, v]$  and  $p_2^* > v$ . Consider Player 1 and notice that  $\bar{p}_1^* < p_1^*$ . She can profitably deviate to  $\bar{p}'_1 = p_1^*$ . The first term of her utility function remains unchanged while the second one increases.  $\square$

**Corollary A.4.1.** *Profiles featuring no trade in at least one contingent market are not equilibria of the bilateral trade game where at least one agent's degree of morality is  $\kappa \in (0, 1)$ .*

**Lemma A.5.** *Consider the bilateral trade game with complete information and  $\kappa \in (0, 1)$ . Profiles featuring full trade where the threshold differs from the price in at least one contingent market are not Nash equilibria of the game between two homo moralis.*

*Proof.* Assume without loss of generality,  $p_1^* < \bar{p}_2^*$ . Then, Player 1 can profitably deviate to  $p'_1 = \bar{p}_2^*, \bar{p}'_1 = \max\{\bar{p}_1^*, \bar{p}_2^*\}$ . The first term of her utility function increases while the second one does not decrease.  $\square$

**Lemma A.6.** *Consider the bilateral trade game with complete information and  $\kappa \in (0, 1)$ . Profiles featuring full trade but at different prices are not Nash equilibria of the game between two homo moralis.*

*Proof.* by Lemma A.5, we can focus on profiles where thresholds equal price in both contingent markets. Assume, without loss of generality, that  $p_2^* = \bar{p}_1^* < p_1^* = \bar{p}_2^*$ . Then, Player 1 can profitably deviate to  $\bar{p}'_1 = p_1^*$ . The first term of her utility function does not decrease while the second one increases.  $\square$

**Lemma A.7.** *Consider the bilateral trade game with complete information and  $\kappa \in (0, 1)$ . Of all symmetric profiles with full trade only those with coordinates inside the  $[r, v]$  interval are Nash equilibria in pure strategies.*

*Proof.* (1) Consider a profile where  $p^* = p_1^* = \bar{p}_2^* = p_2^* = \bar{p}_1^* < r$ . Any such profile is not an equilibrium because (without loss of generality) Player 1 can profitably deviate to  $p'_1 = \bar{p}'_1 > p^*$ . The first term of her utility function increases while the second one remains the same.

(2) Consider next a profile where  $v < p^* = p_1^* = \bar{p}_2^* = \bar{p}_1^* = p_2^*$ . Then, Player 1 can profitably deviate to  $\bar{p}'_1 = \bar{p}'_1 = p^* - \epsilon$ . For small enough  $\epsilon$ , the first term of her utility function increases. The second term remains the same.

(3) Take a profile where  $r \leq p^* = p_1^* = \bar{p}_2^* = \bar{p}_1^* = p_2^* \leq v$ . Consider, without loss of generality, Player 1. Given Player 2's strategy, both terms of her utility function are at their maximum. Therefore, there are no profitable deviations possible. Then, the profile is an equilibrium.  $\square$

**Proposition A.8.** *Consider the bilateral trade game with complete information and  $\kappa = 1$ . The set of Nash equilibria in pure strategies is  $p_i^* \leq \bar{p}_i^* \forall i \in \{1, 2\}$*

*Proof.* Given that  $r < v$ ,  $i$ 's utility function with  $\kappa = 1$  is maximised if and only if  $p_i^* \leq \bar{p}_i^*$ .  $\square$

Agents with degree of morality  $\kappa$  equal to 1 (*homo kantiansis*) will only care about her own strategy in order to compute their utility. That is, they will look for a best response against their own strategy and completely disregard what the other player might do.

Under the parametrisation assumed, the surplus generated by trade is always positive. Therefore, agents are then better off when they propose a price and a threshold such that trade would take place if the other agent were to (hypothetically) post exactly the same ones. Recall, however, that the occurrence or not of trade in this model still depends on whether Player 1's threshold is at least as high as the price proposed by Player 2 and vice-versa.

A game between two *homo kantiansis* will thus feature equilibria of two types. The first kind are full trade equilibria, while in the second one only contingent trade will be attained.

These equilibria will always feature some trade, either contingent or full. This can be easily shown by noticing that it is always be true that either  $\bar{p}_2^* \leq \bar{p}_1^*$  or  $\bar{p}_2^* > \bar{p}_1^*$ . In the first case, this implies that  $p_2^* \leq \bar{p}_1^*$  (in equilibrium it is necessarily true that  $p_2^* \leq \bar{p}_2^* \leq \bar{p}_1^*$ ). This means that Player 2 is selling to Player 1. If, in addition, it is also true that  $\bar{p}_2^* \geq p_1^*$ , then these profiles feature full trade, but this is not a necessary condition for equilibrium.

When comparing this game with the one between two *homo oeconomicus*, an immediate conclusion is that the coordination failure introduced by the simultaneity in price and threshold decisions is mitigated, as no-trade equilibria are no longer possible. Notwithstanding, full morality does not always produce the more materially efficient result. This is because the game may still feature equilibria where trade is only contingent. These are a direct effect of agents only caring about their own threshold and prices while completely disregarding those of their rival. The result thus points to the fact that partially consequentialistic reasoning keeps *homo*

*moralis* “anchored” in reality by actually paying attention to her opponent’s choices and thus producing actual trade.

The existence of completely deontological agents who only care about what would be “the right thing to do” might seem a little far fetched. However, the above result serves to illustrate a broader point already made in the preceding literature (notably in Alger and Weibull (2017)): a higher degree of morality does not always imply a materially more efficient outcome, and whether it does is very context-dependent. In this particular case, the inefficiency arises because *homo kantiensis* loses her grip on reality in the sense that she is not concerned about whether actual trade occurs, for which it is necessary that players set thresholds at least as high as their rival’s prices. In contrast, she is content with just setting a price that she herself would accept, no matter how large or small.

## Appendix B Appendix to Section 3

**Proposition 3.2\*.** *Consider the bilateral trade game with valuation uncertainty and  $\kappa = 0$ . The pure strategy Nash equilibria are:*

- *Full trade/Full trade:*  $\forall i, j \in \{1, 2\}, i \neq j : r \leq p_i^* = \min\{\bar{p}_{jh}^*, \bar{p}_{jl}^*\} \leq v_l, \bar{p}_{ih}^* \in [(1 - \lambda) p_{il}^* + \lambda r, (\frac{1}{\lambda}) p_{il}^* - (\frac{1-\lambda}{\lambda}) r]$
- *A/B with  $A, B \in \{Full\ trade, High\ valuation\}$ ,  $A \neq B$ :* for  $i, j \in \{1, 2\}, i \neq j$ :
  - $r \leq p_i^* = \min\{\bar{p}_{jh}^*, \bar{p}_{jl}^*\} \leq v_l, \bar{p}_{ih}^* \in [(1 - \lambda) p_{il}^* + \lambda r, (\frac{1}{\lambda}) p_{il}^* - (\frac{1-\lambda}{\lambda}) r]$ ;
  - $v_l < p_j^* = \bar{p}_{ih}^* \leq v_h, 0 \leq \bar{p}_i^* \leq \lambda \bar{p}_{ih}^* + (1 - \lambda)r$
- *A/B with  $A, B \in \{Full\ trade, No\ trade\}$ ,  $A \neq B$ :* for  $i, j \in \{1, 2\}, i \neq j$ :
  - $r \leq p_i^* = \min\{\bar{p}_{hj}^*, \bar{p}_{ij}^*\} \leq v_l, \bar{p}_{hi}^* \in [(1 - \lambda) \bar{p}_{hi}^* + \lambda r, (\frac{1}{\lambda}) \bar{p}_{hi}^* - (\frac{1-\lambda}{\lambda}) r]$ ;
  - $\max\{\bar{p}_{li}^*, \bar{p}_{hi}^*\} < r, p_j^* > v_h$
- *A/B with  $A, B \in \{High\ valuation, No\ trade\}$ ,  $A \neq B$ :* for  $i, j \in \{1, 2\}, i \neq j$ :
  - $\max\{\bar{p}_{lj}^*, \bar{p}_{jh}^*\} < r, p_i^* > v_h$
  - $v_l < p_j^* = \bar{p}_{hi}^* < v_h, 0 \leq \bar{p}_i^* \leq \lambda \bar{p}_{hi}^* + (1 - \lambda)r$
- *High valuation/High valuation:*  $\forall i, j \in \{1, 2\}, i \neq j : v_l < p_j^* = \bar{p}_{hi}^* < v_h, 0 \leq \bar{p}_i^* \leq \lambda \bar{p}_{hi}^* + (1 - \lambda)r$
- *No-trade in both contingent markets:*  $\forall i, j \in \{1, 2\}, i \neq j : \max\{\bar{p}_{lj}^*, \bar{p}_{hj}^*\} < r, p_i^* > v_h$

*Proof.* The price set by a given player is completely independent of her thresholds and only affects her utility through its interaction with her rival's thresholds. Thus, the bilateral trade game's equilibrium set is all the pairs of profiles that can be formed from the equilibria of the contingent game.

Consider now the contingent game where the seller sets price  $p$  and the buyer sets thresholds  $\bar{p}_l$  and  $\bar{p}_h$ . The buyer's best reply correspondences in the contingent game are:

$$\bar{p}_{BR}^l = \operatorname{argmax}_{\bar{p}^l} U^B(\bar{p}^l, \bar{p}^h; p) = \begin{cases} \bar{p}_{BR}^h = p & \text{if } p \leq v_h \\ \bar{p}_{BR}^l = p & \text{if } p > v_h \end{cases}$$

$$\bar{p}_{BR}^h = \operatorname{argmax}_{\bar{p}^h} U^B(\bar{p}^l, \bar{p}^h; p) = \begin{cases} \bar{p}_{BR}^l = p & \text{if } p \leq v_l \\ \bar{p}_{BR}^h = p & \text{if } p > v_l \end{cases}$$

The seller's best reply correspondence in the contingent game is:

$$\begin{cases} p^{BR} > \max\{\bar{p}^h, \bar{p}^l\} & \text{if } \max\{\bar{p}^h, \bar{p}^l\} < r \\ p^{BR} = \max\{\bar{p}^h, \bar{p}^l\} & \text{if } \max\{\bar{p}^h, \bar{p}^l\} \geq r, \min\{\bar{p}^h, \bar{p}^l\} < r \\ p^{BR} = \min\{\bar{p}^h, \bar{p}^l\} & \text{if } \bar{p}^l \geq r, (1-\lambda)\bar{p}^l + \lambda r \leq \bar{p}^h \leq \frac{1}{\lambda}\bar{p}^l - \frac{1-\lambda}{\lambda}r \\ p^{BR} = \bar{p}^l & \text{if } \bar{p}^l \geq r, \bar{p}^h < (1-\lambda)\bar{p}^l + \lambda r \\ p^{BR} = \bar{p}^h & \text{if } \bar{p}^l \geq r, \bar{p}^h > \frac{1}{\lambda}\bar{p}^l - \frac{1-\lambda}{\lambda}r \end{cases}$$

They intersect at the following profiles:

- Both consumer types are served:  $r \leq p^* = \min\{\bar{p}_l, \bar{p}_h\} \leq v_l; \bar{p}_l^* \leq \bar{p}_h^* \leq \frac{\bar{p}_l^*}{\lambda} - \frac{1-\lambda}{\lambda}r$  or  $\frac{\bar{p}_h^* - \lambda r}{1-\lambda} \leq \bar{p}_h^* \leq \bar{p}_l^*$
- Only high valuation consumers are served:  $v_l < p^* = \bar{p}_h^* \leq v_h, \bar{p}_l^* \leq \lambda \bar{p}_h^* + (1-\lambda)r$
- No consumers are served:  $\max\{\bar{p}_l^*, \bar{p}_h^*\} < r, p^* > v_h$

□

**Proposition 3.3\*.** *Consider the bilateral trade game with heterogeneous buyer valuations and  $\kappa \in (0, 1)$ . The Nash equilibria in pure strategies are  $\bar{p}_{hi}^* \in [(1-\lambda)p_{li}^* + \lambda r, (\frac{1}{\lambda})p_{li}^* - (\frac{1-\lambda}{\lambda})r] \forall i, j \in \{1, 2\}, i \neq j$ .*

*Proof.* The proof follows directly from Lemmas B.1 to B.10, presented below. □

**Lemma B.1.** *Consider the bilateral trade game with heterogeneous buyer valuations and  $\kappa \in (0, 1)$ . Profiles where at least one of the contingent markets features no-trade with the high (low) valuation consumer and price below her valuation are not Nash equilibria in pure strategies.*

*Proof.* Assume, without loss of generality, that  $\bar{p}_{1Q}^* < p_2^* \leq v_Q$  for some  $Q \in \{h, l\}$ . Then, Player 1 can profitably deviate by setting  $\bar{p}'_{1Q} = p_2^*$ . Her utility function's first term increases while the second one does not decrease. □

**Lemma B.2.** *Consider the bilateral trade game with heterogeneous buyer valuations and  $\kappa \in (0, 1)$ . Profiles where at least one of the contingent markets features no-trade and a threshold above cost are not Nash equilibria in pure strategies.*

*Proof.* Assume, without loss of generality, that  $r \leq \max\{\bar{p}_{h2}^*, \bar{p}_{l2}^*\} < p_1^*$ . Then, Player 1 can profitably deviate by setting  $p_1' = \max\{\bar{p}_{h2}^*, \bar{p}_{l2}^*\}$ . Her utility function's first term increases while the second one does not decrease.  $\square$

**Lemma B.3.** *Consider the bilateral trade game with heterogeneous buyer valuations and  $\kappa \in (0, 1)$ . Profiles featuring no-trade in both contingent markets are not Nash equilibria in pure strategies.*

*Proof.* By Lemmas B.1 and B.2, we can focus only on profiles where thresholds are below cost and prices above valuation:  $\forall i, j \in \{1, 2\}, i \neq j : \max\{\bar{p}_{lj}^*, \bar{p}_{hj}^*\} < r, p_i^* > v_h$ . Then, without loss of generality, Player 1 can profitably deviate to  $r < p_1' = \bar{p}_{h1}' < v_l$ . Her utility function's second term increases while the first one does not decrease.  $\square$

**Lemma B.4.** *Consider the bilateral trade game with heterogeneous buyer valuations and  $\kappa \in (0, 1)$ . Profiles featuring no-trade and exclusion are not Nash equilibria in pure strategies.*

*Proof.* By Lemmas B.1 and B.2, we can focus only on profiles where the contingent market with no trade features thresholds below cost and price above valuation. Suppose first that  $\max\{\bar{p}_{l1}^*, \bar{p}_{h1}^*\} < r, p_2^* > v_h$  and  $\bar{p}_{Q2}^* < p_1^* \leq \bar{p}_{R2}^* < r; Q, R \in \{h, l\}, Q \neq R$ . Notice that Player 1 is selling at a price below cost. She can profitably deviate to  $r < \bar{p}_1' = \bar{p}_{h1}' = p_1' < v_l$ . Her utility function's first term increases while the second one does not decrease.

