J Bus Econ (2017) 87:125–153
DOI 10.1007/s11573-015-0790-4

ORIGINAL PAPER

Strategic risk in supply chain contract design

Abdolkarim Sadrieh[1] · **Guido Voigt**[1]

Published online: 24 November 2015

Abstract Supply chains facing asymmetric information can either operate in a cooperative mode with information and benefit sharing or can choose a non-co-operative form of interaction and align their incentives via screening contracts. In the cooperative mode, supply chain efficiency can be achieved, but high levels of trust and trustworthiness are required. In the non-cooperative mode, the contract mechanism guarantees a second best supply chain performance, but only if all parties choose their equilibrium strategies without trembles. Experimental evidence, however, shows that both operating modes often fail due to strategic risk. Cooperation is disrupted by deceptive signals and the lack of trust, whereas non-cooperative strategies suffer from persistent out-of-equilibrium behavior. We present two means to reduce strategic risk. First, a punishment mechanism leads to a better matching of trust and trustworthiness and supports the cooperative operating mode. Second, an enforcement of self-selection supports the non-cooperative equilibrium by increasing the attractiveness of screening contracts. We find that supply chain performance can benefit from reduced strategic risk in either operating mode.

Keywords Behavioral operations management · Contracting · Asymmetric information · Punishment

JEL Classification C7 · C72 · D22 · D82 · L210 · M210

Electronic supplementary material The online version of this article (doi:10.1007/s11573-015-0790-4) contains supplementary material, which is available to authorized users.

✉ Guido Voigt
guido.voigt@ovgu.de

Abdolkarim Sadrieh
sadrieh@ovgu.de

[1] Otto-von-Guericke University Magdeburg, Universitätsplatz 2, 39016 Magdeburg, Germany

1 Introduction

Supply chains facing asymmetrically distributed information basically have two opposite means executing their operations. On the one hand, they can choose to operate in a cooperative mode in which private information is communicated truthfully and the benefits are shared. On the other hand, they can choose a non-cooperative form of interaction and align their incentives via complex contract schemes without communicating private information. In the cooperative mode, supply chain efficiency can be achieved, but high levels of trust and trustworthiness are required. In the non-cooperative mode, the contract mechanism guarantees a second best supply chain performance, but only if all parties choose their equilibrium strategies without trembles. However, basically all of the experimental literature on behavior in supply chains shows that neither a perfectly cooperative nor a perfectly non-cooperative outcome can be achieved. Cooperation is frequently shown to be disrupted by deceptive signals or the lack of trust in truthful signals (see Özer et al. 2011; Inderfurth et al. 2013). The non-cooperative equilibrium and profit maximization often fail due to some persistent degree of out-of-equilibrium behavior that may be due to misperceptions, bounded rationality, or social preferences (see Schweitzer and Cachon 2000; Lim and Ho 2007; Bolton and Katok 2008; Katok and Wu 2009; Wu 2013; Kremer et al. 2010; Becker-Peth et al. 2013). In the context of asymmetric information, we define strategic risk as the risk that the supplier faces either concerning the consistency of the buyers' shared information and contract choice behavior or concerning the buyers adherence to payoff maximizing behavior.

In this paper we present an experimental study in which we compare a baseline experimental treatment in supply chain interaction with two other treatments involving reduced strategic risk. In one of the two treatment variations strategic risk is lower due to the introduction of a punishment option that allows suppliers to punish apparently uncooperative buyers. In the other treatment variation, the strategic risk of the supplier is lower than in the baseline due to the fact that buyers' profit-maximization is enforced in screening contracts.

We consider an adverse selection problem in a simple serial supply chain in which the buyer (she) holds private information. The supplier (he) offers either a simple coordinating contract (in the following referred to as "simple contract") or a screening contract (also referred to as "a menu of contracts"). The simple contract is first best, only if the buyer (she) shares her private information truthfully. However, if the buyer strategically misrepresents her private information, the supplier does best by ignoring the signals and offering the screening contract (i.e. the screening contract is inefficient, but second best). While a number of studies have analyzed the effectiveness of different contracts formats in supply chains with asymmetric information (see Özer et al. 2011, 2014; Inderfurth et al. 2013; Spiliotopoulou et al. 2015), this study is the first that examines the endogenous choice of these contract types. This endogenous choice of the contract type allows us to identify the role of strategic risk in the different operational modes. The implications for our research provide valuable insights for supply chain managers concerning contract type decisions.

Knowing that the predictability of the buyer's behavior is a key driver when suppliers choose between contract types, we systematically investigate how the reduction of each type of strategic risk affects the subtle interactions between contract offers, information sharing, and contract choices. First, we provide the supplier with a punishment mechanism that enables him to punish apparently uncooperative behavior and, thus, to substantially reduce the incentives for deception.[1] Second, we enforce the buyers' self-selection once a screening contract was offered by the supplier. This enforcement eliminates the strategic risk that the buyer destroys supplier's payoffs by contract choices that entail a small loss for the buyer but a substantial loss for the supplier.

We find that reducing strategic risk improves supply chain performance in both the cooperative and the non-cooperative mode of operation. With the suppliers' punishment mechanism, buyers' send more truthful signals leading to a better alignment of trust and trustworthiness. The enforcement of self-selection leads to a significantly greater number of screening contract offers which reduces efficiency losses from a misalignment of trust and trustworthiness. Thus, strategic risk management is profitable for the supply chain no matter if using cooperative or non-cooperative contract design.

Our results are highly relevant for the supply chain coordination and contracting literature. Supply chain managers should seek to reduce the variability and increase the predictability of the supply chain partners' behavior in adverse selection environments. Punishment mechanisms—such as terms of contracts that penalize hostile deviations from "business as usual" between supply chain partners—can increase the trustworthiness of communication and can boost performance of supply chains operating under relatively simple contract formats. However, in business interactions lacking punishment options, tailoring contracts to shared information may cause huge losses. In these cases, suppliers should offer screening contracts, but set the incentives for the buyers high enough to guarantee the desired self-selection and to reduce the strategic risk of buyers' choices deviating from profit maximization. Therefore, suppliers may consider adding a slack that increases the cost of deviation for buyers and facilitates behaviorally robust self-selection (see Sect. 3 for a more formal argument that is in line with Laffont and Martimort 2002, Chapter 9.8.1 and Voigt 2015).

2 Literature review

The present study is most closely related to game-theoretic and behavioral work in the field of supply chain coordination via contracts. We review both game-theoretic and behavioral work on contracting under full and asymmetric information.

[1] Apparently uncooperative behavior is defined as an inconsistency between the signaled cost and the selected contract option or if the high cost signal is send significantly too often.

2.1 Game theoretic models on contracting in supply chains

In non-cooperative supply chains, there is a large body of work showing that the less informed supply chain party maximizes his expected profit by offering a sophisticated menu of contracts, i.e., screening contract (see, for example, Corbett et al. 2004, and the references therein). These screening contracts align the incentives of the supply chain members in a way that the holder of the private information reveals her information by the contract choice. Nonetheless, the outcome is inefficient from a supply chain perspective. In contrast, cooperative supply chains may engage in truthful information sharing and trusting information processing. In such a cooperative operating mode, simple coordinating contracts (e.g., two part tariffs) are tailored to the truthfully shared information, resulting in supply chain efficient outcomes (see e.g. Goyal 1977). This cooperative view stresses that communication (e.g. Cachon and Fisher 2000) and trust (Moore 1998; Zaheer et al. 1998) are necessary for successful supply chain management.

2.2 Behavioral studies in full information context

There are numerous studies investigating supply chains that face uncertain demand (e.g. in the newsvendor context) and operate under simple wholesale price contracts. In most of these studies, the wholesale price is an exogenous parameter and is used as a focus variable for identifying decision biases for high- and low-profit situations (see, e.g., Schweitzer and Cachon 2000; Katok and Wu 2009; Kremer et al. 2010). In contrast, Keser and Paleologo (2004) investigate the supply chain behavior in the newsvendor context when the wholesale price is not exogenously determined, but set by the supplier who has full information regarding the distribution of stochastic demand. Becker-Peth et al. (2013) show that wholesale prices in combination with a buy-back component (buy-back contract) can be systematically manipulated in order to account for decision biases of the buyer. Lim and Ho (2007) experimentally investigate the effect of contract design on the inefficiencies resulting from double marginalization in a deterministic demand setting. Their primary focus lies on how the number of price blocks in a quantity discount scheme under full information impacts the supply chain performance.

Note that some of the papers mentioned above study decisions in supply chains without interaction. In these games strategic risk obviously plays no role. However, these studies establish that there is considerable behavioral variance even in the individual decision making setting. In this study, we take the literature a step further by focusing on the question how behavioral variance plays out in a truly interactive setting.

2.3 Behavioral studies in asymmetric information context

Inderfurth et al. (2013) study the impact of information sharing on the supplier's screening contract offers in a dyadic supplier–buyer supply chain. They do not allow for an endogenous choice of the contract type. Kalkanci et al. (2011) present an experimental analysis of the impact of contract complexity under asymmetric demand information in a dyadic supplier–buyer supply chain. They analyze how the supplier sets the price-breaks for an all-unit quantity discount. In contrast to our experimental setup, however, buyers' decisions are automated and strategic risk in the interaction is not considered. Özer et al. (2011, 2014) investigate the interaction of supply chain members given a simple wholesale price contract and asymmetric information. They find that there is partial truth-telling and trust although theory predicts that all communication accounts to no more than cheap talk. In contrast to our experimental setup, the supplier in Özer et al. (2011, 2014) was limited to only offering wholesale price contracts, while we allow an endogenous choice of the contract type.

3 Outline of the model

We briefly review a strategic lot sizing model and refer the interested reader for a broader motivation and more concise derivation of the game theoretic benchmarks to Voigt and Inderfurth (2012) and Voigt (2011).

We assume a serial supply chain consisting of a supplier delivering a product to a buyer. We assume that the buyer's demand is deterministic and constant over time and, without loss of generality, it is standardized to one unit per period. Hence, unit profits equal period profits. The product is sold at an exogenously determined price to the end-customer, i.e., the impact of ordering and pricing decisions on the buyer-customer interface are not considered in the underlying model. The supplier incurs fixed cost, f, per replenishment and therefore prefers high order sizes. In contrast, the buyer prefers low order sizes, because she incurs holding cost h for every unit stored per period. The holding cost may vary from period to period depending on a number of parameters (e.g., cost of capital, handling and storage cost, etc.). Instead of modeling the holding cost explicitly, we assume that the holding cost is a random variable. The distribution of the holding cost is known to both parties, but the actual realization is only known to the buyer. We assume that all external customer demand must be immediately fulfilled by the buyer, i.e., backlogging is not allowed.

The model captures the basic conflict of interest in supply chain management that buyers prefer low order sizes, while suppliers prefer high order sizes (see Corbett and de Groote 2000; Sucky 2006). In case of uncoordinated actions, the buyer's profit maximizing order size is arbitrarily close to zero (i.e., no inventory holding due to just-in-time delivery) and the supplier's profit maximizing order size is infinite (i.e., arbitrarily small fixed cost per period). Thus, both individual optimal order policies are contrary and some form of negotiation has to take place to agree upon acceptable order sizes for both parties.

We assume that the buyer negotiates the terms of delivery and asks the supplier to ship in smaller lots in order to reduce holding cost. The supplier, in turn, tries to induce a higher order size to lower his average cost per unit. The supplier has to take into account the buyer's outside option, i.e., sourcing from an alternative supplier at cost R. The supplier has fixed revenues of Y_s. The buyer has fixed revenues of

$Y_b^{'} = Y_b + Y_s$, i.e., the buyers revenues are a markup on the suppliers fixed revenues.[2]

Note that our research is not strictly limited to this specific operations management environment, but can be generalized to other adverse selection problems. The first essential assumption is that the supplier has convex cost in the contract variables. The second is that the buyer has a linear cost function that is private information. As an example, an analogous model setup follows when the buyer's demand rate (which we standardized to one in our setup) is private information. In this case, the high demand buyer would try to mimic the low demand buyer in order to claim a higher compensation for larger lot sizes.[3]

3.1 Full information and simple coordinating contracts

Under full information, the supplier has knowledge of the buyer's realization of holding cost (h). Hence, the supplier can offer the buyer an optimal price per unit, w, to promote a higher order size, q. This leads to the contract $\langle w, q \rangle$ as an outcome of the following optimization problem:

$$\text{Problem FI: max } \pi_s = w - \frac{f}{q}$$
$$s.t.\ \pi_b = Y_b^{'} - w - \frac{h}{2} \cdot q \geq Y_b^{'} - R.$$

Since demand is standardized to one unit per period, π_s and π_b, denote the supplier's and the buyer's unit profit margins, respectively. The supplier's objective function maximizes his unit profit. In the optimal solution to problem FI, the buyer's participation constraint is binding. Solving the participation constraint for w, we insert w in the objective function and optimize optimal order quantity, and get $q^{FI} = \sqrt{\frac{2f}{h}}$. The optimal order size, thus, resembles the classical economic order quantity with demand per period standardized to one unit. The optimal unit price, w^{FI}, is set to satisfy the buyer's participation constraint, ensuring that the buyer's profits are not smaller than in her outside option, i.e., $w^{FI} = R - \frac{h}{2} \cdot q^{FI}$. The optimal contract parameters under full information $\langle w^{FI}, q^{FI} \rangle$ not only optimize the supplier's profit, but also the overall supply chain performance.

[2] The fixed revenues $Y_b{'}$ are from a modelling perspective irrelevant since the supply chain performance is only determined by the average fixed cost and holding cost per period. We introduced the fixed revenues to move the subject's payoffs in the following experiments from the loss to the profit domain.

[3] Another interesting situation that fits the model is a quality game. The supplier has increasing marginal cost in the quality level of the product. The buyer's production cost decreases linearly in the input quality. In such a context, the failure rate and therefore the total cost of production for any given quality level is private information. The theoretical analysis of the model is equivalent the one provided in this paper. The interpretation of the results, however, relates to the conflict concerning the quality of the delivered inputs.

3.2 Asymmetric information and screening contracts

Under asymmetric information, the holding cost realization is only known to the buyer, but not to the supplier. The supplier only has the information on the probability distribution $p_i, i = 1, \ldots, n$ over possible values of the buyer's holding cost $h_i, i = 1, \ldots, n;\ h_1 > \cdots > h_n$.

The basic screening idea is that the profit maximizing buyer reveals her private information with her contract choice. Let $\pi_{b,i}(q_j) = Y_b' - w_j - h_i/2 \cdot q_j$ denote the unit profit margin of the buyer facing holding cost h_i, and choosing the contract $\langle w_j, q_j \rangle$. Information revelation is ensured by the incentive constraint $\pi_{b,i}(q_i) \geq \pi_{b,i}(q_j),\ \forall i \neq j;\ i,j = 1, \ldots, n$. The buyer facing holding costs h_i will always choose the offer $\langle w_i, q_i \rangle$ as any other contract $\langle w_j, q_j \rangle$ will result in a lower unit profit margin. The participation constraint $\pi_{b,i}(q_i) \geq Y_b' - R,\ \forall i = 1, \ldots, n$ ensures that the buyer will not benefit from choosing the alternative supplier.

Let $\pi_{s,j} = w_j - f/q_j$ denote the supplier's unit profit margin if the buyer chooses the contract q_j. Due to the incentive constraint the supplier knows that the buyer will choose the contract $\langle w_i, q_i \rangle$ with probability p_i. Hence, the supplier can maximize his expected unit profit margin with the following optimization problem[4]:

$$\text{Problem AI: } \max E[\pi_s] = \sum_{i=1}^{n} p_i \cdot \pi_{s,i}$$

$$s.t. \quad \pi_{b,i}(q_i) \geq \pi_{b,i}(q_j), \quad \forall i \neq j;\ i,j = 1, \ldots, n$$

$$\pi_{b,i}(q_i) \geq Y_b' - R, \quad \forall i = 1, \ldots, n$$

The following notation is used to refer to the supplier's optimal menu of contracts (screening contract) $A = (A_i | i = 1, \ldots, n)$ where $A_i = \langle w_i^{AI}, q_i^{AI} \rangle,\ \forall i = 1, \ldots, n$. Furthermore, $F_i = \langle w_i^{FI}, q_i^{FI} \rangle,\ \forall i = 1, \ldots, n$ denotes the supply chain's optimal contract when the buyer faces holding costs h_i. We refer to Voigt (2011) for a derivation of the optimal menu of contracts.

One important feature of the screening contracts is that order sizes are increasing with decreasing holding cost levels, i.e., $q_i^{AI} \geq q_{i+1}^{AI}$ and the unit prices are decreasing with decreasing holding cost levels, i.e. $w_i \leq w_{i+1},\ \forall i = 1, \ldots, n$. The menu of contracts can, therefore, be interpreted as a quantity discount that is inefficient, since all order sizes except q_1^{AI} are downward distorted. If the supplier would have full information, he could offer q_i^{FI} instead of letting the buyer self-select the distorted order size q_i^{AI} and, thus, enhance the supply chain performance. A numerical example follows in the next section on the basis of the parameters we use in our experiment.

We derive the theoretical benchmark along with our research hypotheses in Sect. 4.2. Note, that our model only captures those elements that are relevant from a game-theoretic perspective. For our experiment, however, we add a few game-

[4] An additional slack as mentioned in the introduction renders for example the incentive constraint to $\pi_{b,i}(q_i) - \delta_i \geq \pi_{b,i}(q_j)$, where δ_i is the cost of deviating from self-selection (Laffont and Martimort 2002, Chapter 9.8.1 and Voigt 2015).

theoretically irrelevant features (punishments and rewards, additional incentives for self-selection) to identify behavioral effects. In the discussion of the hypotheses, we explain why those features have no impact from a game-theoretic perspective (Hypotheses 1a–d, game-theoretic benchmark), but may have a behavioral impact (Hypotheses 2a–c, punishment treatment, and Hypothesis 3, enforced self-selection treatment).

Finally, note that the supplier might want to limit himself by offering only one contract even under asymmetric information. If he wants to make offers that leaves all cost types at least the reservation profit, then the contract F_H would be optimal. This can be easily confirmed by solving problem AI with $\langle w_i, q_i \rangle = \langle w, q \rangle,\ \forall i = 1, \ldots, n$. In turn, the assumption that all buyers receive an acceptable offer may be lifted (frequently referred to as a “cut-off” policy). Under these assumptions, the contract $\mathrm{F_M}$ turns out to be the supplier’s optimal simple contract offer with expected profits of 47.22 followed by F_L with 46.76 and F_H with 36.9.

4 Experimental design, implementation, and research hypotheses

The experimental software was implemented with the toolbox z-Tree (Fischbacher 2007). Participants were recruited online using ORSEE (Greiner 2015) and randomly distributed over the treatments. The subject pool contains graduate and undergraduate students of a mid-scale university in Germany.