Suppose alternatively that  $\max\{\bar{p}_{l1}^*, \bar{p}_{h1}^*\} < r, p_2^* > v_h$  but  $\max\{\bar{p}_{Q2}^*, r\} < p_1^* \leq \bar{p}_{R2}^* < p_2^*; Q, R \in \{h, l\}, Q \neq R$ . Notice that this means that  $\max\{\bar{p}_{l1}^*, \bar{p}_{h1}^*\} < p_1^*$ . Then, Player 1 can profitably deviate to  $\bar{p}_{h1}' = p_1^*$ . Her utility function's second term increases while the first one does not decrease.

Finally, assume that  $\max\{\bar{p}_{l1}^*, \bar{p}_{h1}^*\} < r, p_2^* > v_h$  and  $\bar{p}_{Q2}^* < p_1^*, p_2^* \leq p_1^* \leq \bar{p}_{R2}^*; Q, R \in \{h, l\}, Q \neq R$ . Notice that this means that Player 2 is buying at a price above the highest valuation  $v_h$ . She can deviate to  $\bar{p}_{Q2}' \leq \bar{p}_{R2}' = p_2' = p_1^* - \epsilon$ , with  $\epsilon \rightarrow 0$ . She thus makes a material loss as a consequence of the price reduction. This loss can be made arbitrarily small. Then, her utility function's first term increases for a sufficiently small  $\epsilon$  while the second one does not decrease.  $\square$

**Lemma B.5.** *Consider the bilateral trade game with heterogeneous buyer valuations and  $\kappa \in (0, 1)$ . Profiles featuring no-trade and full-trade are not Nash equilibria in pure strategies.*

*Proof.* By Lemmas B.1 and B.2, we can focus only on profiles where the contingent market with no trade features thresholds below cost and price above valuation. Assume without loss of generality that  $\max\{\bar{p}_{l1}^*, \bar{p}_{h1}^*\} < r, p_2^* > v_h$  and  $p_1^* \leq \min\{\bar{p}_{l2}^*, \bar{p}_{l2}^*\}$ .



Suppose first that  $p_1^* < r$ . Then, Player 1 can profitably deviate to  $r < p'_1 = \bar{p}'_{l1}, \bar{p}'_{h1} < p_2^*$ . Her utility function's first term increases while the second one does not decrease.

Alternatively, assume that  $r \leq p_1^* \leq v_l$ . Notice then that  $\max\{\bar{p}_{l1}^*, \bar{p}_{h1}^*\} < p_1^* < p_2^*$ . Player 1 can profitably deviate by setting  $\bar{p}'_{l1}, \bar{p}'_{h1} = p_1^*$ . She makes a moral gain and sustains no material losses.

Finally, suppose that  $p_1^* > v_l$ . Player 2 can profitably deviate to  $r < p'_2 = \bar{p}'_{l2} < p_1^*$ . Her utility function's first term increases while the second one does not decrease.  $\square$

**Lemma B.6.** *Consider the bilateral trade game with heterogeneous buyer valuations and  $\kappa \in (0, 1)$ . Profiles featuring any exchange at a price above the high valuation are not Nash equilibria in pure strategies.*

*Proof.* Assume, without loss of generality, that  $v_h < p_2^* \leq \bar{p}_1^{Q*}$  for some  $Q \in \{h, l\}$ . Suppose first that  $p_1^* < p_2^*$ . Then, Player 1 can deviate to  $\bar{p}_1^{Q'} = p_1^*$ . Her utility function's first term increases while the second one does not decrease. In turn, assume that  $p_1^* > p_2^*$ . If  $p_2^* \leq \bar{p}_1^{Q*} < p_1^*$ , then Player 1 can deviate to  $p'_1 = \bar{p}_1^{Q*}$  and obtain a moral gain without suffering material losses. If  $p_2^* < p_1^* \leq \bar{p}_1^{Q*}$ , then it is Player 2 who can deviate by setting  $p'_2 = \bar{p}_2^{R'} = \bar{p}_1^{Q*} - \epsilon$  for all  $R \in \{h, l\}$ , with  $\epsilon < \bar{p}_1^{Q*} - p_2^*$ . Her utility function's first term increases while the second one does not decrease.

Next, suppose that prices are equal, so  $p_1^* = p_2^*$ . Then, Player 1 can profitably deviate by setting  $\bar{p}_1^{Q'} = p'_1 = p_2^* - \epsilon$ , with  $\epsilon \rightarrow 0$ . The deviation entails a strictly positive material gain from ceasing to buy at  $p_2^* > v_h$  and a potential loss from selling at the lower price  $p_2^* - \epsilon$  and no moral losses. Since the material loss can be made arbitrarily low by choosing a small enough  $\epsilon$ , the deviation is profitable. Her utility function's first term increases while the second one does not decrease.  $\square$

**Lemma B.7.** *Consider the bilateral trade game with heterogeneous buyer valuations and  $\kappa \in (0, 1)$ . Profiles with at least some trading in at least one contingent market are not Nash equilibria if the price is below cost.*

*Proof.* Assume, without loss of generality that  $p_1^* < r, p_1^* \leq \bar{p}_2^{Q*}$  for some  $Q \in \{h, l\}$ . Then, if  $\bar{p}_1^{R*} < r$  for some  $R \in \{h, l\}$ , Player 1 can profitably deviate by setting  $p'_1 = \bar{p}_1^{R'} = r$ , whereas if  $\bar{p}_1^{R*} > r$  for all  $R \in \{h, l\}$ , she can deviate to  $p'_1 = r$ . Her utility function's first term increases while the second one does not decrease.  $\square$

**Lemma B.8.** *Consider the bilateral trade game with heterogeneous buyer valuations and  $\kappa \in (0, 1)$ . Profiles featuring exclusion in one contingent market and full trade in the other one are not Nash equilibria in pure strategies.*

*Proof.* Assume, without loss of generality that  $\bar{p}_2^{Q*} < p_1^* < \bar{p}_2^{R*}$  for  $Q, R \in \{h, l\}, Q \neq R$ , and  $p_2^* \leq \min\{\bar{p}_1^{h*}, \bar{p}_1^{l*}\}$ . By Lemmas B.6 and B.7, we can focus on profiles where  $r \leq p_i^* \leq v_h \forall i \in \{1, 2\}$ . Then:

(1) Profiles where  $p_2^* < p_1^*$  are not equilibria. Indeed, Player 1 will always be able to deviate from a profile where  $p_2^* < p_1^*$  and  $\bar{p}_1^{Q*} < p_1^*$  for some  $Q \in \{h, l\}$  to  $\bar{p}_1^{Q'} \geq p_1^*$ , as she obtains a

moral gain and suffers no material losses. That only leaves  $p_2^* < p_1^* \leq \min\{\bar{p}_1^{h*}, \bar{p}_1^{l*}\}$ . But then, it is Player 2 who can deviate to  $p_2' = \bar{p}_2^{Q'} = p_1^* - \epsilon$ , with  $\epsilon \rightarrow 0$ . Her utility function's first term increases while the second one does not decrease. Therefore, we cannot have  $p_2^* < p_1^*$ .

(2) Profiles where  $p_1^* \leq p_2^*$  are not equilibria. Indeed, Player 1 will always be able to deviate from a profile where  $p_1^* \leq p_2^*$  and  $v_l < p_2^*$  by setting  $p_1' = \bar{p}_1^{l'} = p_2^* - \epsilon$ , with  $\epsilon \rightarrow 0$ . Her utility function's first term increases while the second one does not decrease. That only leaves  $p_1^* \leq p_2^* \leq v_l$ . But then, it is Player 2 who can deviate to  $p_2' = \bar{p}_2^{Q'} = p_1^*$ . Her utility function's first term increases while the second one does not decrease. Therefore, we cannot have  $p_1^* \leq p_2^*$ .  $\square$

**Lemma B.9.** *Consider the bilateral trade game with heterogeneous buyer valuations and  $\kappa \in (0, 1)$ . Profiles featuring exclusion in both contingent markets are not Nash equilibria in pure strategies.*

*Proof.* Consider profiles where  $\bar{p}_{Qi}^* < p_j^* < \bar{p}_{Rj}^*; \forall i, j \in \{1, 2\}, i \neq j$  and  $Q, R \in \{h, l\}, Q \neq R$ . By Lemma B.1, we can focus only on the subset of profiles where  $v_l < p_i^* \leq v_h \forall i \in \{1, 2\}$ . Then:

(1) Profiles where prices are different are not equilibria. Suppose, without loss of generality, that  $p_1^* < p_2^*$ . Then, if  $\bar{p}_{h2}^* < p_2^*$ , Player 2 can deviate to  $\bar{p}_{h2}' = p_2^*$  to obtain a moral gain while suffering no material losses. Thus, we can only have  $\bar{p}_{h2}^* \geq p_2^*$ . But then, Player 1 finds it profitable to deviate to  $p_1' = \bar{p}_{l1}' = p_2^* - \epsilon$ , with  $\epsilon < p_2^* - p_1^*$ . Her utility function's first term increases while the second one does not decrease.

(2) But neither are equilibria profiles where prices are the same. Take  $\max\{\bar{p}_{l1}^*, \bar{p}_{l2}^*\} < p_1^* = p_2^* \leq \min\{\bar{p}_{h1}^*, \bar{p}_{h2}^*\}$ . Then, without loss of generality, Player 1 can deviate to  $p_1' = \bar{p}_{l1}' = p_2^* - \epsilon$ , with  $\epsilon \rightarrow 0$ . She thus obtains a strictly positive moral gain that she trades off against a material loss that can be made arbitrarily small. Her utility function's second term increases while the first one does not decrease.  $\square$

**Lemma B.10.** *Consider the bilateral trade game with heterogeneous buyer valuations and  $\kappa \in (0, 1)$ . Profiles featuring pooling in both contingent markets are Nash equilibria only if  $\bar{p}_{hi}^* \in [(1 - \lambda)p_{li}^* + \lambda r, (\frac{1}{\lambda})p_{li}^* - (\frac{1-\lambda}{\lambda})r] \forall i, j \in \{1, 2\}, i \neq j$ .*

*Proof.* By Lemmas B.6 and B.7, we can focus on profiles where prices are weakly above cost and below the high valuation. Take a profile with full trade in both contingent markets, so that  $\min\{\bar{p}_{hi}^*, \bar{p}_{li}^*\} < p_j^*$  and  $r \leq p_j^* \leq v_h, \forall i, j \in \{1, 2\}, i \neq j$ .

(1) If prices are different, the profile is not a Nash equilibrium. To see it, assume without loss of generality that  $p_1^* < p_2^*$ . Then, if  $\bar{p}_{Q2}^* < p_2^*$  for some  $Q \in \{h, l\}$ , Player 2 can profitably deviate to  $\bar{p}_{Q2}' \geq p_2^*$ . She obtains a moral gain without suffering any material losses. In turn, if  $\bar{p}_{Q2}^* \geq p_2^*$  for all  $Q \in \{h, l\}$  are not Nash equilibria either, as Player 1 can profitably deviate to  $p_1' = p_2^*$ , obtaining a material gain without experiencing moral losses (her utility function's first term increases while the second one does not decrease).

(2) If prices are equal but above the low valuation, the profile is not a Nash equilibrium. Indeed, suppose that  $v_l < p_1^* = p_2^*$ . Then, without loss of generality, Player 1 can deviate to  $\bar{p}_1' = p_1' = p_2^* - \epsilon$ , with  $\epsilon \rightarrow 0$ . She obtains a strictly positive material gain by not buying from Player 2 in the low-valuation case. This gain outweighs the material loss brought about by the price reduction if  $\epsilon$  is small enough. In addition, she suffers no moral losses. Her utility function's first term increases while the second one does not decrease.

(3) With equal prices, profiles where a player's lowest threshold is above the common price are not Nash equilibria. Indeed, suppose without loss of generality that  $p_1^* = p_2^* < \min\{\bar{p}_{l2}^*, \bar{p}_{h2}^*\}$ . Then, Player 1 can deviate to  $p_1' = \bar{p}'_{h1} = \bar{p}'_{l1}$ . Her utility function's first term increases while the second one does not decrease.

(4) With equal prices, if  $(\bar{p}_{li}^*, \bar{p}_{hi}^*)$  are such that  $\bar{p}_{hi}^* \notin [(1 - \lambda)p_{li}^* + \lambda r, (\frac{1}{\lambda})p_{li}^* - (\frac{1-\lambda}{\lambda})r]$  for some  $i \in \{1, 2\}$  then the profile is not an equilibrium. Indeed, suppose without loss of generality that  $\bar{p}_{h2}^* \notin [(1 - \lambda)p_{l2}^* + \lambda r, (\frac{1}{\lambda})p_{l2}^* - (\frac{1-\lambda}{\lambda})r]$ . Then, Player 1 can deviate to  $p_1' = \bar{p}'_{h1} = \bar{p}'_{l1} = \max\{p_{l2}^*, p_{h2}^*\}$ . Her utility function's first term increases while the second one does not decrease.