Upon arrival, each participant received written instructions that were read out aloud (see online-supplement). The instructions for both roles (supplier/buyer) were identical. All remaining questions were answered privately in the subject’s cubical at the experimental laboratory. Subjects were paid according to their performance in the experiment.

Subject’s payments were the total cumulative points earning during the experiment multiplied with a conversion rate of 0.01, i.e. every experimental monetary unit exchanges to 1 cent. The average earnings were 9.7 EUR (Max: 18.5 EUR/Min: 7 EUR). Sessions lasted no longer than 90 min. The experiment consisted of three treatments with a total of 72 subjects. Subject were recruited using a standard recruiting software and were randomly assigned to treatments. From a subject pool of almost 2000, most subjects were undergraduates in a mixture of all study programs. About 47 % of the subjects were female. Each subject participated in only one treatment (between subjects design). All subjects played 30 rounds.

We consider a serial supply chain consisting of one buyer and one supplier. The subjects were randomly assigned the role of buyer or supplier after the instructions were read. The matching of buyers and suppliers remains unchanged over time.

In the following, we present the experimental design for our three treatments (see Table 1). We compare the baseline treatment to two behavioral variations with reduced strategic risk. 24 subjects participated in each treatment (12 buyers and 12 suppliers), leaving twelve independent observations per treatment.

Table 1 Treatment summary

	No. of participants (no. of independent observations)	Change to baseline
Baseline treatment	24 (12)	–
Punishment treatment	24 (12)	Supplier may costly reduce buyer's payoffs
Enforced self-selection treatment	24 (12)	Buyer can only choose profit maximizing contract or reject
Total	72 (36)	

4.1 Treatment design

4.1.1 Baseline treatment

4.1.1.1 Parameters There are three holding cost realizations $h_L = 1$, $h_M = 5$ and $h_H = 9$ that occur with the corresponding a priori probabilities $p_L = 0.4$, $p_M = 0.3$ and $p_H = 0.3$. The holding costs are drawn independently in every round according to the distribution function which is common knowledge. Thus, frequencies of holding cost realizations might differ across treatments. The supplier's total fixed cost is $f = 800$ and the buyer's unit cost of sourcing from the alternative supplier is $R = 157$. The buyer's revenue is fixed at $Y_b' = 160$ per round. The supplier's revenue is fixed at $Y_s = 155$ per round.[5]

4.1.1.2 Decision sequence *(t = 1)* At the beginning of each round, the buyer sends a signal S that may communicate her holding cost to the supplier, where $S \in [S_L = h_L, S_M = h_M, S_H = h_H, S_{No} \doteq No\ Signal]$.[6]

(t = 2) After the buyer sends her signal, the supplier offers a contract. Contract offers are restricted to (i) $F_L = \langle w_L^{FI}, q_L^{FI} \rangle$, (ii) $F_M = \langle w_M^{FI}, q_M^{FI} \rangle$, (iii) $F_H = \langle w_H^{FI}, q_H^{FI} \rangle$ and (iv) $A = (A_L, A_M, A_H)$ where $A_i = \langle w_i^{AI}, q_i^{AI} \rangle$, $i \in L, M, H$. Thus, the supplier may either offer one of the three tailored contracts F_i (i.e. one of the contracts that are optimal under full information, see Problem FI) or the screening contract A that is optimal (but second best) under asymmetric information (see Problem AI).

(t = 3) If the supplier offers a tailored contract in t = 2, then the buyer can either accept or reject this offer. If the supplier offers the screening contract A in t = 2, the buyer chooses A_L, A_M, A_H, or rejection. If the buyer rejects the offer, she sources

[5] Note that in our experimental instructions (see online-supplement) we do not refer directly to the unit prices. Instead, we present the unit price separated into a fixed revenue part Y_s and a compensation part Z_i that is derived optimally from the game parameters. In our model context, this directly translates to $w_i = Y_s - Z_i$.

[6] Since our signals are costless, they are cheap-talk and cannot be used from a game-theoretic perspective to credibly share information (see Crawford 1998). Hence, our game is not a signaling game.

from the alternative supplier at the cost R, the supplier realizes zero profits, and the next round starts.

(t = 4) The supplier has the option to give the buyer a reward. This is, because the truth-telling buyer will only receive her reservation profit under a tailored F_i—contract. The reward allows the supplier to compensate the buyer for the profit difference between the screening contract and the reservation profit. Because of the efficiency gap between screening contracts and tailored contracts, the reward can be set such that both parties are better off. The buyer, thus, faces a situation similar to the well-known trust game (see, e.g., Berg et al. 1995): the buyer is vulnerable if she sends a truthful signal and has to trust that she receives a sufficient reward from the supplier that compensates for accepting the efficiency enhancing F_i—contract.

The maximum size of the reward is limited to the supplier's earnings in the respective round. This limit on rewards still gives sufficient leeway for cooperative play resulting in win–win outcomes. In particular, the highest reward that is necessary to implement win–win cooperation occurs when the buyer signals low holding cost and accepts the efficient contract accordingly. In this case, the minimum reward that leaves the buyer indifferent between reporting truthfully and accepting and the respective profit in the screening contract is 43.77.

Finally, note that the true holding cost realization is never revealed to the supplier. The supplier can only check consistency (contract choice in the screening contract matches the signaled holding cost level) or the credibility of the signal by statistical interference.

(t = 5) A new holding cost parameter is drawn in every round. Thus, the supplier can neither infer the buyer's holding cost parameter of the next round through the buyer's signal nor through the buyer's action in the current period.

Table 2 depicts the parameter values and the resulting payoffs for our experiment. For example, if the buyer has low holding costs of h_L and accepts the contract F_M, she realizes a profit of $\pi_b = 38.88$ (net of the additional reward in t = 4) and the supplier realizes $\pi_s = 67.46$ (net of the additional reward in t = 4).

Note that the participation constraint makes the buyer of type i (with holding cost h_i) indifferent between her outside option and the contract F_i. To avoid indifference in the experiment, we add 0.1 to the buyer's profit when accepting the contract F_i. Similarly, we break the tie between accepting the self-selection option of the screening contract for the true holding cost versus the next higher holding cost, by adding 0.1 to the true option.[7] For example, if a buyer with h_M chooses the contract A_H, she earns 0.1 less than she earns by choosing the self-selection option A_M. In this case, we refer to A_H as the *indifference contract* and to A_M as the *self-selection contract*.

The following extensive game form in Fig. 1 illustrates the effects of trust and mistrust in our experimental design. The buyer's decision node is denoted with 'B' and the supplier's with 'S', respectively. Additionally, 'N' denotes the so-called nature's decision node, i.e. the nature decides with respect to the a priori probabilities which holding cost realization is assigned to the buyer. The buyer

[7] The small incentive we add does not alter the fact that screening contracts yield the highest expected payoffs, when there is only incomplete information on the holding cost for the supplier.

Table 2 Contracts and corresponding profits

	Order size: q_i	Unit price: w_i	Profit supplier: π_s	Profit buyer: π_b		
				h_L	h_M	h_H
F_L	40.00	136.90	116.90	3.10	−76.90	−156.90
F_M	17.89	112.18	67.46	38.88	3.10	−32.68
F_H	13.33	96.90	36.90	56.43	29.77	3.10
Screening contract: ***A***						
A_L	40.00	93.13	73.13	46.87	−33.13	−113.13
A_M	12.44	107.01	42.72	46.77	21.88	−3.00
A_H	9.34	114.86	29.23	40.47	21.78	3.10

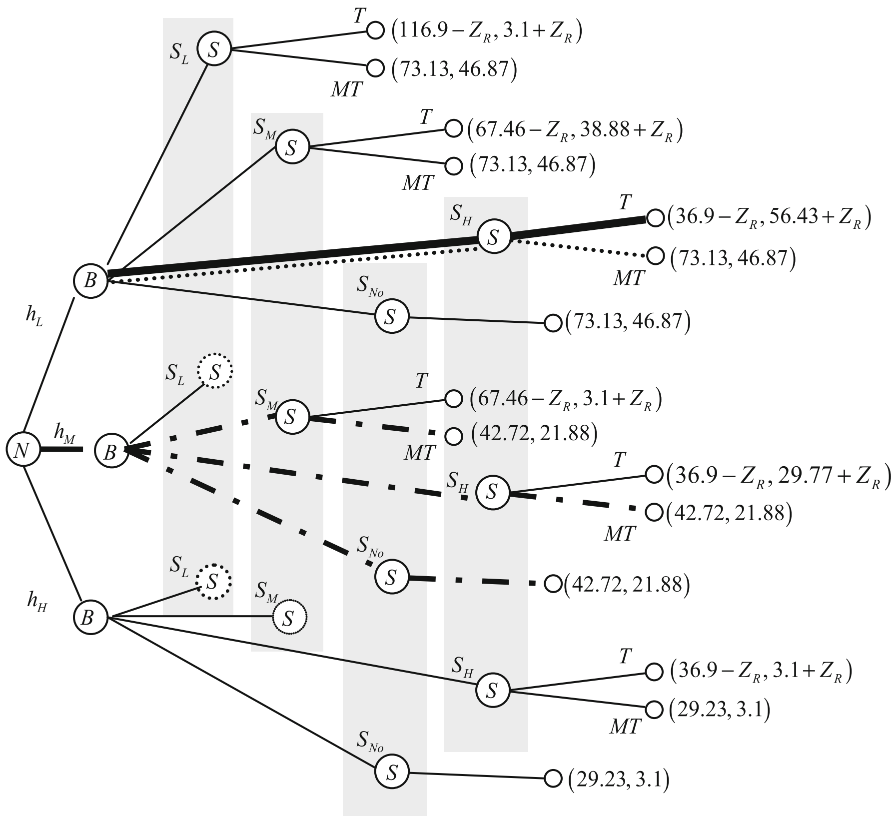

Fig. 1 Payoff consequences of trust and mistrust (adapted from Voigt 2011, p. 29)

gives a signal $S \in (S_L, S_M, S_H, S_{No})$. The supplier cannot distinguish whether a report is truthful or not. This fact is captured by the shaded boxes. Each of these boxes contains the same signal, e.g. S_H, but this signal is not necessarily identical to the holding cost realization, e.g. h_H.