(5) Profiles where  $r \leq p_i^* = p_j^* \leq v_l$  and  $\bar{p}_{hi}^* \in [(1 - \lambda)p_{li}^* + \lambda r, (\frac{1}{\lambda})p_{li}^* - (\frac{1-\lambda}{\lambda})r] \forall i, j \in \{1, 2\}, i \neq j$  are equilibria. Indeed, notice that given  $(p_j^*, p_{lj}^*, p_{hj}^*), (p_i^*, p_{li}^*, p_{hi}^*)$  maximises ???. The strategies are then mutual best responses and therefore, these profiles are Nash equilibria.  $\square$

**Proposition B.11.** *The set of Nash equilibria of the bilateral trade game with heterogeneous buyer valuations and  $\kappa = 1$  is  $p_i^* \leq \min\{\bar{p}_{hi}^*, \bar{p}_{li}^*\} \forall i \in \{1, 2\}$*

*Proof.* Given that  $r < v_l < v_h$ , Expression (6) with  $\kappa = 1$  is maximised if and only if  $p_i^* \leq \min\{\bar{p}_{hi}^*, \bar{p}_{li}^*\}$ .  $\square$

## Appendix C Appendix to Section 4.1

**Proposition 4.2\*.** *Consider the bilateral trade game with with adverse selection between two homo oeconomicus and assume that  $0 < r_l < v_l < r_h < v_h$ . The Nash equilibria in pure strategies are:*

1. If  $\lambda \geq \lambda^e$ :

- *Full-trade:*  $\bar{p}_i^* = p_{jl}^* = p_{jh}^* \in [r_h, v^e]; \forall i, j \in \{1, 2\}, i \neq j$
- *Adverse selection and pooling:*  $\bar{p}_i^* = p_{jl}^* \in [r_l, v_l], p_{jh}^* > v_h, \bar{p}_j^* = p_{il}^* = p_{ih}^* \in [r_h, v^e]; i, j \in \{1, 2\}, i \neq j$
- *No trade and pooling:*  $\bar{p}_i^* < r_l, p_{jh}^* > \min\left\{\max\left\{p_{jl}^*, \frac{1-\lambda}{\lambda}(v_l - p_{jl}^*)\right\}, v_h\right\}, p_{jl}^* > v_l, \bar{p}_j^* = p_{il}^* = p_{ih}^* \in [r_h, v^e]; i, j \in \{1, 2\}, i \neq j$
- *Adverse selection (both cont. markets):*  $\bar{p}_i^* = p_{jl}^* \in [r_l, v_l], p_{jh}^* > v_h, \forall i, j \in \{1, 2\}, i \neq j$

- *Adverse selection and no-trade*:  $\bar{p}_i^* = p_{jl}^* \in [r_l, v_l], p_{jh}^* > v_h,$   
 $\bar{p}_j^* < r_l, p_{ih}^* > \min \left\{ \max \left\{ p_{il}^*, \frac{1-\lambda}{\lambda}(v_l - p_{il}^*) \right\}, v_h \right\}, p_{il}^* > v_l; i, j \in \{1, 2\}, i \neq j$
- *No-trade*:  $\bar{p}_i^* < r_l, p_{jh}^* > \min \left\{ \max \left\{ p_{jl}^*, \frac{1-\lambda}{\lambda}(v_l - p_{jl}^*) \right\}, v_h \right\}, p_{jl}^* > v_l; \forall i, j \in \{1, 2\}, i \neq j$

2. If  $\lambda < \lambda^e$ , the Nash equilibria in pure strategies of the main game are:

- *Adverse selection (both cont. markets)*:  $\bar{p}_i^* = p_{jl}^* \in [r_l, v_l], p_{jh}^* > v_h, \forall i, j \in \{1, 2\}, i \neq j$
- *Adverse selection and no-trade*:  $\bar{p}_i^* = p_{jl}^* \in [r_l, v_l], p_{jh}^* > v_h,$   
 $\bar{p}_j^* < r_l, p_{ih}^* > \min \left\{ \max \left\{ p_{il}^*, \frac{1-\lambda}{\lambda}(v_l - p_{il}^*) \right\}, v_h \right\}, p_{il}^* > v_l; i, j \in \{1, 2\}, i \neq j$
- *No-trade*:  $\bar{p}_i^* < r_l, p_{jh}^* > \min \left\{ \max \left\{ p_{jl}^*, \frac{1-\lambda}{\lambda}(v_l - p_{jl}^*) \right\}, v_h \right\}, p_{jl}^* > v_l; \forall i, j \in \{1, 2\}, i \neq j$

When agents are *homo oeconomicus*, their utility function reduces to their expected material payoff, shown in Expression (3) (substituting for the appropriate seller and buyer payoffs shown in (4)). Notice that the terms inside the first square brackets of Player  $i$ 's utility function are only affected by her threshold  $\bar{p}_i$  and Player  $j$ 's prices  $p_{jh}$  and  $p_{jl}$ . Meanwhile, the second square brackets of Player  $i$ 's utility function are only affected by her prices  $p_{ih}$  and  $p_{il}$  and by Player  $j$ 's threshold  $\bar{p}_j$ . I then study the “contingent” game between a buyer who sets  $\bar{p} \in \mathbb{R}_+$  and a seller who decides  $(p_{2g}, p_{2b})$ . Their utilities are respectively  $U^B$  and  $U^S$ :

$$U^B(\bar{p}; p_h, p_l) = \lambda \mathbb{1}\{p_h \leq \bar{p}\}(v_h - p_h) + (1 - \lambda) \mathbb{1}\{p_l \leq \bar{p}\}(v_l - p_l), \quad (13)$$

$\in \mathbb{R}_+^2$  and has utility:

$$U^S(p_h, p_l; \bar{p}) = \lambda (\mathbb{1}\{p_h \leq \bar{p}\}(p_h - r_h)) + (1 - \lambda) (\mathbb{1}\{p_l \leq \bar{p}\}(p_l - r_l)). \quad (14)$$

As with the symmetric information case, I find the Nash equilibria in pure strategies of the contingent game by writing down the best-response correspondences and looking for their intersection. The result is summarised in Lemma C.1, where the term  $v^e$  is used to refer to expected quality  $\lambda v_h + (1 - \lambda)v_l$ .

**Lemma C.1.** *Assume that  $0 < r_l < v_l < r_h < v_h$ . Then, the Nash equilibria in pure strategies of the contingent game between a seller who chooses prices  $(p_h^*, p_l^*)$  and a buyer who chooses threshold  $\bar{p}$  are:*

(1) If  $r_h \leq v^e$ :

- *Pooling*:  $\bar{p}_1^* = p_g^* = p_l^* \in [r_h, v^e]$
- *Adverse selection (only bad quality traded)*:  $\bar{p}_1^* = p_l^* \in [r_l, v_l], p_h^* > v_h$

- No trade:  $\bar{p}^* < r_l, p_l^* > v_l, p_h^* > \min \left\{ \max \left\{ p_l^*, \frac{1-\lambda}{\lambda}(v_l - p_l^*) \right\}, v_h \right\}$

(2) If  $r_h > v^e$ :

- Adverse selection (only bad quality traded):  $\bar{p}_1^* = p_l^* \in [r_l, v_l], p_g^* > v_h$

- No trade:  $\bar{p}^* < r_l, p_l^* > v_l, p_h^* > \min \left\{ \max \left\{ p_l^*, \frac{1-\lambda}{\lambda}(v_l - p_l^*) \right\}, v_h \right\}$

*Proof.* The seller's best reply correspondences in the contingent game are:

$$p_h^{BR} = \operatorname{argmax}_{p_h} U^S(p_h, p_l; \bar{p}) = \begin{cases} p_h^{BR} = \bar{p}_1 & \text{if } \bar{p} \geq r_h \\ p_h^{BR} > \bar{p}_1 & \text{if } \bar{p} < r_h \end{cases}$$

$$p_l^{BR} = \operatorname{argmax}_{p_l} U^S(p_h, p_l; \bar{p}) = \begin{cases} p_l^{BR} = \bar{p} & \text{if } \bar{p} \geq r_l \\ p_l^{BR} > \bar{p} & \text{if } \bar{p} < r_l \end{cases}$$

The buyer's best reply correspondence in the contingent game is:

$$\begin{cases} \bar{p}^{BR} \geq \max\{p_h, p_l\} & \text{if } 0 \leq p_l \leq v_l, 0 \leq p_h \leq v_h \\ \bar{p}^{BR} \in [p_l, p_h) & \text{if } 0 \leq p_l \leq v_l, v_h < p_h \\ \bar{p}^{BR} \in [p_h, p_l) & \text{if } v_l < p_l, p_h \leq \min\{v_h, p_l\} \\ \bar{p}^{BR} < \min\{p_h, p_l\} & \text{if } v_l < p_l, p_h > \min \left\{ \max \left\{ v_h + \frac{1-\lambda}{\lambda}(v_l - p_l), p_l \right\}, v_h \right\} \\ \bar{p}^{BR} \geq \max\{p_h, p_l\} & \text{if } v_l < p_l \leq p_h \leq v_h, 0 \leq p_h \leq v_h + \frac{1-\lambda}{\lambda}(v_l - p_l) \end{cases}$$

If  $r_h \leq v^e$ , the best replies intercept whenever:

$$\bar{p}^* = p_h^* = p_l^* \in [r_h, \lambda v_h + (1-\lambda)v_l], \bar{p}^* = p_l^* \in [r_l, v_l], p_h^* > v_h \text{ and}$$

$$\bar{p}^* < r_l, p_l^* > v_l, p_h^* > \min \left\{ \max \left\{ p_b^*, \frac{1-\lambda}{\lambda}(v_l - p_b^*) \right\}, v_h \right\}.$$

If  $r_h > v^e$ , they do so when:

$$\bar{p}^* = p_l^* \in [r_l, v_l], p_h^* > v_h \text{ and } \bar{p}^* < r_l, p_l^* > v_l, p_h^* > \min \left\{ \max \left\{ p_b^*, \frac{1-\lambda}{\lambda}(v_l - p_l^*) \right\}, v_h \right\}.$$

□

### Proof of Proposition 4.2:

*Proof.* Immediate from the fact that the main game is composed of two contingent games and Lemma C.1. □

**Proposition C.2.** Consider the bilateral trade game with with adverse selection between two fully moral agents and assume that  $0 < r_l < v_l < r_h < v_h$ . The Nash equilibria in pure strategies are of the form  $\bar{p}_i^* \geq \max\{p_{ih}^*, p_{il}^*\} \forall i \in \{1, 2\}$ .

*Proof.* Given that  $r_h < v_h$  and  $r_l < v_l$ , utility is maximised  $\bar{p}_i^* \geq \max\{p_{ih}^*, p_{il}^*\}$ ,  $\forall i \in \{1, 2\}$ .  $\square$

**Corollary C.2.1.** *The equilibria between two homo kantiensis always feature some trade.*

Since it is always true that either  $\bar{p}_1^* \geq \bar{p}_2^*$  or  $\bar{p}_2^* \geq \bar{p}_1^*$ , in equilibrium there will always be a player who is selling to the other one in both states of the world. In contrast, the high-threshold agent (who is buying in both states of the world) may or may not post prices low enough to sell to the other player. Hence, equilibria may feature full trade, trade of both qualities when one player is the seller and no trade when she is the buyer, or trade of both qualities when one player is the seller and trade of only one quality when roles are reversed.

**Proposition 4.3\*.** *Consider the bilateral trade game with adverse selection between two homo moralis and assume  $0 < r_l < v_l < r_h < v_h$ . If  $\lambda \geq \lambda^e$ , the only Nash equilibria are profiles where  $r_h \leq \bar{p}_i^* = p_{ih}^* = p_{il}^* \leq v^e$ ; for all  $i \in \{1, 2\}$ . In turn, if  $\lambda < \lambda^e$ , there exist no Nash equilibria in pure strategies.*

*Proof.* The proof follows directly from Lemmas B.1 to B.10, presented below.  $\square$

**Lemma C.3.** *Consider the bilateral trade game with adverse selection between two homo moralis and assume  $0 < r_l < v_l < r_h < v_h$ . Then, profiles featuring no trade in at least one contingent market where either of conditions (1) to (4) are true, are not Nash equilibria in pure strategies. The conditions are: (1) the low quality price below is (weakly) the low valuation, (2) the low quality price is (weakly) above low quality valuation but below the high quality price and the expected price is not above expected valuation (3) the high quality price is below both the low quality price and the high valuation, (4) the threshold is weakly above low-quality cost.*

*Proof.* Assume without loss of generality, that  $\bar{p}_1^* < \min\{p_{2h}^*, p_{2l}^*\}$ . Then:

(1) Assume also that  $p_{2l}^* \leq v_l$ . Player 1 can profitably deviate to  $\bar{p}'_1 = p_{2l}^*$ . The first term of her utility function increases because  $\bar{p}'_1 = p_{2l}^* \leq v_l$ . The second term does not decrease.

(2) Assume also that  $v_l \leq p_{2l}^* \leq p_{2h}^* \leq v_h + \frac{1-\lambda}{\lambda}(v_l - p_{2l}^*)$ . Player 1 can profitably deviate to  $\bar{p}'_1 = p_{2h}^*$ . The first term of her utility function increases because  $\bar{p}'_1 = p_{2h}^* \leq v_h + \frac{1-\lambda}{\lambda}(v_l - p_{2l}^*)$ . The second term does not decrease.

(3) Assume also that  $p_{2h}^* < p_{2l}^*, p_{2h}^* \leq v_h$ . Player 1 can profitably deviate to  $\bar{p}'_1 = p_{2h}^*$ . The first term of her utility function increases because  $\bar{p}'_1 = p_{2h}^* \leq v_h$ . The second term does not decrease.

(4) Assume also that  $\bar{p}_1^* \geq r_l$ . Player 2 can deviate to  $p'_{2l} = \bar{p}_1^*$ . The first term of her utility function increases because  $p'_{2l} \geq r_l$ . The second term does not decrease.  $\square$

**Lemma C.4.** *Consider the bilateral trade game with adverse selection between two homo moralis and assume  $0 < r_l < v_l < r_h < v_h$ . Profiles where no trade takes place in either contingent market are not Nash equilibria in pure strategies.*

*Proof.* No-trade profiles are described by  $\bar{p}_i^* < \min\{p_{jh}^*, p_{jl}^*\}$  and  $\bar{p}_j^* < \min\{p_{ih}^*, p_{il}^*\}$  (with  $i, j \in \{1, 2\}, i \neq j$ ). From Lemma C.3, only profiles featuring  $\bar{p}_i^* < r_l, \bar{p}_j^* < r_l, p_{il}^* > v_l, p_{jl}^* > v_l, p_{ih}^* >$

$\min\{\max\{p_{il}^*, \frac{1-\lambda}{\lambda}(v_l - p_{il}^*)\}, v_h\}$  and  $p_{jh}^* > \min\{\max\{p_{jl}^*, \frac{1-\lambda}{\lambda}(v_l - p_{jl}^*)\}, v_h\}$  remain as candidate equilibria. Next, consider, without loss of generality, Player 1. She can profitably deviate by setting  $\bar{p}'_1 \in (\bar{p}_2^*, \min\{p_{2l}^*, p_{2h}^*\})$  and new prices  $\{(p'_{1l}, p'_{1h}) : p'_{1l} \in (\bar{p}_2^*, \bar{p}'_1], p'_{1h} \in (\bar{p}_2^*, \bar{p}'_1]\}$ . The first term of her utility function remains the same, while the second one increases.  $\square$

**Lemma C.5.** *Consider the bilateral trade game with adverse selection between two homo moralis and assume  $0 < r_l < v_l < r_h < v_h$ . Profiles with trade of only the good quality in at least one contingent market are not Nash equilibria in pure strategies.*

*Proof.* Assume without loss of generality that  $p_{1h}^* \leq \bar{p}_2^* < p_{1l}^*$ . Then, if in addition:

(1)  $\bar{p}_1^* \leq \bar{p}_2^* \leq r_l$ , Player 1 can deviate to  $p'_{1h} \in (\bar{p}_2^*, r_l)$  and  $\bar{p}'_1 = p'_{1h}$ . The first term of her utility function increases while the second one does not decrease.