The supplier's decision to trust (i.e. tailor a F_i contract to the signal) is denoted with 'T' and to mistrust (i.e. offer a screening contract) with 'MT' respectively. For the sake of clarity, Fig. 1 only shows the payoff consequences for the cases in which the buyer reports truthfully, overstates his holding costs or refuse to give a signal. Understatements of holding cost realizations are indicated by dotted circles. All payoffs that result for a specific action sequence do directly follow from Table 2 and Z_R denotes the amount of reward from the supplier to the buyer. For the sake of clarity, we assume that no reward is paid if the supplier mistrusts (i.e. $Z_R = 0$). Furthermore, the game-tree does not capture the buyer's choice to choose the alternative supplier.

Obviously, cooperative behavior (i.e. report truthfully and trust) can lead to Pareto improvements compared to mistrust, as long as the additional side-payment is sufficiently high. For example, if the buyer faces cost h_M, reports truthfully S_M, and the supplier trusts, then the total supply chain profit is $67.46 + 3.1 = 70.56$.[8] Yet, if the supplier mistrusts the signal and offers the menu of contracts, the profits under self-selection are $42.72 + 21.88 = 64.6$. The efficiency loss is therefore $70.56 - 64.6 = 5.96$. The side-payment resulting in a win–win outcome can be calculated from $67.46 - Z_R \geq 42.72$ and $3.1 + Z_R \geq 21.88$, i.e. $18.78 \leq Z_R \leq 24.74$.

In turn, consider a buyer who faces holding costs h_L and reports S_H (see bolded line in Fig. 1). If the supplier trusts and offers F_H than the supplier yields profits of $\pi_s = 36.9 - Z_R$ and the buyer $\pi_{b,l} = 56.43 + Z_R$, respectively. The supply chain profits result from $36.9 + 56.43 = 93.33$. Yet, if the supplier mistrusts (i.e. offering the menu of contracts) the supply chain profits would amount to $73.13 + 46.87 = 120$ (see dotted line in Fig. 1). Hence, trust in comparison to mistrust leads to an efficiency loss of $120 - 93.33 = 26.67$.

Finally, note that communication can only be an effective coordination instrument as long as the supplier does not simply ignore all signals. In this case, the menu of contracts would be offered regardless of the buyer's signal. The interrupted lines in Fig. 1 highlights this for the holding cost realization h_M. Obviously, the outcome is independent of the buyer's signal and communication is not effective.

4.1.2 *Punishment treatment*

In the punishment treatment, the supplier has the option to punish the buyer or to reward the supplier as in the baseline treatment. The punishment option provides the buyer a means to retaliate uncooperative behavior. Yet, since the true holding cost level is never revealed to the buyer, the supplier's intent to play cooperative or uncooperative can only be inferred from contract choice behavior (e.g., inconsistencies between the signaled cost and the selected contract option indicate uncooperative behavior) or the signaling strategy (e.g., too frequent high cost signals indicate uncooperative behavior).

[8] Note that the additional side-payment does not influence the overall supply chain performance, as it is simply a transfer payment between the supplier and the buyer.

The supplier may punish the buyer by arbitrarily reducing her profits in t = 4. Punishment, however, is costly. Every unit of punishment costs the supplier 0.2. The maximum punishment per round is limited to 40. We conducted this treatment since punishments are a realistic threat in supply chain relationships (see Cui et al. 2007 for a discussion on punishments in supply chain relationships). As Kahneman et al. (1986) put it: "Even profit-maximizing firms […] are willing to […] punish unfair firms at some cost […]" (see Kahneman et al. 1986, p. 285). In practice, such a punishment may be operationalized by terms of contract that penalize deviations from business-as-usual.

4.1.3 Enforced self-selection treatment

The only difference between the baseline treatment and this treatment is that the buyers in this treatment are forced to choose the profit-maximizing self-selection contract in t = 3 whenever offered a menu of contracts. In particular, indifference contract choices are not possible which means that the self-selection mechanism of the screening contract works perfectly, substantially reducing suppliers' strategic risk.

In practice, self-selection can be enhanced by adding slack into the incentive constraint (see footnote 4). Although our treatment cannot be used to assess the optimal threshold for the buyers' self-selection incentives, it provides a clear benchmark on the effects of sufficiently high slack in incentive constraints.

4.2 Research hypotheses

This section provides the game theoretic benchmark as well as the behavioral research hypotheses for our three treatments. Note, that the game-theoretic benchmark is identical in all three treatments. This is, because we only test behavioral features that are irrelevant on the equilibrium path. We present the game-theoretic benchmark in Hypotheses 1a–1e, the behavioral research hypotheses for the punishment treatment in Hypotheses 2a–2c, and the behavioral research hypothesis for the enforced self-selection treatment in Hypothesis 3.

Our game-theoretic benchmark is derived from a one-shot game. This approach is in line with a number of other experimental papers (see Boles et al. 2000; Croson et al. 2003) that study repeated interaction games in which the one-shot game equilibrium is also the sequentially rational equilibrium of the finitely repeated game.[9]

[9] There is ample experimental evidence on cooperation (deviation from sequentially rational play) in social dilemmas in finitely repeated games. Typically, cooperation rates are higher in partner designs than in stranger designs (see Andreoni and Miller 1993). One frequently employed explanation dating back to Kreps et al. (1982) is that the probability that some player deviate from rational, self-interested, and profit maximizing choices may foster cooperation. While our finitely repeated design may lead to higher levels of cooperation than in a one-shot game, we believe that it resembles the multi-period interaction of most real supply chains.

Hypothesis 1a: Communication is uninformative Since the preferences in the underlying game are perfectly opposed (i.e., the buyer always tries to convince the supplier that she has high holding cost), and since private information is independently distributed between periods, credible information sharing cannot take place in the non-cooperative equilibrium (see Crawford 1998; Fudenberg et al. 1990).

Hypothesis 1b: Suppliers ignore signals As the fully rational supplier anticipates that all signals are uninformative (see Hypothesis 1a), we will not observe any correlation between the buyer's signals and the supplier's behavior.

Hypothesis 1c: Suppliers only offer screening contracts Without informative signals (see Hypothesis 1a), the screening contract maximizes the supplier's expected profits. Hence, we will only observe screening contracts.

Hypothesis 1d: Buyers choose the self-selection option of the screening contract The optimal choice of profit maximizing buyers, who are offered a screening contract (see Hypothesis 1c), is the self-selection option.

Hypothesis 1e: Punishments and rewards are not observed Since suppliers offer screening contracts and buyers self-select into their profit maximizing option (Hypotheses 1a–1d), we only observe equilibrium play, in which—by definition of a game-theoretic equilibrium—payoffs cannot be increased using the reward and punishment options (see, e.g., Fudenberg and Tirole 1992).

There is ample evidence in experimental economics that punishments options have a positive effect on cooperation in social dilemmas such as ultimatum bargaining, trust games or voluntary contribution to public goods. Instead of reviewing this extensive literature here, we refer interested readers to a recent meta-analysis on the effects of punishment and rewards on cooperation in social dilemmas by Balliet et al. (2011). They report that a punishment option is particularly effective when punishment is costly and administered in repeated interaction. Since both conditions hold in our punishment treatment, we expect a cooperation enhancing effect of the punishment option. Cooperation in our setup is characterized by supply chain profit maximizing actions (see Hypotheses 2a, 2b, and 2c, below) resulting from contracts that are tailored to truthfully shared information.

Hypothesis 2a: The frequency of truthful signals is higher in the punishment than in the baseline treatment Our experimental design entails a cheap talk game, in which a sender (the buyer) gives a signal based on her private information about her type (i.e. the holding cost), with the receiver (the supplier) reacting to this signal. We refer to Crawford (1998), for a comprehensive survey on cheap talk theory. Brandts and Charness (2003) report an experimental laboratory study in which the receiver observes ex-post whether the signal was deceptive or honest. They report that receivers show a strong and significant tendency to punish deceptive senders. Sánchez-Pagés and Vorsatz (2007) replicate the result that and show that senders are significantly less deceptive with punishment than without and tend to give no signals (similar to our notion of S_{No}) more frequently. Peeters et al. (2013) report significantly higher levels of truthfulness and less deception when a punishment option is available.

In contrast to above mentioned experimental studies, our experimental design does not allow to verify ex-post if a signal was honest or deceptive when a simple contract is chosen. However, supplier can observe signal consistency with screening contracts and can discover statistically excessive high cost reporting. Hence, we hypothesize that buyers send truthful signals more often in the punishment treatment compared to the baseline, because they fear to be punished more frequently for apparently uncooperative behavior when suppliers have a punishment option. We define buyers' behavior as apparently uncooperative if the high cost signal is send significantly more often as statistically expected or if there is an inconsistency between the signaled cost and the selected contract option.