(2)  $p_{1h}^* \leq \bar{p}_2^* < p_{1l}^*$ ,  $\bar{p}_2^* < \bar{p}_1^*$  and  $\bar{p}_2^* \leq r_l$ , Player 1 can deviate to  $p'_{1h} \in (\bar{p}_2^*, \bar{p}_1^*]$ . The first term of her utility function increases while the second one stays the same.

(3)  $p_{1h}^* \leq \bar{p}_2^* < p_{1l}^*$  and  $\bar{p}_2^* > r_l$ , Player 1 can deviate to  $p'_{1l} = \bar{p}_2^*$ . The first term of her utility function increases while the second one does not decrease.  $\square$

In contrast with the *homo oeconomicus* variant of the model, profiles where one contingent market features no trade at all and the other one presents trade of only the bad quality or trade of both qualities are not equilibria of the game between two *homo moralis* agents. Lemmas C.6 and C.7 leverage Lemma C.3 in order to show this, checking only profiles in which the contingent market with no trade has prices above valuations and threshold below costs.

**Lemma C.6.** *Consider the bilateral trade game with adverse selection between two homo moralis and assume  $0 < r_l < v_l < r_h < v_h$ . Then, profiles with full trade in one contingent market and only trade of the bad quality in the other one are not equilibria.*

*Proof.* Assume without loss of generality that  $\bar{p}_1^* < \min\{p_{2h}^*, p_{2l}^*\}$ . From Lemma C.3, only when  $\bar{p}_1^* < r_l$ ,  $p_{2l}^* > v_l$  and  $p_{2h}^* > \min\{\max\{p_{2l}^*, \frac{1-\lambda}{\lambda}(v_l - p_{2l}^*)\}, v_h\}$  can these profiles be equilibria. Then, assume alternatively that:

(1)  $p_{1l}^* \leq \bar{p}_2^* < r_l$ . Player 1 can deviate to  $p'_{1l} \in (\bar{p}_2^*, r_l)$  and  $\bar{p}'_1 = p'_{1l}$ . The first term of her utility function increases because  $r_l > p_{1l}^*$  while the second one remains the same (if  $\bar{p} \geq p_{1l}^*$ ) or increases (if  $\bar{p}_1^* < p_{1l}^*$ ).

(2)  $v_l < p_{1l}^* \leq \bar{p}_2^*$ . Player 2 can deviate to  $\bar{p}'_2 \in (v_l, p_{1l}^*)$  and  $p'_{2l} \in (v_l, \bar{p}'_2]$  if  $\bar{p}'_2 < p_1 b^*$  and to  $p'_{2g} \in (v_l, \bar{p}'_2]$  if  $\bar{p}'_2 < p_{1h}^*$ . The first term of her utility function increases because  $v_l < p_{1l}^*$  while the second one remains the same.

(3)  $r_l \leq p_{1l}^* \leq \bar{p}_2^* \leq v_l$ . Player 2 can deviate to  $p'_{2g} = p'_{2l} \in (\bar{p}_1^*, \bar{p}_2^*]$ . The first term of her utility function remains the same while the second one increases.  $\square$

Notice that the last point of the proof of Lemma C.6 resorts to a deviation that has no effect with *homo oeconomicus* agents, as it does not lead to more or less trade nor does it modify the prices at which it is carried out.

**Lemma C.7.** *Consider the bilateral trade game with adverse selection between two homo moralis and assume  $0 < r_l < v_l < rh < v_h$ . Then, profiles with full trade in one contingent market and no trade in the other one are not Nash equilibria in pure strategies.*

*Proof.* Assume, without loss of generality, that  $\bar{p}_1^* < \min\{p_{2h}^*, p_{2l}^*\}$  and  $\max\{p_{1h}^*, p_{1l}^*\} \leq \bar{p}_2^*$ . From Lemma C.3, only if  $\bar{p}_1^* < r_l$ ,  $p_{2l}^* > v_l$  and  $p_{2h}^* > \min\{\max\{p_{2l}^*, \frac{1-\lambda}{\lambda}(v_l - p_{2l}^*)\}, v_h\}$  can these profiles be equilibria. Then, assume alternatively that:

(1)  $\max\{p_{1g}^*, p_{1b}^*\} \leq \bar{p}_2^* < r_l$ . Player 1 can deviate by setting  $p_{1g}'$ ,  $p_{1b}'$  and  $\bar{p}_1'$  such that  $p_{1h}' \in (\bar{p}_2^*, \bar{p}_1']$ ,  $p_{1l}' \in (\bar{p}_2^*, \bar{p}_1']$  and  $\bar{p}_1' < p_{2l}^*$ . The first term of her utility function increases because  $p_{1h}^* < r_l$  and  $p_{1l}^* < r_l$  while the second one does not decrease.

(2) Assume alternatively to (2), that  $v_l < p_{1l}^* \leq \bar{p}_2^*$ . Player 2 can deviate to  $\bar{p}_2' \in (\bar{p}_1^*, p_{1b}^*)$  and also lowering  $p_{2b}$  to  $p_{2l}' \in (\bar{p}_1^*, \bar{p}_2']$ . The first term of her utility function increases because  $v_l < p_{1l}^*$  while the second one increases does not decrease.

(3) Assume, in turn, that  $r_l \leq p_{1l}^* \leq \bar{p}_2^* \leq v_l$ . Player 2 can deviate to  $p_{1l}' \in (\bar{p}_1^*, \bar{p}_2^*]$ . The first term of her utility function remains the same while the second one increases.  $\square$

**Lemma C.8.** *Consider the bilateral trade game with adverse selection between two homo moralis and assume  $0 < r_l < v_l < rh < v_h$ . Then, profiles with only trade of the bad quality in both contingent markets are not Nash equilibria in pure strategies.*

*Proof.* Assume, without loss of generality, that  $\bar{p}_1^* < p_{2h}^*$  and  $\bar{p}_2^* < p_{1h}^*$ . Next, consider a profile where  $p_{1l}^* \leq \bar{p}_2^* < r_l$ . Player 1 can deviate to  $p_{1l}' \in (\bar{p}_2^*, r_l)$  and  $\bar{p}_1' = p_{1l}'$  if  $\bar{p}_1^* < p_{1l}'$ . The first term of her utility function increases because  $p_{1l}^* < r_l$  while the second one remains the same. This shows that only profiles with  $r_l \leq p_{1l}^* \leq \bar{p}_2^*$  and  $r_l \leq p_{2l}^* \leq \bar{p}_1^*$  remain as equilibrium candidates.

(1) Assume now that  $r_l \leq p_{1l}^* \leq \bar{p}_2^*$  and  $r_l \leq p_{2l}^* \leq \bar{p}_1^*$  and, without loss of generality, that  $\bar{p}_1^* \leq \bar{p}_2^*$ . If  $p_{2l}^* < \bar{p}_1^*$ , Player 2 can deviate to  $p_{2l}' = \bar{p}_1^*$ . The first term of her utility function increases because  $p_{2l}^* < p_{2l}'$  while the second one remains the same. Thus, if  $\bar{p}^* \leq \bar{p}_2^*$ , only profiles with  $p_{2l}^* = \bar{p}_1^*$  remain as equilibrium candidates.

(2) Assume that  $r_l \leq p_{1l}^* \leq \bar{p}_2^*$  and  $r_l \leq p_{2l}^* = \bar{p}_1^* \leq \bar{q}_2^*$ . If  $\bar{p}^* < \bar{q}_2^*$ , Player 1 can deviate to  $p_{1l}'$  and  $\bar{p}$  to  $\bar{p}_1' = \bar{p}_2^*$ . The first term of her utility function increases because  $p_{1l}^* < p_{1l}'$  while the second one remains the same. Thus, only profiles with  $r_l \leq p_{2l}^* = p_{2l}^* = \bar{p}_1^* = \bar{p}_2^*$  remain as equilibrium candidates.

(3) Consider now profiles where  $v_l < p_{1l}^* = p_{2l}^* = \bar{p}_1^* = \bar{p}_2^*$ . Player 1 can deviate by marginally reducing  $\bar{p}_1$  to  $\bar{p}_1' = \bar{p}_1^* - \epsilon$  (with  $\epsilon \rightarrow 0$ ) while at the same time lowering  $p_{1b}$  to  $p_{1l}' = \bar{p}'$ . The first term of her utility function increases because  $p_{1l}^* - p_{1l}'$  can be made arbitrarily small, whereas the quantity  $q_2 b^* - v_l$  is fixed. The second term of her utility function remains the same. Thus, only profiles with  $r_l \leq p_{2l}^* = p_{1l}^* = \bar{p}_1^* = \bar{p}_2^* \leq v_l$  remain as equilibrium candidates.

(4) Assume  $r_l \leq p_{2l}^* = p_{1l}^* = \bar{p}_1^* = \bar{p}_2^* \leq v_l$  and  $p_{1h}^* \leq \min\{\max\{p_{2l}^*, v_h + \frac{1-\lambda}{\lambda}(v_l - p_{2l}^*)\}, v_h\}$ . Then, Player 1 can deviate to  $\bar{p}_1' = p_{2h}^*$ . The first term of (??) increases, whereas the second one increases (if  $p_{1h}^* \in (\bar{p}_1^*, p_{2h}^*]$ ) or remains the same (if  $p_{1h}^* > p_{2h}^*$ ).

(5) Consider now profiles where  $r_l \leq p_{2l}^* = p_{1l}^* = \bar{p}_1^* = \bar{p}_2^* \leq v_l$  and  $p_{2h}^* > \min\{\max\{p_{2l}^*, v_h + \frac{1-\lambda}{\lambda}(v_l - p_{2l}^*)\}, v_h\}$ . Then, Player 1 can deviate to  $\bar{p}_1' \in (\bar{p}_2^*, p_{2h}^*)$ . The first term of her utility function remains the same, whereas the second one increases.  $\square$



**Lemma C.9.** *Consider the bilateral trade game with adverse selection between two homo moralis and assume  $0 < r_l < v_l < r_h < v_h$ . Then, profiles with only trade of the bad quality in one contingent market and full trade in the other one are not Nash equilibria in pure strategies.*

*Proof.* Assume, without loss of generality, that  $p_{1l} \leq \bar{p}_2^* < p_{1h}^*$  and  $\max\{p_{2l}^*, p_{2h}^*\} \leq \bar{p}_1^*$ . Then:

(1) Suppose that  $\bar{p}_1^* \leq \bar{p}_2^*$ . Profiles with  $\max\{p_{2l}^*, p_{2h}^*\} < \bar{p}_1^*$  are not equilibria. Player 2 can set  $p_{2Q} = \bar{p}_1^*$ ,  $Q \in \{h, l\}$ . The first term of her utility function while the second one remains the same because  $\bar{p}_1^* \leq \bar{p}_2^*$ .

(2) Suppose that  $p_{2l}^* = p_{2h}^* = \bar{p}_1^* \leq \bar{p}_2^*$  and  $p_{1l}^* \leq \bar{p}_2^* < p_{1h}^*$ . If  $p_{1l}^* < \bar{p}_2^*$ , Player 1 can deviate to  $p'_{1l} = \bar{p}_1^* = \bar{p}_2^*$ . The first term of her utility function while the second one remains the same.

(3) Assume that  $p_{2l}^* = p_{2h}^* = \bar{p}_1^* = p_{1l}^* = \bar{q}^* < r_h$ . Player 2 can deviate to  $p'_{2g} > \bar{p}_1^*$  and  $\bar{p}'_2 = p'_{2g}$ . The first term of her utility function while the second one remains the same.

(4) Assume, alternatively to (3), that  $r_h \leq p_{2l}^* = p_{2h}^* = \bar{p}_1^* = p_{1l}^* = \bar{p}_2^* < p_{1h}^*$ . Player 1 can deviate to  $p'_{1h} = p_{2l}^* = p_{2h}^* = \bar{p}_1^* = p_{1l}^* = \bar{p}_2^*$ . Both terms of her utility function increase.

(5) Assume  $p_{1l} < \bar{p}_1^* < p_{1h}^*$ ,  $\max\{p_{2l}^*, p_{2h}^*\} \leq \bar{p}_1^*$  and  $\bar{p}_1^* > \bar{p}_2^*$ . Player 1 can deviate to  $p'_{1l} = \bar{p}_2^*$ . The first term of her utility function increases because  $p'_{1l} = \bar{p}_2^* < \bar{p}_1^*$ , while the second one remains the same.

(6) Assume that  $p_{1l}^* = \bar{p}_2^* < \min\{r_l, p_{1h}^*\}$ ,  $\max\{p_{2l}^*, p_{2h}^*\} \leq \bar{p}_1^*$  and  $\bar{p}_1^* > \bar{p}_2^*$ . Player 1 can deviate to  $p'_{1l} \in (\bar{p}_2^*, \bar{p}_1^*]$ . The first term of her utility function increases because  $p'_{1l} < r_l$ , while the second one remains the same because  $p'_{1l} \leq \bar{p}_1^*$ .

(7) Assume that  $p_{1l}^* = \bar{p}_2^* > v_l$ ,  $\bar{p}_2^* < p_{1h}^*$ ,  $\max\{p_{2l}^*, p_{2h}^*\} \leq \bar{p}_1^*$  and  $\bar{p}_1^* > \bar{p}_2^*$ . If either  $p_{2h}^* < \bar{p}_2^* < p_{2l}^*$ ,  $p_{2l}^* < \bar{p}_2^* < p_{2h}^*$  or  $\max\{p_{1l}^*, p_{2h}^*\} < \bar{p}_2^*$ , Player 2 can deviate to  $\bar{p}'_2 \in [p_{2h}^*, p_{1l}^*]$ ,  $\bar{p}'_2 \in [p_{2l}^*, p_{1l}^*]$  or  $\bar{p}'_2 \in [\max\{p_{2l}^*, p_{2h}^*\}, p_{1l}^*]$  respectively. The first term of her utility function increases because  $p_{1l}^* > v_l$ , while the second one remains the same. In turn, if  $\min\{p_{2l}^*, p_{2h}^*\} > \bar{p}_2^*$ , then Player 2 can profitably deviate  $\bar{p}'_2 < p_{1l}^*$ . The first term of her utility function increases because  $p_{1l}^* > v_l$ , while the second one remains the same.

If  $p_{2l}^* = \bar{p}_2^*$  or  $p_{2h}^* = \bar{p}_2^*$  (or both), Player 2 can deviate to  $\bar{p}'_2 = p_{1l}^* - \epsilon$  (with  $\epsilon \rightarrow 0$ ),  $p'_{2l} = \bar{p}'_2$  and/or  $p_{2g} \rightarrow p'_{2g} = \bar{p}'_2$ . The first term of her utility function increases because  $p_{2l}^* - p'_{2l}$  can be made arbitrarily small, whereas the quantity  $p_{2l}^* - v_l$  is fixed. The second term of her utility function remains the same.

(8) Assume  $\bar{p}_2^* = p_{1l}^* \in [r_l, v_l]$ ,  $\bar{p}_2^* < p_{1h}^*$ ,  $p_{2h}^* \leq \bar{p}_1^*$ ,  $\bar{p}_2^* < \bar{p}_1^*$  and  $p_{1h}^* < v_h$ . Player 2 can deviate to  $\bar{p}'_2 \geq p_{1h}^*$ . The first term of her utility function increases because  $p_{1h}^* < v_h$ , while the second one increases (if  $q_{2b} \in (\bar{p}_2^*, \bar{p}'_2]$  or  $p_{2g} \in (\bar{p}_2^*, \bar{p}'_2]$ ) or remains the same (if  $p_{2b} \leq \bar{p}_2^*$  and  $p_{2g} \leq \bar{p}_2^*$  or  $p_{2b} > \bar{p}'_2$  and  $p_{2g} > \bar{p}'_2$ ).

(9) Assume that  $\bar{p}_2^* = p_{1l}^* \in [r_l, v_l]$ ,  $\bar{p}_2^* < p_{1h}^*$ ,  $v_h < p_{1h}^*$ ,  $\max\{p_{2l}^*, p_{2h}^*\} \leq \bar{p}_1^*$ ,  $\bar{p}_2^* < \bar{p}_1^*$  and  $\min\{p_{2h}^*, p_{2l}^*\} > v_h$ . Player 1 can deviate to  $\bar{p}'_1 < \min\{p_{2h}^*, p_{2l}^*\}$  to  $p'_{1h} = \bar{p}'_1$  (if  $\bar{p}'_1 < p_{1h}^*$ ). The first term of her utility function increases because  $\min\{p_{2g}^*, p_{2b}^*\} > v_h$ , while the second one remains the same.

(10) Assume that  $\bar{p}_2^* = p_{1l}^* \in [r_l, v_l]$ ,  $\bar{p}_2^* < p_{1h}^*$ ,  $v_h < p_{1h}^*$ ,  $\max\{p_{2l}^*, p_{2h}^*\} \leq \bar{p}_1^*$ ,  $\bar{p}_2^* < \bar{p}_1^*$ ,  $p_{2h}^* < p_{2l}^* \leq \bar{p}_1^*$  and  $p_{2l}^* > v_l$ . Player 1 can deviate to  $\bar{p}'_1 \in [p_{2h}^*, p_{2l}^*]$  and  $p'_{1h} = \bar{p}'_1$  (if  $\bar{p}'_1 < p_{1h}^*$ ). The first term of her utility function increases because  $p_{2b}^* > v_l$ , while the second one increases

if.

(11) Assume that  $\bar{p}_2^* = p_{1l}^* \in [r_l, v_l]$ ,  $\bar{p}_2^* < p_{1h}^*$ ,  $v_h < p_{1h}^*$ ,  $\max\{p_{2l}^*, p_{2h}^*\} \leq \bar{p}_1^*$ ,  $\bar{p}_2^* < \bar{p}_1^*$  and  $p_{2h}^* \leq \bar{p}_1^* < r_h$ . Player 2 can deviate to  $p'_{2g} > \bar{p}^*$  while at the same time augmenting  $\bar{q}$  to  $\bar{p}'_2 \in [p'_{2g}, p_{1h}^*)$ . The first term of her utility function increases because  $p_{2l}^* > v_l$ , while the second one increases if.

(12) Assume that  $\bar{p}_2^* = p_{1l}^* \in [r_l, v_l]$ ,  $\bar{p}_2^* < p_{1h}^*$ ,  $v_h < p_{1h}^*$ ,  $\max\{p_{2l}^*, p_{2h}^*\} \leq \bar{p}^*$ ,  $\bar{p}_2^* < \bar{p}^*$  and  $p_{2l}^* \leq \bar{p}_1^* < r_l$ . Player 2 can deviate to  $p'_{2l} > \bar{p}_1^*$  and  $\bar{p}'_2 \in [p'_{2l}, p_{1h}^*)$ . The first term of her utility function increases because  $p_{2l}^* > v_l$ , while the second one increases if.

(13) Assume that  $\bar{p}_2^* = p_{1l}^* \in [r_l, v_l]$ ,  $\bar{p}_2^* < p_{1h}^*$ ,  $v_h < p_{1h}^*$ ,  $\max\{p_{2l}^*, p_{2h}^*\} \leq \bar{p}_1^*$  and  $\bar{p}_2^* < \bar{p}_1^*$  while at the same time,  $r_l \leq p_{2l}^* \leq p_{2h}^*$  and  $r_h \leq p_{2h}^* \leq v_h$ . Notice that  $r_h \leq p_{2h}^*$  and  $\bar{p}_2^* = p_{1l}^* \in [r_l, v_l]$  implies that  $\bar{p}_2^* < p_{2h}^*$ . Player 2 can deviate to  $\bar{p}'_2 = p_{2h}^*$ . The first term of her utility function remains the same, while the second one increases. □

**Lemma C.10.** *Consider the bilateral trade game with adverse selection between two homo moralis and assume  $0 < r_l < v_l < r_h < v_h$ . Then, profiles with full trade in at least one contingent market at a price below high quality cost are not equilibria in pure strategies of the game between two homo moralis.*

*Proof.* Assume, without loss of generality, that  $\max\{p_{1l}^*, p_{1h}^*\} \leq \bar{p}_2^* < r_h$ . Then, if in addition,  $\bar{p}_2^* < \bar{p}_1^*$ , Player 1 can profitably deviate to  $p'_{1h} \in (\bar{p}_2^*, \bar{p}_1^*]$ . The first term of her utility function increases while the second one remains the same. In turn, assume that  $\bar{p}_1^* \leq \bar{p}_2^*$ . Then, Player 1 can profitably deviate to  $p'_{1h} > \bar{p}_2^*$  and  $\bar{p}'_1 = p'_{1h}$ . The first term of her utility function increases while the second one does not decrease. □

**Lemma C.11.** *Consider the bilateral trade game with adverse selection between two homo moralis and assume  $0 < r_l < v_l < r_h < v_h$ . Then, profiles where with full trade in at least one contingent market at a price above expected valuation are not Nash equilibria in pure strategies of the game between two homo moralis.*

*Proof.* Take a profile where  $v^e < \bar{p}_i^* = p_{iQ}^*$ ; for all  $i \in \{1, 2\}$ , and  $Q \in \{h, l\}$ . Then, without loss of generality, Player 1 can profitably deviate to  $\bar{p}'_1 < \bar{p}_2^*$  and  $p'_i = p'_{1l} = p'_{1h} = \bar{p}'_1$ , where  $p^* - \bar{p}'_1 = \epsilon > 0, \epsilon \rightarrow 0$ . There exists  $\epsilon > 0$  sufficiently small such that  $p^* - \bar{p}'_1 < v^e - p^*$ . Therefore, the first term of her utility function increases while the second one remains the same. □

**Corollary C.11.1.** *If  $v^e < r_h$ , there are no equilibria in pure strategies in the bilateral trade game with quality uncertainty and desirable trade between two moralis.*

**Lemma C.12.** *Consider the bilateral trade game with adverse selection between two homo moralis and assume  $0 < r_l < v_l < r_h < v_h$  and  $r_h \leq v^e$ . Then, symmetric full trade profiles with prices between high quality cost and expected valuation are equilibria.*

*Proof.* Consider, without loss of generality, Player 1. Given Player 2's strategy, both terms of her utility function are maximised. The strategies are then mutual best responses and thus the profile is a Nash equilibrium. □

## Appendix D Appendix to Section 4.2

### Proof of Proposition 4.4

The prices set by a given player are completely independent of her threshold and only affect her utility through their interaction with her rival's threshold. Thus, the bargaining game's equilibrium set is all the pairs of profiles that can be formed from the equilibria of the contingent game.

Consider now the contingent game. The seller's best reply correspondences are:

$$p_h^{BR} = \operatorname{argmax}_{p_h} U^S(p_h, p_l; \bar{p}) = \begin{cases} p_h^{BR} = \bar{p} & \text{if } \bar{p}_1 \geq r_h \\ p_h^{BR} > \bar{p} & \text{if } \bar{p} < r_h \end{cases}$$

$$p_l^{BR} = \operatorname{argmax}_{p_l} U^S(p_h, p_l; \bar{p}) = \begin{cases} p_l^{BR} = \bar{p} & \text{if } \bar{p} \geq r_l \\ p_l^{BR} > \bar{p} & \text{if } \bar{p} < r_l \end{cases}$$

The buyer's best reply correspondence is:

$$\left\{ \begin{array}{ll} \bar{p}^{BR} \geq \max\{p_h, p_l\} & \text{if } 0 \leq p_l \leq v_l, 0 \leq p_h \leq v_h \\ \bar{p}^{BR} \geq p_h & \text{if } v_l < p_l \leq v_e, p_l \leq p_h \leq v_h + \frac{1-\lambda}{\lambda}(v_l - p_l) \\ \bar{p}^{BR} \in [p_l, p_h) & \text{if } 0 \leq p_l \leq v_l, v_h < p_h \\ \bar{p}^{BR} \in [p_h, p_l) & \text{if } v_l < p_l, 0 \leq \min\{p_l, v_h\} \\ \bar{p}^{BR} < \min\{p_h, p_l\} & \text{if } v_l < p_l, p_h > \min\left\{\max\left\{v_h + \frac{1-\lambda}{\lambda}(v_l - p_l), p_l\right\}, v_h\right\} \end{array} \right.$$

If  $\lambda \geq \lambda_e$ , the best replies intercept whenever:

$$\bar{p}^* = p_h^* = p_l^* \in [r_h, v_e], \text{ and}$$

$$\bar{p}^* < r_l, p_l^* > v_l, p_h^* > \min\left\{\max\left\{p_l^*, v_h + \frac{1-\lambda}{\lambda}(v_l - p_l^*)\right\}, v_h\right\}.$$

If  $\lambda < \lambda_e$ , they do so only when:

$$\bar{p}^* < r_l, p_l^* > v_l, p_h^* > \min\left\{\max\left\{p_l^*, v_h + \frac{1-\lambda}{\lambda}(v_l - p_l^*)\right\}, v_h\right\}.$$

**Lemma D.1.** *Consider the quality uncertainty game between two homo moralis and assume that  $0 < v_l < r_l < r_h < v_h$ . Then, no-trade profiles where the weakly largest threshold is below own high quality price or above own bad quality price are not Nash equilibria in pure strategies.*

*Proof.* Assume that  $\bar{p}_i^* < \min\{p_{jh}^*, p_{jl}^*\} \forall i, j \in \{1, 2\}, i \neq j$  and, without loss of generality, that  $\bar{p}_2^* \leq \bar{p}_1^*$ . Then:

(1) Suppose additionally that  $\bar{p}_1^* < p_{1h}^*$ . If  $\bar{p}_2^* < \bar{p}_1^*$ , Player 1 can profitably deviate to  $p'_{1h} = \bar{p}_1^*$ . In turn, if  $\bar{p}_2^* = \bar{p}_1^*$ , she can profitably deviate to  $\bar{p}'_{1h} = p'_1 \in (\bar{p}_1^*, \min\{p_{1l}^*, p_{2h}^*, p_{2l}^*\})$ .

In both deviations, the first term of her utility function does not change while the second one strictly increases.

(2) Suppose, alternatively to (1), that  $p_{1l}^* \leq \bar{p}_1^*$ . Then, for  $\bar{p}_2^* < \bar{p}_1^*$ , Player 1 can profitably deviate to  $p'_{1l} > \bar{p}_1^*$ . The first term of her utility function does not change while the second one strictly increases.  $\square$

**Lemma D.2.** *Consider the quality uncertainty game between two homo moralis and assume that  $0 < v_l < r_l < r_h < v_h$ . Then, profiles where there is no trade in either contingent market are not Nash equilibria in pure strategies.*

*Proof.* Assume that  $\bar{p}_i^* < \min\{p_{jh}^*, p_{jl}^*\} \forall i, j \in \{1, 2\}, i \neq j$  and, without loss of generality that  $\bar{p}_1^* \leq \bar{p}_2^*$ . First, consider the case where  $\bar{p}_1^* < \bar{p}_2^*$ . Then, by Lemma D.1, only profiles where  $p_{2h}^* \leq \bar{p}_2^* < p_{2l}^*$  can be candidate equilibria. The profiles left when  $\bar{p}_1^* < \bar{p}_2^*$  are then those where  $\bar{p}_1^* < p_{2h}^* \leq \bar{p}_2^* < \min\{p_{2l}^*, p_{1h}^*, p_{1l}^*\}$ . However, notice that if  $p_{2h}^* \leq v_h$ , then Player 1 finds it (weakly) profitable to increase her own threshold to  $\bar{p}'_1 = p_{2h}^*$ . After this deviation, the only change is that Player 1 now acquires Player 2's high quality good at price  $p_{2h}^* \leq v_h$ , without buying the bad quality item. The first term of her utility function increases while the second one does not change. Thus, no-trade profiles featuring  $\bar{p}_1^* < \bar{p}_2^*$  and  $\bar{p}_2^* \leq v_h$  are not equilibria.

It remains to check whether profiles such that  $\bar{p}_1^* < \bar{p}_2^*$  and  $p_{2h}^* > v_h$  are equilibria. The answer is also negative because we have that  $p_{2h}^* \leq \bar{p}_2^* \implies \bar{p}_2^* > v_h > r_h$ . With Player 2's threshold above  $r_h$ , Player 1 can profitably deviate by lowering her own high quality price to  $p'_{1h} = \bar{p}_2^*$  and sell her good quality item for a profit, without experiencing neither "moral" losses nor gains, as both her prices remain above her threshold. The first term of her utility function increases while the second one does not change. I can then affirm that no-trade profiles where players set different thresholds are not equilibria of this game.

Finally, suppose that  $\bar{p}_1^* = \bar{p}_2^*$ . Player 1 can profitably deviate to  $p'_{1h} = \bar{p}'_1 \in (\bar{p}_2^*, \min\{p_{1l}^*, p_{2l}^*, p_{2h}^*\})$ . In this way, she gets to sell the high quality good to "herself" without modifying anything else, thus obtaining a "moral" gain. The first term of her utility function does not change while the second one strictly increases.  $\square$

**Lemma D.3.** *Consider the quality uncertainty game between two homo moralis and assume that  $0 < v_l < r_l < r_h < v_h$ . Then, profiles with trade of both qualities in one contingent market and no trade in the other one where agents setting different thresholds are not Nash equilibria in pure strategies.*

*Proof.* Assume without loss of generality that  $\bar{p}_1^* < \min\{p_{2h}^*, p_{2l}^*\}$ ,  $\bar{p}_2^* \geq \max\{p_{1h}^*, p_{1l}^*\}$  and  $\bar{p}_1^* \neq \bar{p}_2^*$ . Then:

(1) Suppose first that  $\bar{p}_2^* < \bar{p}_1^*$ . This implies that  $\bar{p}_2^* < \min\{p_{2h}^*, p_{2l}^*\}$ . But then, Player 2 can profitably deviate to  $\bar{p}'_2 < \bar{p}_2^* = p'_{2h} < p'_{2l}$ . Her utility function's first term remains unchanged while the second one increases.