Hypothesis 2b: The supplier offers the contract tailored to the signal more frequently (i.e., he shows a higher level of trust) in the punishment treatment than in the baseline treatment Sánchez-Pagés and Vorsatz (2007) and Peeters et al. (2013) report a significantly higher level of receiver's trust in the sender's signal when a punishment option is available. We therefore expect the supplier to show a higher level of trust in the punishment treatments, in which the threat of punishment enhances truthful and consistent signals. A higher level of trust is characterized by a higher frequency of simple contracts that are tailored to the respective signal.

Hypothesis 2c: The supply chain performance is higher in the punishment treatment than in the baseline treatment We have two reasons to believe that the punishment option may enhance supply chain performance. First, we expect the punishment option to increase truthful signaling (2a) and trust (2b). Thus, supply chain actions are better coordinated, because simple contracts fit to the actual holding cost level. Second, we expect the punishment option to decrease the frequency of inconsistent indifference contract choices.

Hypothesis 3: Screening contracts are offered more frequently in the enforced self-selection treatment than in the baseline treatment Consider a buyer giving a deceptive signal, e.g., she signals high holding costs although she only incurs medium holding costs. In case the supplier offers the menu of contracts, the buyer can easily pretend to be honest by choosing the high cost contract that fits to her signal (but not to her holding cost realization). If the supplier anticipates that the buyer may choose an indifference contract in order to cover up her deceptive signal, then offering a simple tailored contract might be profit maximizing. In these cases, offering a simple contract does not result from trust in the signal, but from mistrust towards the buyers' self-selection behavior. Thus, eliminating the strategic risk of not choosing the self-selection contract allows us to disentangle trust from strategic risk avoidance in this treatment.

5 Experimental results

We present the results of the experiments in the sequence of the decisions taken, i.e., the signal, the contract offers, the contract choices, and the cooperation facilitating role of punishments and rewards. We conclude the analysis with analyzing the impact of observed behavior on supply chain performance.

5.1 Buyers' signaling behavior

Table 3 summarizes the frequencies of signals across treatments. If all buyers' reported truthfully, we would expect a frequency of signals according the a priori distribution, i.e., 40, 30 and 30 % for the low cost, medium cost and high cost level respectively. We observe a clear shift to the strategically most relevant high cost signals in all treatments, however, the shift is more pronounced in the baseline treatment than in the two control treatment. Omitting all S_{No} observations, we find a highly significant correlation between holding cost levels and signals (Spearman, $p < 0.01$, $\rho = 0.4$). Thus, buyers give significantly higher cost signals if holding cost increase and vice versa. We therefore reject hypothesis 1a ("Communication is uninformative").

The following Fig. 2 shows the development of truthfulness over time. The horizontal line indicates the average level of truthfulness across all treatments, i.e., 53 %. A visual inspection of the graphs tends to support our hypothesis that truthfulness is higher in punishment treatment than in the baseline treatment.

In order to further assess if the shift in signals is due to a higher level of truthful signals, we run the following random effects probit regression estimating the buyer's tendency to signal truthfully or not:

$\text{Probit}(Y_{it} = 1) = \beta_0 + \beta_1 \times \text{Punish} + \beta_2 \times \text{Enforced} + u_i + \varepsilon_{it}$ where $Y = 1$ if the buyer reports truthfully (i.e., $S_i = h_i$) and $Y = 0$ otherwise. The variables "Enforced" and "Punish" are the treatment dummies. We run separate regressions for each holding cost level. The marginal effects and respective statistics of the probit regression are summarized in Table 4.

Overall, the regression results support directionally but not significantly hypothesis 2a, i.e. we observe a higher propensity to report truthfully in the punishment treatment. Taking all holding cost realization into account (column "All") we see a higher chance to observe truthful behavior in the punishment treatment. A closer look on the separate regressions (columns "h_{low}, h_{med}, h_{high}" denote the respective selection of included holding cost realizations) shows that this effect is mainly driven by the subjects' tendency to report more truthfully in the punishment treatment when facing the holding realization h_M. Taking the average truthfulness given h_M per subject as an independent observation, the latter observation is weakly significant (Mann–Whitney U (MWU), $p < 0.1$, one sided). Thus, buyers with holding cost h_M tend to report more truthfully instead of giving the deceptive high cost signal in the punishment treatment.

Table 3 Frequencies of signals in treatments

	S_L (%)	S_M (%)	S_H (%)	S_{No} (%)
Baseline	13.06	13.89	62.50	10.56
Punishment	19.72	26.06	43.33	13.89
Enforced	9.17	18.89	54.44	17.50

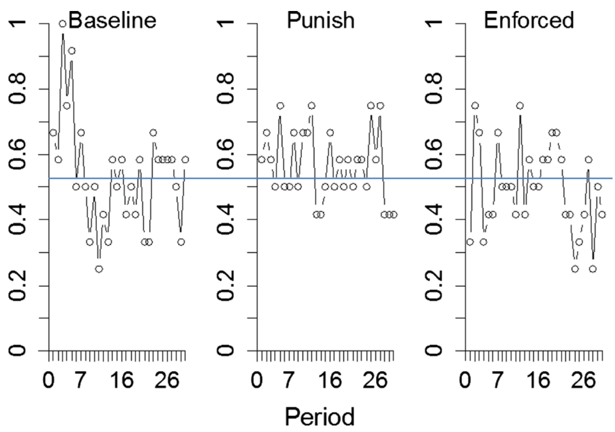

Fig. 2 Average truthfulness by period and treatment

Table 4 Probit regression of treatment effects on truthfulness

	h_L	h_M	h_H	All
Enforced	−0.17 (0.67)	0.55 (0.52)	0.17 (0.75)	−0.03 (0.35)
Punish	−0.41 (0.66)	0.69 (0.52)	−0.08 (0.73)	0.14 (0.35)
σ_i	1.47 (0.28)	1.11 (0.21)	1.16 (0.43)	0.81 (0.12)
LL	−164.35	−159.20	−43.15	−530.01
Prob > Chi^2	0.67	0.37	0.94	0.87
N	347	287	295	929

Marginal effects of probit regression, SD in parentheses (*** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$)

5.2 Suppliers contract offers

The previous section shows that buyers share their cost information to some extent in all treatments. Yet, even if all buyers truthfully report their holding cost, there would be no effect of communication if the signals are ignored by the suppliers. Table 5 summarizes the frequency of contract offers by types. The high degree of variance in the suppliers' contract offer behavior shows that we have no empirical support for hypothesis 1c (only screening contracts are offered). A one side binomial test with the alternative hypothesis that the frequency of screening contract is smaller than one is highly significant ($p < 0.01$) for all subjects in all treatments.

Simply comparing the ratio of F_i—contracts to screening-contracts across treatments is not a good measure for the suppliers' trust in the signal, since a F_i—contract offer is not necessarily related to the respective signal. Furthermore, note that the frequency of signal consistent F_i—contracts does not necessarily measure the suppliers' trust in the buyers' signals either. For suppliers to offer signal tailored F_i—contracts, it suffices that the suppliers trust in the buyers' consistent signal-choice behavior, even if the signals are not truthful. Hence, we must cautiously distinguish between trust in the truthfulness of the signal, trust in signal consistent

Table 5 Frequency of F_i- and screening-contract offers

	F_L (%)	F_M (%)	F_H (%)	Menu (%)
Baseline	10.56	16.39	28.33	44.72
Punishment	4.44	16.67	34.44	44.44
Enforced	5.00	8.33	18.33	68.33

contract choices, and the lack of strategic risk resulting from indifference contract choices. The first concept plays no role if suppliers offer screening contracts. It is, however, essential for coordinating the supply chain based on signal tailored contracts. The second concept is essential for suppliers' payoffs, because suppliers' payoffs only depend on the signal-choice consistency, but not on the truthfulness of the signal. The third concept mainly plays a role when suppliers offer screening contracts, because with a screening contract suppliers' payoffs can be substantially affected by buyers' out-of-equilibrium play at minor costs for the buyers.[10] Comparing the two contract types, we can conjecture that an increase in (either form of) trust or an increase in strategic risk (in the screening contract) will lead to an increased use of tailored contracts (see also hypothesis 3).

The following Fig. 3 shows the development signal tailored contracts over time. The horizontal line indicates the average level of truthfulness across all treatments (33 %). A visual inspection of the graphs clearly highlights fewer signal tailored contracts in the enforced self-selection contract. In turn, hypothesis 2b stating that we observe a higher level of signal tailored contracts in the punishment treatment is not supported.

In order to further assess the supplier's contract offer behavior, we run the following random effects probit regression that estimates the supplier's tendency to tailor the contract to the signal or not. We also run a separate regression for S_L, S_M, and S_H. The regression results and respective statistics are summarized in Table 6.

Overall, we do not find support for our hypothesis that the punishment option (hypothesis 2b) increases trust in signals. However, the above regression indicates that tailored contracts are less likely to be observed in the enforced self-selection than in the baseline treatment. A closer look on the frequency of contract offers (see Table 7) indicates that the lower level of tailored contracts is compensated by a higher level of screening contract offers. A visual inspection of the frequency of screening contract offers by period and treatment in Fig. 4 supports the hypothesis of more screening contract offers in the enforced self-selection treatment. The horizontal line depicts the average level of screening contract offers across all treatments (52.5 %).

We run a random effects probit regression to test for differences in screening contracts offers across treatments (i.e., $Y = 1$ if the supplier offered the screening contract and $Y = 0$ otherwise). The regression results are significantly in favor of hypothesis 3, i.e., the suppliers in the enforced self-selection treatment show a

[10] Additionally, suppliers face the strategic risk that an offer is rejected. However, this kind of strategic risk cannot explain treatment differences, because this option is available in all treatments.

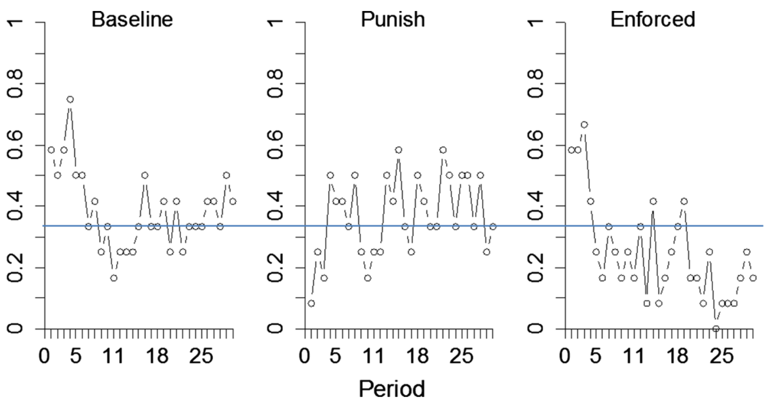

Fig. 3 Frequency of tailored F_i contract offers

Table 6 Probit regression of treatment effects on tailored contracts

	S_L	S_M	S_H	All
Enforced	0.25 (1.44)	−0.52 (0.63)	−0.47 (0.33)	−0.37 (0.32)
Punish	−1.08 (1.36)	0.02 (0.65)	0.49 (0.33)	0.03 (0.32)
σ_i	2.57	1.25 (0.28)	0.72	0.74 (0.11)
LL	−46.54	−66.03	−329.57	−530.97
Prob > Chi^2	0.57	0.07	0.02	0.39
N	151	201	577	929

Marginal effects of probit regression, SD in parentheses (*** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$)

Table 7 Probit regression of treatment effects on screening contract offers

Enforced	0.71 (0.28)***
Punish	−0.01 (0.28)
σ_i	0.65 (0.09)
LL	−646.52
Prob > Chi^2	0.01
N	1080

Marginal effects of probit regression, SD in parentheses (*** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$)

significantly higher propensity to offer screening contracts in the enforced self-selection treatment than in the baseline treatment.

We finally test over all treatments for a dependency of signals to contract offers with random effect probit regressions by contract offers (see Table 8). The results show that contract offers significantly depend on the respective signals, i.e., the likelihoods of observing an F_i—contract are significantly higher if the buyer sends the respective signal S_i. The likelihood of observing screening contract, in turn, is significantly higher when observing the signal S_{No} (i.e. the marginal effects for S_M

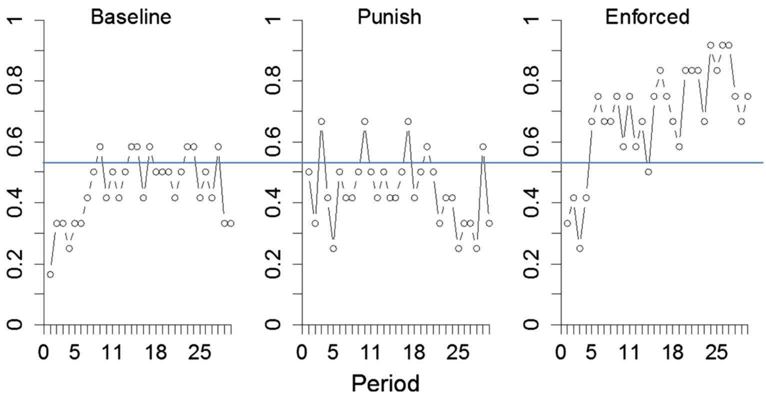

Fig. 4 Frequency of screening contract offers by period and treatment

and S_H are significantly negative and for S_L not significantly different from zero). Thus, Hypothesis 1b ("Suppliers ignore signals") is rejected.

5.3 Buyers contract choice behavior

Table 9 gives an overview of buyers' contract choices. We divide the buyers' contract choices into five categories: (1) the buyer chooses the profit maximizing contract, (2) the buyer chooses the indifference-contract, i.e. she chooses A_M (A_H) although she faces holding costs of $h_L(h_M)$, (3) the buyer chooses the alternative supplier when (almost) indifferent between the offered contract and the outside option, i.e., whenever she faces holding costs h_H or is offered the tailored contract that fits her actual holding cost $F_i = h_i, i \in (L, M)$, (4) the buyer chooses the alternative supplier because the participation constraint is not satisfied, which is only possible if the supplier offers F_L or F_M, and (5) the buyer has a loss greater than 0.1. The values in Table 9 indicate the percentages of the cases in each treatment.

First note that only 3 % of all choices fall in category 5 and therefore 97 % of all observations are optimal or nearly optimal contract choices. However, we find overall a substantial amount of only nearly profit maximizing behavior (i.e., categories (2) and (3)). Overall, the results only partly confirm hypothesis 1d (i.e. profit maximization of buyers), leaving a substantial number of non-profit maximizing choices ($\approx$23 %; cases 2–5). In most of these cases, however, the buyer incurs only a marginal loss ($\approx$20 %, cases 2–4). A one sided binomial test with the null hypothesis that only self-selection contracts are chosen once the screening was offered is rejected for 26 of 36 subjects ($p < 0.01$, one-sided). On an aggregate level (i.e. over all subjects), the null hypothesis is rejected. Thus, while we observe that some buyers always self-select (10 of 36), we have to reject hypothesis 1d ("Buyers self-select").

In order to test for treatment effects, we compared the average contract choice frequencies of the buyers in the respective treatments with a Mann–Whitney U test.

Table 8 Probit regression of signaled holding cost on contract offers

	F_L	F_M	F_H	Menu
S_L	1.07 (0.29)***	−0.13 (0.24)	−0.63 (0.35)*	−0.28 (0.18)
S_M	−1.74 (0.48)***	1.12 (0.20)***	0.49 (0.24)**	−0.75 (0.17)***
S_H	−1.12 (0.26)***	−0.22 (0.19)	1.88 (0.22)***	−0.92 (0.15)***
σ_i	1.30 (0.27)	0.64 (0.11)	0.79 (0.12)	0.73 (0.10)
LL	−173.49	−355.85	−392.61	−623.52
Prob > Chi^2	0.00	0.00	0.00	0.00
N	1080	1080	1080	1080

Marginal effects of probit regression, SD in parentheses (*** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$)

Table 9 Buyers' contract choices

Treatment	Profit maximum (1) (%)	Indifference contract (2) (%)	Reject-indifference contract (3) (%)	Reject-participation constraint (4) (%)	Profit loss >0.1 (5) (%)
Baseline	70.56	7.50	10.28	7.78	3.89
Punishment	82.50	8.33	7.22	1.11	0.83
Enforced	76.67	0.00 (by design)	14.44	5.28	3.61

We did not find any significant differences in the contract choice behavior except the following. There are significantly less rejections in case of violated participations constraints (category 4) in the enforced self-selection treatment than in the baseline treatment (MWU, $\alpha < 0.1$, two-sided). This is due to the fact that violations of the participation constraints are only possible in case of (mis) tailored contracts. Because tailored contracts are offered significantly less often in the enforced-self-selection treatment (see Table 5), we observe such instances less often.

We further analyze the buyer's contract choice behavior by analyzing the interaction between consistency and profit maximizing behavior by contract offer, i.e., F_i contracts (Table 10) and menu of contracts (Table 11). We skip S_{No} observations (151 observations) from this analysis. Clearly, inconsistencies are rather low under F_i contracts and only observed along with profit maximizing contract choices. In these cases, the buyer rejects F_i offers although their reported holding cost level suggests otherwise. In turn, inconsistencies are frequently observed under menus of contracts and highest in the enforced self-selection treatment, highlighting that inconsistencies are more likely to occur when deceptive signals cannot be covered up.

Figure 5 shows signal consistent contract choices by treatment and period. The horizontal line depicts the average level of signal consistent contract choice, i.e., 79 %. A visual inspection confirms the overall lower level of inconsistency in the enforced self-selection treatment and shows a negative trend as time goes by.

We finally present random effects probit regressions by signal to provide further insights on the consistency of signals (see Table 12). The analysis confirms that

Table 10 Buyers' contract choices by consistency and (non) profit maximization for F_i contracts

	Consistent			Inconsistent		
	Profit maximizing (%)	Non profit maximizing (%)	Sum (%)	Profit maximizing (%)	Non profit maximizing (%)	Sum (%)
Baseline	71.96	24.87	96.83	3.17	0	3.17
Punishment	78.38	18.38	96.76	3.24	0	3.24
Enforced	70.30	27.72	98.02	1.98	0	1.98

Table 11 Buyers' contract choices by consistency and (non) profit maximization for menu of contracts

	Consistent			Inconsistent		
	Profit maximizing (%)	Non profit maximizing (%)	Sum (%)	Profit maximizing (%)	Non profit maximizing (%)	Sum (%)
Baseline	54.14	19.55	73.69	14.29	12.03	26.32
Punishment	53.60	10.40	64.00	21.60	14.40	36.00
Enforced	48.98	0	48.98	37.24	13.78	51.02

consistency in enforced self-selection treatment is significantly lower than in the baseline treatment. However, consistency between the baseline and the punishment treatment is not significantly different. Furthermore, the analysis shows that inconsistency in the enforced self-selection treatment is strongly driven by the strategically most relevant high cost signal.[11]

5.4 Cooperation, contract rejection, punishments, and rewards

Table 13 summarizes the number of contract offers by categories and treatments, how many of these contracts are tailored to the signal, the rejection rate, the average reward and the absolute number of rewards given in the respective contract offer state.