(2) Assume now that  $\bar{p}_1^* < \bar{p}_2^*$ . Notice first that a profile complying with these assumptions and also with  $p_{1l}^* < \bar{p}_2^*$  cannot be an equilibrium. Indeed, Player 1 can profitably deviate to

$p'_{1l} = \bar{p}_2^*$  and make a material improvement while experiencing no moral losses, or even gains (her utility function's first term increases while the second one does not decrease). Next, remark that profiles where  $p_{2l}^* \leq \bar{p}_2^*$  or  $p_{2h}^* > \bar{p}_2^*$  are not equilibria. In the former case, Player 2 can  $p'_{2l} > \bar{p}_2^*$  whilst in the latter she can do so to  $p'_{2h} \leq \bar{p}_2^*$ . The deviations entail a moral gain without material losses (her utility function's first term remains unchanged while the second one increases).

We are thus left with profiles where  $\bar{p}_1^* < p_{2h}^* \leq \bar{p}_2^* = p_{1l}^* < p_{2l}^*$  and  $p_{1h}^* \leq \bar{p}_2^*$ . No profiles of this kind are equilibria if  $p_{1h}^* < \bar{p}_2^*$ . Player 1 can profitably deviate to  $p'_{1h} = \bar{p}'_1 = p_{2h}^* - \epsilon$ . She makes a material gain and sustains no moral losses.

Profiles with  $p_{1h}^* = \bar{p}_2^*$  are not equilibria either. Consider first the case where  $p_{1h}^* = \bar{p}_2^* \leq v_h$ . Then we must also have  $p_{2h}^* = \bar{p}_2^*$  since if  $p_{2h}^* < \bar{p}_2^*$ , Player 1 can deviate to  $\bar{p}'_1 = p_{2h}^*$  and make a material gain. Therefore, the remaining candidates are  $p_{1l}^* = p_{1h}^* = p_{2h}^* = \bar{p}_2^* < p_{2l}^*$ . These are not equilibria, as Player 1 can deviate to  $p'_{1h} = \bar{p}'_1 = \bar{p}_2^* - \epsilon$ . She makes a moral gain and sustains a material loss that can be made smaller by choosing the appropriate  $\epsilon$ . Assume in turn that  $p_{1h}^* = \bar{p}_2^* > v_h$ . Then, Player 2 can deviate to  $\bar{p}'_2 = p'_{2h} \in (\bar{p}_1^*, \bar{p}_2^*)$ . She makes a material gain while avoiding moral losses. □

**Lemma D.4.** *Consider the quality uncertainty game between two homo moralis and assume that  $0 < v_l < r_l < r_h < v_h$ . Then, profiles with trade of both qualities in one contingent market, no-trade in the other one and agents setting the same threshold are not Nash equilibria in pure strategies.*

*Proof.* Assume without loss of generality that  $\min\{p_{1h}^*, p_{1l}^*\} \leq \bar{p}_2^* = \bar{p}_1^* < \min\{p_{2h}^*, p_{2l}^*\}$ . Then, Player 2 can profitably deviate to  $(\bar{p}'_2, p'_{2h}, p'_{2l})$  such that  $\bar{p}_2^* = \bar{p}_1^* < p'_{2h} \leq \bar{p}'_2 < p'_{2l}$ . This simple deviation allows her to achieve a ‘moral’ gain by commencing to trade with ‘herself’ the good quality object, thus breaking the equality in thresholds assumed at first. □

**Lemma D.5.** *Consider the quality uncertainty game between two homo moralis and assume that  $0 < v_l < r_l < r_h < v_h$ . Then, profiles where only the bad quality is exchanged in at least one contingent market at a price strictly below the threshold are not Nash equilibria in pure strategies.*

*Proof.* Assume without loss of generality that  $p_{1l}^* < \bar{p}_2^* < p_{1h}^*$ . Then, Player 1 can profitably deviate to  $p'_{1l} = \bar{p}_2^*$ . She makes a material gain without suffering moral losses. □

**Lemma D.6.** *Consider the quality uncertainty game between two homo moralis and assume that  $0 < v_l < r_l < r_h < v_h$ . Then, profiles where only the bad quality is exchanged in at least one contingent market at a price above the low quality valuation are not Nash equilibria in pure strategies.*

*Proof.* Assume without loss of generality that  $v_l < p_{2l}^* \leq \bar{p}_1^* < p_{2h}^*$ . By Lemma D.5, we can focus on the case  $v_l < p_{2l}^* = \bar{p}_1^* < p_{2h}^*$ . Then:

(1) Suppose in addition that  $p_{1h}^* < \bar{p}_1^*$ . Then Player 1 can profitably deviate to  $\bar{p}'_1 \in (p_{1h}^*, \bar{p}_1^*)$ . She makes a material gain while avoiding a moral loss.

(2) Assume, alternatively, that  $p_{1h}^* = \bar{p}_1^*$ . Then Player 1 can deviate to  $\bar{p}'_1 = \bar{p}_1^* - \epsilon$ . The deviation entails a material gain from stopping the trade of the bad quality item at a price above valuation and a potential loss from reducing her own high quality price. However, the loss can be made smaller than the gain by choosing an appropriate  $\epsilon$ . Since there are no moral losses involved, the deviation is profitable.

(3) Assume, in turn, that  $p_{1h}^* > \bar{p}_1^*$ . Then, Player 1 can profitably deviate to  $\bar{p}'_1 < \bar{p}_1^*$ . She makes a material gain while avoiding a moral loss.  $\square$

**Lemma D.7.** *Consider the quality uncertainty game between two homo moralis and assume that  $0 < v_l < r_l < r_h < v_h$ . Then, profiles where only the bad quality is exchanged in at least one contingent market at prices below low quality cost are not Nash equilibria in pure strategies.*

*Proof.* Assume without loss of generality that  $p_{2l}^* \leq \bar{p}_1^* < \min\{r_l, p_{2h}^*\}$ . By Lemma D.5, we can focus on the case  $p_{2l}^* = \bar{p}_1^* < \min\{r_l, p_{2h}^*\}$ . Then, Player 2 can profitably deviate to  $p'_{2l} > r_l$ . She makes a material gain while avoiding a moral loss.  $\square$

**Corollary D.7.1.** *If we assume  $v_l < r_l$ , profiles where only the bad quality is exchanged in at least one contingent market are not equilibria of this game. Indeed, Lemmas D.5, D.6 and D.7 show that the threshold of the contingent market featuring only trade of the bad quality cannot be neither above  $v_l$  nor below  $r_l$  to be equilibria.*

**Lemma D.8.** *Consider the quality uncertainty game between two homo moralis and assume that  $0 < v_l < r_l < r_h < v_h$ . Then, profiles with only good quality trade in one contingent market and no trade in the other one are not Nash equilibria in pure strategies.*

*Proof.* Assume without loss of generality that  $\bar{p}_1^* < \min\{p_{2h}^*, p_{2l}^*\}$  and  $p_{1h}^* \leq \bar{p}_2^* < p_{1l}^*$ .

(1) Consider the case where, in addition,  $\bar{p}_2^* < \bar{p}_1^*$ . Firstly, Player 1 can profitably deviate from any profile where  $\bar{p}_1^* \geq p_{1l}^*$  to  $p'_{1l} > \bar{p}_1^*$ . She makes a moral gain without sustaining material losses. Therefore, I can restrict attention to those where it is also true that  $\bar{p}_1^* < p_{1l}^*$ . But then, notice that Player 2 can make a “moral” gain by deviating to  $p'_{2l} > \bar{p}_2^* = p'_{2h} \in (\bar{p}_1^*, p_{1l}^*)$ .

(2) In turn, suppose that  $\bar{p}_1^* < \bar{p}_2^*$ . Firstly, notice that if  $p_{2l}^* \leq \bar{p}_2^*$  Player 2 can profitably deviate to  $p'_{2l} > \bar{p}_2^*$  and make a moral gain without suffering material losses. I then focus on cases where  $p_{2l}^* > \bar{p}_2^*$ . If  $\bar{p}_2^* < r_h$ , Player 1 can profitably deviate to  $p'_{1h} \in (\bar{p}_2^*, \min\{p_{2h}^*, p_{2l}^*\})$ . She makes a material gain without sustaining moral losses. If  $\bar{p}_2^* \geq r_h$ , Player 1 can profitably deviate to  $p'_{1l} = \bar{p}_2^*$ . Again, she makes a material gain without sustaining moral losses.

(3) Assume now that  $\bar{p}_2^* = \bar{p}_1^*$ . Then, Player 2 can profitably deviate to  $p'_{2l} > \bar{p}'_2 = p'_{2h} \in (\bar{p}_1^*, p_{1l}^*)$  and  $p'_{2h} = \bar{p}'_2$ . she makes a material gain without sustaining moral losses.  $\square$

**Lemma D.9.** *Consider the quality uncertainty game between two homo moralis and assume that  $0 < v_l < r_l < r_h < v_h$ . Then, only profiles where players set the same high quality price and threshold remain as equilibrium candidates when there is only good quality trade in both contingent markets.*

*Proof.* Assume without loss of generality that  $\bar{p}_1^* < \bar{p}_2^*$ . Then, Player 2 can always profitably deviate from a profile where  $p_{2l}^* \leq \bar{p}_2^*$  to  $p'_{2l} > \bar{p}_2^*$ , making a moral loss without suffering material losses. Therefore, I consider only profiles where  $\bar{p}_2^* < p_{2l}^*$ .

Next, notice that Player 1 can profitably deviate from any profile where  $p_{1h}^* < \bar{p}_2^*$  by setting  $\bar{p}'_1 = p'_{1h} = \bar{p}_2^*$ . She makes a material gain without experiencing moral losses. As a consequence, I further restrict attention to profiles where  $p_{2h}^* \leq \bar{p}_1^* < p_{1h}^* = \bar{p}_2^* < \min\{p_{2l}^*, p_{1l}^*\}$ . Furthermore, Player 2 can profitably deviate whenever  $p_{2h}^* < \bar{p}_1^* < p_{1h}^* = \bar{p}_2^* < \min\{p_{2l}^*, p_{1l}^*\}$  to  $p'_{2h} = \bar{p}_1^*$ , which leaves only  $p_{2h}^* = \bar{p}_1^* < p_{1h}^* = \bar{p}_2^* < \min\{p_{2l}^*, p_{1l}^*\}$ . But notice that Player 1 can then deviate to  $\bar{p}'_1 = p_{1h}^* = \bar{p}_2^*$  and obtain a material gain without sustaining moral losses.

It follows that out of all possible profiles complying with  $p_{ig}^* \leq \bar{p}_j^* < p_{ib}^* \forall i, j \in \{1, 2\}, i \neq j$ , the only equilibrium candidates are those where  $p_{ig}^* = \bar{p}_j^* = p_{jh}^* = \bar{p}_i^* < \min\{p_{ib}^*, p_{jl}^*\}$ .  $\square$

**Lemma D.10.** *Consider the quality uncertainty game between two homo moralis and assume that  $0 < v_l < r_l < r_h < v_h$ . Then, only profiles where players set the same high quality price and threshold and these are between high quality cost and valuation remain as equilibrium candidates when there is only good quality trade in both contingent markets.*

*Proof.* Define  $p^* : p^* = p_{jh}^* = \bar{p}_i^* = p_{ig}^* = \bar{p}_j^* \forall i, j \in \{1, 2\}, i \neq j$ . Suppose first that  $v_h < p^* < \min\{p_{1l}^*, p_{2l}^*\}$ . Then, without loss of generality, Player 1 can profitably deviate to  $\bar{p}'_1 = p'_{1h} < p^* - \epsilon$ . She makes a material loss (she sells the high quality item at a lower price) that can be made smaller than the material gain by choosing an appropriate  $\epsilon$ . In addition, she does not make any “moral” losses.

In turn, assume that  $p^* < \min\{r_h, p_{2l}^*, p_{1l}^*\}$ . Without loss of generality, Player 1 can profitably deviate by setting  $\bar{p}'_1 = p'_{1h} \in (p^*, \min\{p_{2l}^*, p_{1l}^*\})$ . She obtains a material gain without sustaining moral losses.  $\square$

**Lemma D.11.** *Consider the quality uncertainty game between two homo moralis and assume that  $0 < v_l < r_l < r_h < v_h$ . Define  $p^* : p^* = p_{ig}^* = \bar{p}_j^* = p_{jh}^* = \bar{p}_i^* \forall i, j \in \{1, 2\}, i \neq j$ . Then, profiles where  $r_h \leq p^* \leq \min\left\{(1 - \lambda)r_l + \lambda v_h, \frac{r_l - \kappa v_l}{1 - \kappa}\right\}$ ,  $p^* < \min\{p_{ib}^*, p_{jl}^*\}$  are Nash equilibria in pure strategies if and only if:*

$$\lambda v_h + (1 - \lambda)r_l \geq r_h \iff \lambda \geq \frac{r_h - r_l}{v_h - r_l}, \text{ and}$$

$$\frac{r_l - \kappa v_l}{1 - \kappa} \geq r_h \iff \kappa \geq \frac{r_h - r_l}{r_h - v_l}.$$

*Proof.* Consider a profile where  $r_h \leq p^* \leq v_h$  and  $p^* < \min\{p_{1l}^*, p_{2l}^*\}$ . Take, without loss of generality, Player 1. She has only two potentially profitable deviations available. On the one hand, she can deviate to  $\bar{p}'_1 = p^*$ . This allows her to sell the bad quality object to Player 1 at a price above cost  $r_l$ , but also entails a “moral” loss as she would trade the bad quality item with “herself”.

On the other hand, she can deviate to  $\bar{p}'_1 = p^*$  and  $p'_{1h} = \bar{p}'_1 = p^* - \epsilon$ . This allows her to sell the low quality good to Player 2 at a price above cost while not selling it to “herself”

(avoiding moral losses). In contrast, she stops buying the good quality item from Player 2 at a price below her valuation (because she set a threshold below the latter's price), thus making a material loss. Finally, there is also a material loss that stems from lowering the high quality price but this can be made smaller than the material gain by choosing an appropriate value for  $\epsilon$ .

Utility at the initial profile is:

$$U^{hm}((p^*, p_{1l}^*, p^*), (p^*, p_{2l}^*, p^*)) = \frac{1-\kappa}{2} \left[ \lambda(v_h - p^*) + \lambda(p^* - r_h) \right] + \frac{\kappa}{2} \left[ \lambda(v_h - r_h) \right]. \quad (15)$$

Meanwhile, utilities obtained at respectively the first and second deviations described are:

$$U^{hm}((p^*, p^*, p^*), (p^*, p_{2l}^*, p^*)) = \frac{1-\kappa}{2} \left[ \lambda(v_h - p^*) + \lambda(p^* - r_h) + (1-\lambda)(p^* - r_l) \right] + \frac{\kappa}{2} \left[ \lambda(v_h - r_h) + (1-\lambda)(v_l - r_l) \right], \quad (16)$$

$$U^{hm}((p'_{1h}, p^*, \bar{p}'_1), (p^*, p_{2l}^*, p^*)) = \frac{1-\kappa}{2} \left[ \lambda(p'_{1h} - r_h) + (1-\lambda)(p^* - r_l) \right] + \frac{\kappa}{2} \left[ \lambda(v_h - r_h) \right]. \quad (17)$$

Comparing (15) with (16), we conclude that the first deviation is profitable if and only if:

$$p^* > \frac{r_l - \kappa v_l}{1 - \kappa}. \quad (18)$$

Meanwhile, the comparison of (15) with (17) yields that the second one is beneficial if and only if:

$$p^* > r_l + \frac{\lambda}{1-\lambda}(v_h - p'_{1h}). \quad (19)$$

By taking into account that  $p'_{1h}$  can be made arbitrarily close to  $p^*$  we can take the limit of the expression at the right-hand side of (19) and get:

$$\lim_{p'_{1h} \rightarrow p^*} r_l + \frac{\lambda}{1-\lambda}(v_h - p'_{1h}) = r_l + \frac{\lambda}{1-\lambda}(v_h - p^*). \quad (20)$$

The second deviation is then profitable if and only if:

$$p^* > (1-\lambda)r_l + \lambda v_h. \quad (21)$$

Therefore, the game has equilibria in pure strategies where only the good quality is traded in both contingent markets if and only if:

$$\lambda v_h + (1-\lambda)r_l \geq r_h \iff \lambda \geq \frac{r_h - r_l}{v_h - r_l}, \text{ and} \quad (22)$$

$$\frac{r_l - \kappa v_l}{1 - \kappa} \geq r_h \iff \kappa \geq \frac{r_h - r_l}{r_h - v_l}, \quad (23)$$

They are characterised by:

$$r_h \leq p^* = p_{1h}^* = p_{2h}^* = \bar{p}_2^* = \bar{p}_1^* \leq \min \left\{ (1-\lambda)r_l + \lambda v_h, \frac{r_l - \kappa v_l}{1 - \kappa} \right\} \text{ and } p^* < \min \{ p_{2l}^*, p_{1l}^* \}.$$

□



**Lemma D.12.** *Consider the quality uncertainty game between two homo moralis and assume that  $0 < v_l < r_l < r_h < v_h$ . Then, profiles with trade of the good quality in one contingent market and full trade in the other one are not equilibria if players set different thresholds and prices for the traded qualities.*

*Proof.* Assume without loss of generality, that  $p_{2h}^* \leq \bar{p}_1^* < p_{2l}^*$ ,  $\max\{p_{1h}^*, p_{1l}^*\} \leq \bar{p}_2^*$ .

(1) Suppose first that in addition to the above,  $\bar{p}_2^* < \bar{p}_1^*$ . Then Player 2 could always deviate profitably from profiles where  $p_{2h}^* < \bar{p}_1^*$  just by setting  $p'_{2h} = \bar{p}'_2 = \bar{p}_1^*$ . This is enough to discard profiles with  $\bar{p}_2^* < \bar{p}_1^*$ .

(2) Assume in turn, that  $\bar{p}_1^* < \bar{p}_2^*$ . No profile where  $p_{2l}^* \leq \bar{p}_2^*$  can be an equilibrium, as Player 2 can deviate to  $\bar{p}_2^* < p'_{2l}$  and obtain a moral gain without suffering material losses. Therefore, I can circumscribe to profiles where  $p_{2l}^* > \bar{p}_2^*$ . Moreover,  $p_{1l}^* < \bar{p}_2^*$  cannot be part of an equilibrium, as Player 1 can profitably deviate to  $p'_{1l} = \bar{p}_2^*$ . She obtains a moral gain without suffering material losses. So we are left with  $p_{2h}^* \leq \bar{p}_1^* < \bar{p}_2^* = p_{1l}^* < p_{2l}^*$  and  $p_{1h}^* \leq \bar{p}_2^*$ . If  $p_{1h}^* < \bar{p}_2^*$ , Player 1 can profitably deviate by setting  $p'_{1h} = \bar{p}'_1 \in (p_{1h}^*, \bar{p}_2^*)$ . She makes a material gain without suffering moral losses. In turn, if  $p_{1h}^* = \bar{p}_2^*$ , she can deviate to  $p'_{1h} = \bar{p}'_1 = \bar{p}_2^* - \epsilon$ . She experiences a strictly positive moral gain and a material loss. However, the latter can be made smaller than the former by choosing an appropriate  $\epsilon$ . Then, the thresholds cannot be different.

(3) Finally, suppose that  $\bar{p}_1^* = \bar{p}_2^*$  but  $p_{1h}^*$ ,  $p_{1l}^*$  or  $p_{2h}^*$  are strictly below  $\bar{p}_1^* = \bar{p}_2^*$ . Then, the agent can just increase her price up to the thresholds and make a material gain while sustaining no moral losses.  $\square$

**Corollary D.12.1.** *A necessary condition for the existence of equilibria with trade of the good quality in one contingent market and full trade in the other one is that  $\lambda v_h + (1 - \lambda)v_l \geq r_h \iff \lambda \geq \lambda_e$ .*

**Lemma D.13.** *Consider the quality uncertainty game between two homo moralis and assume that  $0 < v_l < r_l < r_h < v_h$ . Take all profiles with trade of the good quality in one contingent market and full trade in the other one where agents set the same thresholds and prices for traded qualities. Then, only those with prices above high quality cost and below expected valuation remain as equilibrium candidates.*

*Proof.* By Lemma D.13, we can focus on profiles where  $p_{ig}^* = p_{ib}^* = p_{jh}^* = \bar{p}_i^* = \bar{p}_j^* < p_{jl}^*$ ,  $i, j \in \{1, 2\}, i \neq j$ .

Assume without loss of generality, that  $p_{1h}^* = p_{1l}^* = p_{2h}^* = \bar{p}_1^* = \bar{p}_2^* < p_{2l}^*$  and define  $p^* = p_{1h}^* = p_{1l}^* = p_{2h}^* = \bar{p}_1^* = \bar{p}_2^*$ .

Suppose first that  $p^* < r_h$ . Then, without loss of generality, Player 1 can deviate to  $p'_{1h} = \bar{p}'_1 \in (p^*, p_{2l}^*)$ . She makes a material gain and avoids moral losses.

In turn, assume that  $p^* > v^e$ . Then, Player 2 can deviate to  $p'_{2h} = \bar{p}'_2 = \bar{p}_1^* - \epsilon$ . She makes a material gain from ceasing to buy both qualities at a price above the expected valuation and a loss from reducing her price, while experiencing no moral losses. Since the loss can be made smaller than the gain by choosing an appropriate  $\epsilon$ , the deviation is profitable.  $\square$

**Lemma D.14.** *Consider the quality uncertainty game between two homo moralis and assume that  $0 < v_l < r_l < r_h < v_h$ . Then, profiles with full trade in one contingent market and trade of only the good quality in the other one if and only if:*

$$\lambda \geq \lambda_2$$

$$\kappa \in [\kappa_1, \kappa_a(\lambda)]$$

*They are characterised by:*

$$\frac{r_l - \kappa v_l}{1 - \kappa} = p_{ig}^* = p_{jh}^* = p_{ib}^* = \bar{p}_j^* = \bar{p}_i^* < p_{jl}^* \text{ for } i, j \in \{1, 2\}, i \neq j.$$

*Proof.* (1) By Lemmas D.12 and D.13, the remaining candidate equilibria are such that  $r_h \leq p^* = \bar{p}_i^* = \bar{p}_j^* = p_{jh}^* = p_{ig}^* = p_{ib}^* \leq v_e, p^* < p_{jl}^* \forall i, j \in \{1, 2\}, i \neq j$ . Corollary D.12.1 establishes that  $\lambda \geq \lambda_e$  is a necessary condition for the existence of this kind of equilibria.

(2) Assume, without loss of generality, that  $r_h \leq p^* = \bar{p}_1^* = \bar{p}_2^* = p_{2h}^* = p_{1h}^* = p_{1l}^* \leq v_e, p^* < p_{2l}^*$ . Player 1's utility at this kind of profiles is specified in Equation (24), while that of Player 2 in (25).

$$U_1^{hm}((p^*, p^*, p^*), (p^*, p_{2l}^*, p^*)) = \frac{1 - \kappa}{2} \left[ \lambda(v_h - p^*) + \lambda(p^* - r_h) + (1 - \lambda)(p^* - r_l) \right] + \frac{\kappa}{2} \left[ \lambda(v_h - r_h) + (1 - \lambda)(v_l - r_l) \right], \quad (24)$$

$$U_2^{hm}((p^*, p_{2l}^*, p^*), (p^*, p^*, p^*)) = \frac{1 - \kappa}{2} \left[ \lambda(v_h - p^*) + (1 - \lambda)(v_l - p^*) + \lambda(p^* - r_h) + (1 - \lambda)r_l \right] + \frac{\kappa}{2} \left[ \lambda(v_h - r_h) \right]. \quad (25)$$

Player 1 may deviate from the profiles described in two potentially beneficial ways: either by marginally reducing her threshold and high quality price or by increasing her low quality price. The first deviation leads Player 1 to stop trading the bad quality item with “herself”, but also to stop buying the high quality good from Player 2 whilst also selling her own good quality object for a slightly lower price. In turn, the second one leads to her not selling the low quality good to Player 2 but avoiding the exchange of this same item with “herself”. The utility achieved by Player 1 after the first deviation is described in Expression (26), where  $p'_{1h}$  is Player 1's high quality price and threshold after deviating. Meanwhile, Player 1's utility at to the second deviation is shown in Equation (27), with  $p'_{1l}$  representing the low quality price set by Player 1 at the deviation.

$$U_1^{hm}((p'_{1h}, p^*, p'_{1h}), (p^*, p^*, p^*)) = \frac{1-\kappa}{2} \left[ \lambda(p'_{1h} - r_h) + (1-\lambda)(p^* - r_l) \right] + \frac{\kappa}{2} \left[ \lambda(v_h - r_h) \right], \quad (26)$$

$$U_1^{hm}((p^*, p'_{1l}, p^*), (p^*, p^*, p^*)) = \frac{1-\kappa}{2} \left[ \lambda(v_h - p^*) + \lambda(p^* - r_h) \right] + \frac{\kappa}{2} \left[ \lambda(v_h - r_h) \right]. \quad (27)$$

The first deviation is profitable whenever

$$(1-\kappa)(\lambda(v_h - r_h)) + \kappa(1-\lambda)(v_l - r_l) > (1-\kappa)\lambda(p'_{1h} - r_h).$$

Recall that the price reduction from  $p^*$  to  $p'_{1h}$  is a marginal change and Player 1 can make them arbitrarily similar. Taking the limit as  $p'_{1h}$  approaches  $p^*$  of the right hand side in the above expression and solving for  $p^*$  gives:

$$p^* > v_h + \frac{\kappa(1-\lambda)}{(1-\kappa)\lambda}(v_l - r_l) = A. \quad (28)$$

In turn, the second deviation is beneficial to Player 1 whenever:

$$p^* < \frac{r_l - \kappa v_l}{1-\kappa} = B. \quad (29)$$

I now turn to deviations available to Player 2. Firstly, she can reduce her low quality price to  $p^*$ . In this way, she gets to sell her low quality good to Player 1 at price  $p^*$  but also exchanges it with ‘herself’, making a ‘moral’ loss. Secondly, she can, in addition to lowering her low quality good to  $p^*$ , slightly reduce her threshold and high quality price to  $p'_{2h}$ . Thus, she gets to sell her low quality good to Player 1 at price  $p^*$  and avoid trading it with ‘herself’. She also avoids buying Player 1’s low quality object. In contrast, she stops buying the good quality item from Player 1. Player 2’s utility after the first deviation is shown in (30), while the second one is displayed in (31).

$$U_2^{hm}((p^*, p^*, p^*), (p^*, p^*, p^*)) = \frac{1-\kappa}{2} \left[ \lambda(v_h - p^*) + (1-\lambda)(v_l - p^*) + \lambda(p^* - r_h) + (1-\lambda)(p^* - r_l) \right] + \frac{\kappa}{2} \left[ \lambda(v_h - r_h) + (1-\lambda)(v_l - r_l) \right], \quad (30)$$

$$U_2^{hm}((p'_{2h}, p^*, p'_{2h}), (p^*, p^*, p^*)) = \frac{1-\kappa}{2} \left[ \lambda(p'_{2h} - r_h) + (1-\lambda)(p^* - r_h) \right] + \frac{\kappa}{2} \left[ \lambda(v_h - r_h) \right]. \quad (31)$$

The first deviation is profitable to Player 2 whenever:

$$p^* > \frac{r_l - \kappa v_l}{1 - \kappa} = B. \quad (32)$$

In turn, the second one is beneficial to Player 2 when:

$$p^* > \frac{v^e + (1 - \lambda)r_l}{2 - \lambda} = C. \quad (33)$$

From expressions (29) and (32), the only possible value for  $p^*$  in this type of equilibria is:

$$p^* = \frac{r_l - \kappa v_l}{1 - \kappa}.$$

From (28) and (33) we obtain two new conditions for the existence of these equilibria:

$$\begin{aligned} \frac{r_l - \kappa v_l}{1 - \kappa} &\leq v_h + \frac{\kappa(1 - \lambda)}{(1 - \kappa)\lambda}(v_l - r_l), \\ \frac{r_l - \kappa v_l}{1 - \kappa} &\leq \frac{v^e + (1 - \lambda)r_l}{2 - \lambda}. \end{aligned}$$

(3) Putting together all of the above, we have equilibria with full trade in one contingent market and trade of only the good quality in the other one if and only if:

$$B \in [r_h, \min \{A, C, v^e\}]. \quad (34)$$

(4) Notice that:

$$r_h \leq v^e \implies C \leq v^e.$$

(5) We can also show that:

$$B \leq C \implies C \leq A.$$

Firstly,

$$B \leq C \iff \kappa \leq \frac{\lambda(v_h - v_l) + v_l - r_l}{\lambda(v_h - r_l) + r_l - v_l} = \kappa_a(\lambda).$$

Secondly,

$$C \leq A \iff \kappa \leq \frac{2\lambda v_h - \lambda(r_l + v_l)}{2(\lambda(v_h - r_l) + r_l - v_l)} = \kappa_c(\lambda).$$

Finally,  $\kappa_a(\lambda) \leq \kappa_c(\lambda) \iff \lambda \leq 2$ , which is true by definition.