The surprising finding is that the frequency of simple contracts is rather low for the low and medium signals even though buyers cannot use these signals strategically. Since a large fraction of the low and medium signals are actually truthful, we would expect suppliers to offer more tailored contracts after low/medium signals than after high signals. Note, however, that the rejection rates of the tailored contracts for low/medium signals are extremely high. This indicates that buyers who truthfully signal low/medium cost prefer screening contracts to tailored contracts, because with screening they do not pass the entire cost advantage to the

[11] We thank an anonymous reviewer for pointing out that some subject may strategically reveal that their signals are unreliable in order to be offered a menu of contracts which avoids the strategic risk of too low rewards. Obviously, the best way to reveal inconsistency of signals is to always choose the high cost signal.

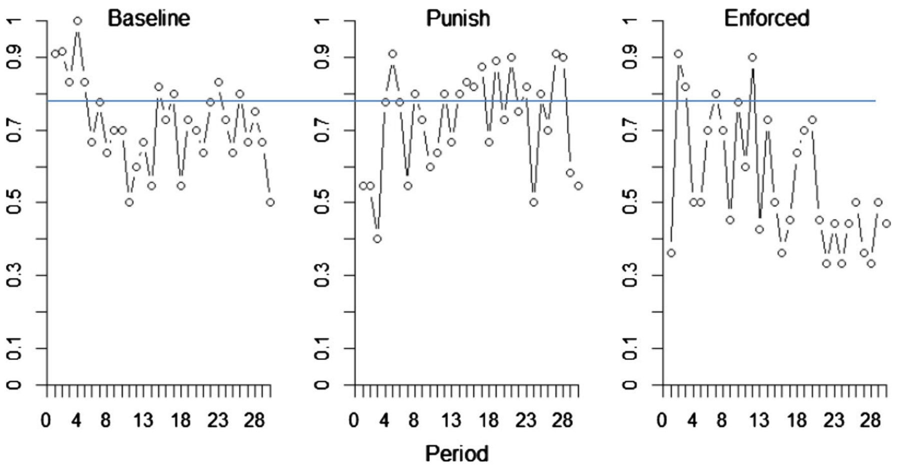

Fig. 5 Signal consistent contract choice

Table 12 Probit regression of treatment effects on consistency

	S_L	S_M	S_H	All
Enforced	0.09 (0.59)	−0.46 (0.39)	−0.78 (0.28)***	−0.59 (0.28)**
Punish	−0.11 (0.54)	−0.46 (0.39)	−0.08 (0.29)	−0.09 (0.28)
σ_i	0.83 (0.30)	0.51 (0.19)	0.58 (0.10)	0.61 (0.09)
LL	−65.19	−98.85	−346.20	−525.75
Prob > Chi^2	0.94	0.42	0.00	0.00
N	151	201	577	929

Marginal effects of probit regression, SD in parentheses (*** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$)

supplier. The high rejection rates expose the suppliers to a high degree of strategic risk, which they avoid by offering screening contracts after receiving a low/medium signal.

An alternative to using screening contracts for profit sharing in the case of low and medium cost would be a tailored contract combined with a sufficiently large reward payment. As we can see in Table 13 there are two reasons why this mechanism does not work in our experiment. First, the tailored contracts are frequently rejected making it impossible for suppliers to share profits by providing rewards. Second, even in those cases in which tailored contracts are implemented, the suppliers often provide rewards that are too low to establish win–win outcomes. In particular, given the contract choice F_L the minimum win–win reward is 43.77 and given the contract choice F_M the minimum win–win reward is 18.78.[12] As shown in Table 13, we see that even in those rounds in which rewards were provided, only 50 % of the rewards lead to a win–win outcome for the supply chain

[12] In the high cost state, any reward will make the buyer better off, since she is only left with her reservation profit for both contract types.

Table 13 Contract offers, rejection rates and awards by treatments and contract type

	Baseline	Punishment	Enforced
F_L			
No. of offers	38	16	18
Tailored to signal (%)	44.73	62.50	50.00
Rejection rate (%)	60.52	56.25	61.11
No. of rewards	13	7	6
No. of win–win rewards	11	0	3
Average rewards	49.23	40.00	38.80
F_M			
No. of offers	59	60	30
Tailored to signal (%)	42.37	45.00	70.00
Rejection rate (%)	45.76	38.33	50.00
No. of rewards	24	21	15
No. of win–win rewards	15	9	5
Average rewards	23.91	15.05	13.68
F_H			
No. of offers	102	124	66
Tailored to signal (%)	95.09	77.42	89.39
Rejection rate (%)	3.92	6.45	16.67
No. of rewards	33	65	33
Average rewards	11.96	8.66	6.82
Menu			
No. of offers	161	160	246
Rejection rate (%)	9.32	4.37	16.67
No. of rewards	60	50	90
Average rewards	7.25	7.31	7.38

parties. Even though we must reject the strictly payoff maximizing hypothesis 1e (binomial test, $p < 0.01$, one-sided), because we observe a substantial number of non-zero rewards, we also cannot detect a cooperation sustaining behavioral effect of rewards.

While we reject hypothesis 1e that punishments are not observed (binomial test, $p < 0.01$, one-sided), it is worth mentioning that the punishment option is only used rarely in. The punishment option was only used in about 6 % of the rounds (21 of 360) with an average punishment size of 25. Only 6 of 12 subjects in the punishment treatment used this option. Nevertheless, the observed differences between the punishment and the baseline treatment indicate that the mere presence of the punishment option has an impact on supply chain behavior.

5.5 Supply chain performance

We compare the observed supply chain performance to three benchmarks in Table 14. The first benchmark is the supply chain efficient outcome (first best). The second benchmark (Screening) is the equilibrium performance based on self-selection in screening contracts (second best). The third benchmark (F_{high}) captures the situation in which the supplier always offers F_H. This benchmark is meaningful because for this contract type the participation constraint of all buyers is satisfied and the variance of the supplier's payoff is zero. Moreover, this contract yields the highest expected supply chain profits among the F_i contracts.

We have calculated each benchmark given the actual realization of the holding cost. Conditioning the benchmarks on the actual realization gives more exact measures than the expected benchmark, because the effect of stochastic cost variations is neutralized. We exclude all observations in which the alternative supplier is chosen, since additional assumptions are required regarding the welfare effects of these choices.

Table 14 shows the average benchmark deviation in percentage of the supply chain optimal solution. The significance levels based on a Wilcoxon test show that we have a significantly lower performance than the supply chain optimum in all three treatments. Yet, the supply chain performance is not significantly lower than the second best benchmark in the enforced self-selection contract. Moreover, when comparing the observed performance to the F_{high} benchmark we see a significant improvement in both the punishment treatment and the enforced self-selection treatment. Thus, while we do not see a significant difference in the baseline treatment to the counterfactual benchmark in which the tailored contract is offered in every period, we see such an improvement in our control treatments.

We additionally performed OLS regression to test treatment effects on the deviation from the screening benchmark (see Table 15). We included dummies for medium and high holding cost realizations. As expected, we observe that the

Table 14 Average deviations from the respective benchmarks

	First best (%)	Second best (%)	F_{high} (%)
Baseline	−12.39***	−7.82**	3.02
Punish	−9.40***	−5.13**	6.74**
Enforced	−6.00***	−0.97	9.66***

Significance levels based on Wilcoxon-tests (two-sided) comparing observed performance to benchmarks

Table 15 OLS Regression–deviation from screening benchmark

	Coefficient
Constant	−12.27 (0.68) ***
Enforced	4.30 (0.80)***
Punishment	1.29 (0.78)*
Medium holding cost	10.15 (0.76)***
High holding cost	14.58 (0.80)***
Prob > F	0.00
R^2	0.32
N	886

SD in parentheses (*** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$)

average deviation of observed efficiency from the screening benchmark is significantly negative (constant = −12.27) in all treatments. Since tailored contracts are only efficiency enhancing for low and medium contracts, we observe significantly positive effects for these cost realizations. Most importantly, we observe a highly significant effect of enforcing self-selection, i.e., the supply chain performance is on average 4.3 points closer to the screening benchmark. We also observe a weakly significant effect of the punishment option, i.e., the performance is on average 1.3 points closer to the screening benchmark. Note, however, that the effect of the punishment option is on the border of significance. Future research—perhaps on variants of the game—will show how robust this effect actually is.

The analysis highlights, that there are basically two means for enhancing supply chain performance. First, if a punishment mechanism can be installed, then the buyer tend to send more trustworthy signals leading to a better alignment of trust and trustworthiness. Second, if self-selection is enforced, the supply chain performance can be enhanced because the effectiveness of the screening contract is improved due to the obliteration of indifference contract choices. Without indifference contract choices, the number of screening contract offers increases. Correspondingly, we observe significantly more screening contracts (enforced self-selection: 68 %; baseline: 45 %) and a significantly better performance in the enforced self-selection treatment than in the baseline treatment. Note that even if no buyer in the baseline treatment had chosen an indifference contract after being offered a screening contract, the performance would still be lower than the screening benchmark. The important issue here is that suppliers offer too few screening contracts in the baseline treatment due to strategic risk of indifference contract choices.