(6) So we can re-write (34) as:

$$B \in [r_h, C] \iff \kappa \in \left[ \frac{r_h - r_l}{r_h - v_l}, \kappa_a(\lambda) \right]. \quad (35)$$

(7) The condition for the above interval to be non-empty is:

$$\lambda \geq \frac{r_h - r_l + r_h - v_l}{r_h - r_l + v_h - v_l}, \quad (36)$$

which is larger than  $\frac{r_h - v_l}{v_h - v_l}$  (the value of  $\lambda$  necessary for  $r_h \leq v^e$ ).  $\square$

**Lemma D.15.** *Consider the quality uncertainty game between two homo moralis and assume that  $0 < v_l < r_l < r_h < v_h$ . Then, of all profiles with full trade, only those where players set the same thresholds and prices remain as candidate Nash equilibria in pure strategies.*

*Proof.* Assume without loss of generality, that  $\bar{p}_1^* < \bar{p}_2^*$ . Then:

(1) Profiles where  $p_{1l}^* < \bar{p}_2^*$  can immediately be discarded as candidate equilibria, since Player 1 can profitably deviate to  $p'_{1l} = \bar{p}_2^*$  and make a material benefit without suffering moral losses.

(2) Within profiles where  $p_{1l}^* = \bar{p}_2^*$ , those where  $p_{1h}^* < \bar{p}_2^*$  are also not equilibria, as Player 1 can profitably deviate to  $p'_{1h} = \bar{p}_1^* \in (p_{1h}^*, \bar{p}_2^*)$  and make a material benefit without suffering moral losses.

(3) From profiles where  $p_{1h}^* = p_{1l}^* = \bar{p}_2^*$ , Player 1 can profitably deviate to  $\bar{p}'_1 = p'_{1h} = \bar{p}_2^* - \epsilon$ . She makes a moral gain and a material loss. The latter can be made smaller than the former by choosing an appropriate value for  $\epsilon$  and thus, the deviation is profitable.

(4) Finally, profiles where  $p_{2Q}^* < \bar{p}_1^* = p_{1h}^* = p_{1l}^* = \bar{p}_2^*$  for some  $Q \in \{g, b\}$  are not equilibria, as Player 2 can profitably deviate to  $p'_{2Q} = \bar{p}_1^*$ , making a material gain without facing moral losses.  $\square$

The proof of Lemma D.15 entails an interesting remark. Indeed, profiles where  $r_h \leq p_{2l}^* = p_{2h}^* = \bar{p}_1^* < p_{1l}^* = p_{1h}^* = \bar{p}_2^* \leq v^e$  are equilibria in the bargaining game between two *homo oeconomicus*. It is *homo moralis*' "impulse" to set a high quality price she herself would be willing to accept that drives her to deviate, even though the deviation has no material consequences on either player, but only a moral gain for the deviating agent.

I turn now to profiles where both players set the same threshold. Lemma D.17 shows that the only equilibria with full trade are those where players set the same prices and thresholds and characterises them.

**Lemma D.16.** *Consider the quality uncertainty game between two homo moralis and assume that  $0 < v_l < r_l < r_h < v_h$ . Then, full trade profiles with symmetric prices and thresholds are not equilibria if they are below the high quality cost or above the expected valuation.*

*Proof.* Define  $p^* = \bar{p}_1^* = p_{1l}^* = p_{1h}^* = \bar{p}_2^* = p_{2l}^* = p_{2h}^*$ .

(1) Assume that  $p^* < r_h$ . Then, without loss of generality, Player 1 can profitably deviate to  $p'_{1h} = \bar{p}'_1 > p^*$ . She makes a material gain without suffering moral losses.

(2) Assume alternatively that  $v^e < p^*$ . Then, without loss of generality, Player 1 can profitably deviate to  $\bar{p}'_1 = p'_{1h} = p^* - \epsilon$ . She makes no moral losses, makes a material gain from not buying both qualities for a price larger than the expected valuation and a loss from reducing her high quality price. However, the loss can be made smaller than the gain by choosing an appropriate value for  $\epsilon$ .  $\square$

**Lemma D.17.** *Consider the quality uncertainty game between two homo moralis and assume that  $0 < v_l < r_l < r_h < v_h$ . Then, full trade profiles are Nash equilibria of pure strategies if and only if:*

$$\lambda \geq \frac{r_h - v_l}{v_h - v_l},$$

$$\kappa \leq \min \{ \kappa_a(\lambda), \kappa_b(\lambda) \}.$$

They are characterised by:

$$p^* = \bar{p}_i^* = p_{jh}^* = p_{jl}^* \in \left[ \max \left\{ r_h, \frac{r_l - \kappa v_l}{1 - \kappa} \right\}, v^e + \frac{\kappa(1 - \lambda)}{1 - \kappa}(v_l - r_l) \right], \forall i, j \in \{1, 2\}, i \neq j.$$

*Proof.* When considering only profiles where  $r_h \leq p^* \leq v^e$ , there are no unilateral deviations that could prove beneficial for all values of  $\kappa$  and  $\lambda$ . Consider (without loss of generality) Player 1. The first potentially profitable deviation could be for her to increase her low quality price. In this way, she would stop trading the bad quality good with “herself” (thus making a “moral” gain), but she would also stop selling it to Player 2, making a material loss. The second and final one is to slightly reduce her threshold and good quality price. This would allow her to stop trading the bad quality item with “herself” and reap a “moral” benefit. In contrast, she would stop buying from Player 2 at prices below the expected valuation  $v^e$  and would also sell her high quality good for an arbitrarily smaller price. Player 1’s utility at the candidate profile is:

$$U^{hm}((p^*, p^*, p^*), (p^*, p^*, p^*)) = \frac{1 - \kappa}{2} \left[ \lambda(v_h - p^*) + (1 - \lambda)(v_l - p^*) + \lambda(p^* - r_h) + (1 - \lambda)(p^* - r_l) \right] + \frac{\kappa}{2} \left[ \lambda(v_h - r_h) + (1 - \lambda)(v_l - r_l) \right]. \quad (37)$$

Meanwhile, her utility at respectably the first and second deviation is shown in expressions (38) and (39):

$$U^{hm}((p^*, p'_{1l}, p^*), (p^*, p^*, p^*)) = \frac{1 - \kappa}{2} \left[ \lambda(v_h - p^*) + (1 - \lambda)(v_l - p^*) + \lambda(p^* - r_h) \right] + \frac{\kappa}{2} \left[ \lambda(v_h - r_h) \right] \quad (38)$$

$$U^{hm}((p'_{1h}, p^*, \bar{p}'_1), (p^*, p^*, p^*)) = \frac{1 - \kappa}{2} \left[ \lambda(p'_{1h} - r_h) + (1 - \lambda)(p^* - r_l) \right] + \frac{\kappa}{2} \left[ \lambda(v_h - r_h) \right] \quad (39)$$

The first deviation is profitable if and only if:

$$p^* < \frac{r_l - \kappa v_l}{1 - \kappa} = B. \quad (40)$$

In turn, the second deviation is profitable whenever:

$$(1 - \kappa) \left[ \lambda(v_h - p^*) + (1 - \lambda)(v_l - p^*) + \lambda(p^* - r_h) \right] + \kappa(1 - \lambda)(v_l - r_l) < (1 - \kappa)\lambda(p'_{1h} - r_h)$$

Taking the limit as  $p'_{1h}$  approaches  $p^*$  of the right hand side in the above expression and solving for  $p^*$  gives:

$$p^* > v^e + \frac{\kappa(1 - \lambda)(v_l - r_l)}{1 - \kappa} = D. \quad (41)$$

Putting together all of the above, we have full-trade equilibria if and only if:

$$p^* \in [\max\{r_h, B\}, \min\{v^e, D\}]. \quad (42)$$

Delving into (42), first notice that  $D \leq v^e$  for all  $\lambda \in [0, 1]$  and  $\kappa \in (0, 1)$ . As a consequence, we can write it as:

$$p^* \in [\max\{r_h, B\}, D]. \quad (43)$$

As a consequence, a necessary and sufficient condition for these equilibria to exist is that both  $B \leq D$  and  $r_h \leq D$ .

We have that  $B \leq D$  if and only if: □

$$\kappa \leq \frac{\lambda(v_h - v_l) + v_l - r_l}{\lambda(v_h - r_l) + r_l - v_l} = \kappa_a(\lambda), \quad (44)$$

while  $r_h \leq D$ :

$$\kappa \leq \frac{\lambda(v_h - v_l) + v_l - r_h}{\lambda(v_h - r_l) + r_l - r_h} = \kappa_b(\lambda). \quad (45)$$

The function  $\kappa_a(\lambda)$  is continuous and strictly increasing in  $\lambda$ . Its root is at  $\lambda_a = \frac{r_l - v_l}{v_h - v_l} \in (0, 1)$ . For  $\lambda > \lambda_a$ ,  $\kappa_b(\lambda) > 0 \iff \lambda > \lambda_e$ . Finally,  $\lambda_a < \lambda_e$ . Therefore, the necessary and sufficient condition to have  $\kappa_a(\lambda) > 0$  and  $\kappa_b(\lambda) > 0$  is  $\lambda > \lambda_e$ .

Notice that given  $\lambda > \lambda_e$ ,  $\kappa_a(\lambda) \geq \kappa_b(\lambda)$  (respectively,  $\kappa_a(\lambda) < \kappa_b(\lambda)$ ) if and only if  $\lambda \leq \lambda_2$  (respectively,  $\lambda > \lambda_2$ ).

We can thus conclude with two necessary and sufficient conditions for the existence of equilibria with full trade:

$$\lambda \geq \lambda_e \text{ and } \kappa \leq \min\{\kappa_a(\lambda), \kappa_b(\lambda)\}.$$

**Proposition D.18.** *Assume that  $0 < v_l < r_l < r_h < v_h$ . Then, the set of Nash equilibria in pure strategies of the game between two homo kantians is  $\{(p_{ig}, p_{ib}, \bar{p}_i) : p_{ig}^* \leq \bar{p}_i < p_{ib}^* \forall i \in \{1, 2\}\}$ .*

*Proof.* Given that  $r_h < v_h$  and  $r_l > v_l$ , the agent's utility function is maximised if and only if  $p_{ig}^* \leq \bar{p}_i < p_{ib}^*$ . □

**Corollary D.18.1.** *The equilibria between two homo kantians always feature trade of the good quality item in at least one contingent market and no trade of the low quality good in the other one.*

Corollary D.18.1 is due to the fact that it will always be true that either  $\bar{p}_1^* \leq \bar{p}_2^*$  or  $\bar{p}_1^* > \bar{p}_2^*$ . As a consequence, these equilibria always feature trade of the good quality item in at least one of the contingent markets and no trade of the low quality good in the other one. I can thus assert that profiles where no player is trading with the other one are not equilibria in this model, and neither are situations where both Players trade both items with each other (full trade).

**Proposition 4.5\*.** *SSS*

## Appendix E Appendix to Section 5

### Proof of Proposition 5.1

The prices set by a given player are completely independent of her threshold and only affect her utility through their interaction with her rival's threshold. Thus, the bargaining game's equilibrium set is all the pairs of profiles that can be formed from the equilibria of the contingent game.

Consider now the contingent game. The buyer's and seller's best reply correspondences are:

$$\bar{p}^{BR} = \operatorname{argmax}_{\bar{p}_1} U^B(\bar{p}; p) = \begin{cases} \bar{p} \geq p & \text{if } (v - p) + \alpha(p - r) \geq 0 \iff \frac{v - \alpha r}{1 - \alpha} = v^{alt} \geq p \\ \bar{p} < p & \text{otherwise} \end{cases}$$

$$p^{BR} = \operatorname{argmax}_p U^S(p; \bar{p}) = \begin{cases} \bar{p} = p & \text{if } (v - p) + \alpha(p - r) \geq 0 \iff \frac{r - \alpha v}{1 - \alpha} = r^{alt} \leq p \\ \bar{p} < p & \text{otherwise} \end{cases}$$

These correspondences intercept whenever  $\bar{p}^* = p^* \in [r^{alt}, v^{alt}]$  or  $\bar{p}^* < r^{alt}, p^* > v^{alt}$ . When  $\alpha \geq \frac{r}{v}$ ,  $r^{alt} \leq 0$  and therefore there does not exist a non-negative threshold that can be lower than  $r^{alt}$ .

### Proof of Proposition 5.2

The prices set by a given player are completely independent of her threshold and only affect her utility through their interaction with her rival's threshold. Thus, the bargaining game's equilibrium set is all the pairs of profiles that can be formed from the equilibria of the contingent game.

The buyer's best reply correspondences in the contingent game are:

$$\bar{p}_{BR}^h = \operatorname{argmax}_{\bar{p}^h} U^B(\bar{p}^l, \bar{p}^h; p) = \begin{cases} \bar{p}_{BR}^h = p & \text{if } p \leq v_h^{alt} \\ \bar{p}_{BR}^h = p & \text{if } p > v_h^{alt} \end{cases}$$

$$\bar{p}_{BR}^l = \operatorname{argmax}_{\bar{p}^l} U^B(\bar{p}^l, \bar{p}^h; p) = \begin{cases} \bar{p}_{BR}^l = p & \text{if } p \leq v_l^{alt} \\ \bar{p}_{BR}^l = p & \text{if } p > v_l^{alt} \end{cases}$$



The seller's best reply correspondence in the contingent game is:

$$\left\{ \begin{array}{ll} p^{BR} > \max\{\bar{p}^h, \bar{p}^l\} & \text{if } \max\{\bar{p}^h, \bar{p}^l\} < r_h^{alt} \\ p^{BR} = \bar{p}^h & \text{if } \bar{p}^h \geq r_h^{alt}, \bar{p}^l \leq \min\{\bar{p}^h, \lambda\bar{p}^h + (1-\lambda)r_l^{alt}\} \\ & \text{or } \bar{p}^h \geq r_e^{alt}, \bar{p}^l \in [\bar{p}^h, \frac{p_j^{h*} - r_h^{alt}}{1-\lambda}] \\ p^{BR} = \bar{p}^l & \text{if } \bar{p}^h \geq r_l^{alt}, \bar{p}^l \in [\frac{p_j^{h*} - r_h^{alt}}{1-\lambda}, \bar{p}^h] \\ & \text{or } \bar{p}^l \geq r_l^{alt}, \bar{p}^h \leq (1-\lambda)\bar{p}^l + \frac{\lambda}{(1-\alpha)}r_h^{alt} \end{array} \right.$$

They intersect at the equilibrium profiles described.