6 Discussion and managerial insights

The main message of our study is that supply chain environments with a high degree of strategic risk, such as the one we study in the baseline treatment, are detrimental to supply chain performance and should be avoided as much as possible. We examine two possible ways to reduce the degree of strategic risk. In our punishment treatment, the supplier has the possibility to punish buyers who exhibit apparently uncooperative behavior, i.e., if the high cost signal is send significantly too frequently or if there is an inconsistency between the signaled cost and the selected contract option. In our enforced self-selection treatment, we technically limit strategic risk by ruling out the buyers' possibility to choose indifference contracts, i.e., contracts that entail a small loss for the buyer but a substantial loss for the supplier.

Our results show that a reliable second best supply chain outcome can only be achieved with screening contracts if the incentives for self-selection are sufficiently high to reduce the behavioral variance in contract choices and, thus, the strategic risk that the suppliers face. Clearly, our experimental design, which provides the most extreme form of incentives for self-selection (i.e. the buyers have no other option to earn income, except by choosing the self-selection contract), cannot be used to assess the optimal threshold for the buyers' self-selection incentives. While theoretical approaches to robust contracts are already available (see Laffont and Martimort 2002, Chapter 9.8.1), the assessment of optimal incentive thresholds in the field will depend on specific market parameters and characteristics.

However, when managers can only choose between classical screening contracts with marginal incentives and tailored contracts, we observe that a punishment option increases trustworthiness of buyers' signals particularly for medium holding cost levels. In this environment, suppliers can tailor contracts to achieve a supply chain performance that is enhanced compared to a baseline without punishment.

In essence, our study shows that contract design should factor in behavioral mechanisms that reduce the strategic risk inherent in many supply chain interactions. On the one hand, allocation mechanisms that seem ineffective from a game-theoretic perspective (such as our punishment option) can reduce the strategic risk resulting from buyers' deceptive signals. The reduction of strategic risk leads to more coordinated supply chain outcomes. On the other hand, contractible incentives that are excessive from a game-theoretic perspective can also reduce behavioral variance, increasing suppliers' willingness to offer screening contracts and, thus, enhancing supply chain coordination. Hence, it seems that from a behavioral perspective the reduction of strategic risk is essential in a supply chain setting no matter which type of contract is considered.

There are several promising directions for future research. It seems interesting to further understand the variance in buyer's contract choice behavior. Theoretical explanations range from bounded rationality (see Basov 2009) to social preferences (see Voigt 2015). A better understanding of the underlying behavioral concepts might help to identify other forms of behavioral mechanisms that can reduce strategic risk. As an example, while we observe that a punishment option seems to be relevant for reducing strategic risk, it seems that the reward mechanism as it was provided in the present study (i.e., voluntary payment after contract choice) is not appropriate for incentivizing cooperative play, since we hardly identify rounds in which win–win communication took place. Thus, different forms of reward mechanisms might be analyzed in future work (e.g., discounts are given before the buyer chooses a contract). Finally, we should point out that the communication technology used in our experiment is rather crude. More elaborate communication, e.g. face-to-face negotiations, may enhance the supply chain outcomes, especially since there is evidence that the means of communication can affect outcomes in other settings (Valley et al. 1998; Brosig et al. 2003).

References

Andreoni J, Miller JH (1993) Rational cooperation in the finitely repeated prisoner's dilemma: experimental evidence. Econ J 103(418):570–585

Balliet D, Mulder LB, Lange Van, Paul AM (2011) Reward, punishment, and cooperation: a meta-analysis. Psychol Bull 137(4):594–615

Basov S (2009) Monopolistic screening with boundedly rational consumers. Econ Rec 85:S29–S33

Becker-Peth M, Katok E, Thonemann UW (2013) Designing buyback contracts for irrational but predictable newsvendors. Manage Sci 59(8):1800–1816

Berg J, Dickhaut J, McCabe K (1995) Trust, reciprocity, and social history. Games Econ Behav 10(1):122–142

Boles TL, Croson Rachel T A, Murnighan JK (2000) Deception and retribution in repeated ultimatum bargaining. Organ Behav Hum Decis Process 83(2):235–259

Bolton GE, Katok E (2008) Learning by doing in the newsvendor problem: a laboratory investigation of the role of experience and feedback. Manuf Serv Oper Manag 10(3):519–538

Brandts J, Charness G (2003) Truth or consequences: an experiment. Manage Sci 49(1):116–130

Brosig J, Weimann J, Ockenfels A (2003) The effect of communication media on cooperation. Ger Econ Rev 4(2):217–241

Cachon GP, Fisher M (2000) Supply chain inventory management and the value of shared information. Manage Sci 46(8):1032–1048

Corbett CJ, de Groote X (2000) Supplier's optimal quantity discount policy under asymmetric information. Manage Sci 46(3):444–450

Corbett CJ, Zhou D, Tang CS (2004) Designing supply contracts: contract type and information asymmetry. Manage Sci 50(4):550–559

Crawford V (1998) A survey of experiments on communication via cheap talk. J Econ Theory 78(2):286–298

Croson R, Boles T, Murnighan JK (2003) Cheap talk in bargaining experiments: lying and threats in ultimatum games. J Econ Behav Organ 51(2):143–159

Cui HT, Raju JS, Zhang ZJ (2007) Fairness and channel coordination. Manage Sci 53(8):1303–1314

Fischbacher U (2007) z-Tree: Zurich toolbox for ready-made economic experiments. Exp Econ 10(2):171–178

Fudenberg D, Tirole J (1992) Game theory, 2nd edn. MIT Press, Cambridge, Mass

Fudenberg D, Holmstrom B, Milgrom P (1990) Short-term contracts and long-term agency relationships. J Econ Theory 51(1):1–31

Goyal SK (1977) An integrated inventory model for a single supplier-single customer problem. Int J Prod Res 15(1):107–111

Greiner B (2015) Subject pool recruitment procedures: organizing experiments with ORSEE. J Econ Sci Assoc 1(1):114–125

Inderfurth K, Sadrieh A, Voigt G (2013) The impact of information sharing on supply chain performance under asymmetric information. Prod Oper Manag 22(2):410–425

Kahneman D, Knetsch JL, Thaler RH (1986) Fairness and the assumptions of economics. J Bus 59(4):285–300

Kalkanci B, Chen K-Y, Erhun F (2011) Contract complexity and performance under asymmetric demand information: an experimental evaluation. Manage Sci 57(4):689–704

Katok E, Wu DY (2009) Contracting in supply chains: a laboratory investigation. Manage Sci 55(12):1953–1968

Keser C, Paleologo GA (2004) Experimental investigation of supplier-retailer contracts: the wholesale price contract. Working paper

Kremer M, Minner S, Wassenhove Van, Luk N (2010) Do random errors explain newsvendor behavior? Manuf Ser Oper Manag 12(4):673–681

Kreps DM, Milgrom P, Roberts J, Wilson R (1982) Rational cooperation in the finitely repeated prisoners' dilemma. J Econ Theory 27(2):245–252

Laffont J-J, Martimort D (2002) The theory of incentives. The principal-agent model. Princeton University Press, Princeton

Lim N, Ho T-H (2007) Designing price contracts for boundedly rational customers: does the number of blocks matter? Mark Sci 26(3):312–326

Moore KR (1998) Trust and relationship commitment in logistics alliances: a buyer perspective. Int J Purch Mater Manag 34(4):24–37

Özer Ö, Zheng Y, Chen K-Y (2011) Trust in forecast information sharing. Manage Sci 57(6):1111–1137

Özer Ö, Zheng Y, Ren Y (2014) Trust, trustworthiness, and information sharing in supply chains bridging China and the United States. Manage Sci 60(10):2435–2460

Peeters R, Vorsatz M, Walzl M (2013) Truth, trust, and sanctions: on institutional selection in sender-receiver games. Scand J Econ 115(2):508–548

Sánchez-Pagés S, Vorsatz M (2007) An experimental study of truth-telling in a sender–receiver game. Games Econ Behav 61(1):86–112

Schweitzer ME, Cachon GP (2000) Decision bias in the newsvendor problem with a known demand distribution: experimental evidence. Manage Sci 46(3):404–420

Spiliotopoulou E, Donohue K, Gürbüz MÇ (2015) Information reliability in supply chains: the case of multiple retailers. Prod Oper Manag. doi:10.1111/poms.12418

Sucky E (2006) A bargaining model with asymmetric information for a single supplier–single buyer problem. Eur J Oper Res 171(2):516–535

Valley KL, Moag J, Bazerman MH (1998) A matter of trust: effects of communication on the efficiency and distribution of outcomes. J Econ Behav Organ 34(2):211–238

Voigt G (2011) Supply chain coordination in case of asymmetric information. Information sharing and contracting in a just-in-time environment. Springer, Heidelberg (Lecture notes in economics and mathematical systems, 650)

Voigt G (2015) Inequity aversion in a joint economic lot sizing environment with asymmetric holding cost information. J Oper Res Soc 66:554–569

Voigt G, Inderfurth K (2012) Supply chain coordination with information sharing in the presence of trust and trustworthiness. IIE Trans 44(8):637–654

Wu DY (2013) The impact of repeated interactions on supply chain contracts: a laboratory study. Int J Prod Econ 142(1):3–15

Zaheer A, McEvily B, Perrone V (1998) The strategic value of buyer-supplier relationships. Int J Purch Mater Manag 34(2):20–26